Mathematical Models of Trading

Thesis by
Angad Singh

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2020
Defended September 17, 2020

Angad Singh
ORCID: 0000-0001-6677-7013

ACKNOWLEDGEMENTS

First and foremost I would like to thank my advisor Jakša Cvitanić for all the support he has provided over the years. His advice and guidance were instrumental in helping me navigate between the worlds of mathematics, economics and finance. I would also like to give a very special thanks to Lawrence Jin for all the interest he took in me. I benefitted enormously from all our discussions.

Thank you to my other committee members Nets Katz and Omer Tamuz for being interested in my research. Thank you to the math department for allowing me the flexibility to pursue my interdisciplinary research interests. And thank you to the entire Caltech community for providing me with such a wonderful and intellectually stimulating environment to grow in. I will forever treasure my time spent here.

Last but not least, I would like to thank all my friends and family who have been there for me throughout the years. The most enormous of thanks go to my mother Simran Kalra, who has supported me unconditionally in all aspects of my life. Nothing would have been possible without her love and care. This thesis is dedicated to her.

ABSTRACT

This thesis presents a mathematical framework to model trading of financial assets on an exchange. The interaction between agents on the exchange is modeled as the Nash equilibrium of a demand schedule auction. The submission of demand schedules in the auction is meant to proxy for the submission of limit and market orders on an exchange. Chapter 1 considers this auction in a one-period setting, highlighting the importance of noisy flow for obtaining a unique Nash equilibrium.

Chapter 2 is the core of the thesis and considers the auction in a continuous time setting. Here the agents trading on the exchange have quadratic-type preferences, and in equilibrium they must clear an exogenously specified stream of market orders. Chapter 3 considers alternative and more realistic dynamics for the exogenous market orders. Chapter 4 endogenizes the market orders by considering an agent executing orders on behalf of noisy clients.. Chapter 5 considers the same model as in Chapter 2, except with a consumption based utility function for each agent.

TABLE OF CONTENTS

RELATED LITERATURE

The contents of this thesis relate to three strands of existing academic literature. The first is the theoretical economics literature on market microstructure. The second is the mathematical finance literature on optimal execution. The third is the finance literature on asset pricing.

The literature on market microstructure is fairly classical and considers various one-period equilibrium auction models. Thorough reviews are given in [OHa11] and [FPR13]. The demand schedule auction used in this thesis was first introduced in [Wil79]. [Kyl89] uses the demand schedule auction to study price impact as a consequence of adverse selection. The contents of Chapter 1 are similar to these papers, though the specific theorems proven are new. One major way in which Chapter 1 (and the rest of the thesis) deviates from [Kyl89] is that price impact is not a consequence of adverse selection. Instead, price impact compensates agents on the exchange for absorbing short term imbalances. This is in line with the idea of liquidity as the price of immediacy, as introduced in [GM88].

The literature on optimal execution considers dynamic partial equilibrium problems for an individual agent trading against a price impact function. This literature was pioneered in [AC01] and has since been extended in a variety of directions, as reviewed in [CJP15] and [Gué16]. The optimal response problems for the models in Chapters 2-5 are very much analogous to the optimization problems in this literature. However, this literature always takes the market price impact function as exogenously given, whereas in this thesis the price impact function is endogenous.

Asset pricing is the standard framework used in the academic finance community to model market prices. Thorough introductions to the subject are given in [Coc05] and [Bac17]. This literature does not explicitly model auctions, as the market microstructure literature does, but instead uses a general equilibrium approach. Furthermore, trading costs are usually ignored in these models. [GP16] and [Bou+18] do consider trading costs in asset pricing settings, but these costs are exogenously specified. Also, the trading costs in these models are not in the form of market price impact, but instead are transaction costs paid on top of market prices. Many of the modeling techniques in this thesis are inspired by these papers, and also by the continuous time CARA-Normal frameworks used in [CK93] and [Bar+15]

This thesis presents a dynamic auction based equilibrium framework to model asset prices. In particular, the endogenous level of prices consists of a price impact component that reflects the quantity being cleared on the exchange. When solving their individual optimization problems, all agents take in to account the market impact of their trades. This allows for a rich asset pricing model with microstructure details and with optimal execution problems for individuals. As such it bridges together the three strands of literature above.

Only two other papers in the literature consider dynamic models similar to the ones in this thesis, and both have a markedly different focus. The first is [SS16], which focuses on heterogenous agents and front running. One section of the paper considers a model of price impact between a large trader and a competitive fringe, which is similar to the model in Chapter 4. However, the model in Chapter 4 features multiple large traders (market makers) and considers an explicit optimization problem for the fringe (liquidity traders trading on behalf of clients). The second paper is [KOW18], which is focused on using private information and overconfidence to get around the classic no-trade theorem in finance.

Chapter 1

ONE PERIOD MODELS

This chapter formulates and discusses some one period models of auctions for shares[1]. The type of auction considered is typically referred to as a conditional uniform price auction, or as the demand schedule game. At the beginning of the auction, each agent submits a demand schedule $D : \mathbb{R} \to \mathbb{R}$, which is a commitment to purchase $D(p)$ shares if the price per share is p. The auctioneer then observes all the schedules, aggregates them into a total demand schedule, and chooses a price such that the auction clears. Agents receive shares based on their individual demand schedules, and finally they derive utility from the payoff associated to each share.

This type of auction is interesting because it's a good way to model agents trading through a limit order book. At a very high level, the decision to place market and limit orders amounts to deciding how aggressively to trade. A trader with a great sense of urgency might submit a large market order, which executes immediately, but is subject to walking the book. On the other hand, a more patient trader might scatter a few limit orders throughout the book, which may or may not execute, but will do so at a good price if they do. Finally, note then when an order executes it must, by definition, execute against another order placed by another agent. Thus agents are not just deciding how aggressively to push their own trading agendas, but also how aggressively to absorb the order flow of other traders.

Returning to the auction, note that selecting a demand schedule also amounts to deciding how aggressively to trade. A vertical demand schedule is equivalent to submitting a market order, as it demands a fixed quantity regardless of the price received. Tilting the slope of the demand schedule downwards is much like placing limit orders. It gives the agent the opportunity to trade at a good price, but this is not a guarantee because there must be another agent willing to accept the other side of

[1]For our purposes a share is simply a unit of an infinitely divisible good.

this trade. Finally, from a game theoretic point of view, when an agent takes others' schedules as given he is a monopsonist facing a supply curve. The suppliers are of course the other agents in the game, so forming an optimal response involves pushing ones own trading agenda while also deciding how much of others' order flow to absorb.

Hence the decision-making process of an agent in the demand schedule game captures many features of trading on a limit order book. An interesting property of the demand schedule game is that it features multiple equilibria when there is no uncertainty in the quantity to be cleared at the auction. From our point of view this is a realistic feature, as the quantity to be cleared on the exchange at any instant is uncertain. This first section below formulates and discusses this result. The second section presents a version of the model where agents have initial inventories, thereby formulating it as a Bayesian game. This version of the model connects better with the dynamic models to follow.

1.1 The Auction Model

There are $N \geq 2$ agents bidding for a total of S outstanding shares. Each share will, after the auction, provide a random payoff of $\tilde{\mu} \sim \mathcal{N}(\mu, \sigma^2)$. Agents are assumed to have CARA preferences over payoffs, with risk aversion parameter $\gamma > 0$. Thus if an agent purchase q shares in the auction for a price of p per share, his expected utility is $\mathbb{E}[-e^{-\gamma q(\tilde{\mu}-p)}]$. We will assume that agents can only submit affine and decreasing demand schedules[2]. So a demand schedule is characterized by a pair $(a, b) \in \mathbb{R} \times (0, \infty)$, corresponding to the commitment to purchase $a - bp$ shares when the price per share is p.[3]

In addition to the N agents there are noise traders who have a random perfectly

[2]The affine assumption is not strictly necessary, as an optimal response to a profile of affine schedules is also given by an affine schedule.

[3]$b = \infty$ corresponds to placing an order that specifies a price but no quantity. Such an order is not possible on an exchange, and thus is not allowed here. Technically we should allow $b = 0$, since it corresponds to placing a market order. However, we only consider symmetric equilibria below, and $b = 0$ can trivially never be such an equilibrium. Thus for the sake of exposition $b = 0$ is omitted from the outset.

inelastic demand for $-\tilde{u}$ shares. Thus the total number of shares the N agents must clear in the auction is $S + \tilde{u}$. By adjusting the mean of $\tilde{u}$ we can simply assume that the total number of shares being auctioned is random and equal to $\tilde{u}$. We assume that $\tilde{u}$ is independent of $\tilde{\mu}$, but make no other assumptions about its distribution.

Given demand schedules $(a^1, b^1), \cdots, (a^N, b^N)$, the auction price p and the shares $q^1, \cdots, q^N$ bought by each agent are implicitly given by

$$a^n - b^n p = q^n \tag{1.1}$$

$$q^1 + \cdots + q^N = \tilde{u}. \tag{1.2}$$

This describes a game in normal form, and we now proceed to study its Nash equilibria. We will focus only on symmetric equilibria, that is on equilibria where all agents submit the same demand schedule. Note that in a symmetric equilibrium all agents purchase the same number of shares in the auction, i.e. $q^1 = \cdots = q^N = \frac{\tilde{u}}{N}$.

Theorem 1.1.1. *Fix exogenous parameters $N \geq 2, \gamma, \sigma > 0$ and $\mu \in \mathbb{R}$.*

If $\tilde{u}$ is degenerate, i.e. $\tilde{u} = u \in \mathbb{R}$ a.s., then there is a one-to-one correspondence between symmetric equilibria and $\lambda > 0$. In equilibrium the price is

$$p = \mu - \frac{\gamma\sigma^2}{N}u - \frac{\lambda}{N}u. \tag{1.3}$$

If $\tilde{u}$ is non-degenerate and $N \geq 3$, then there is a unique symmetric equilibrium with price

$$p = \mu - \frac{\gamma\sigma^2}{N}\tilde{u} - \frac{\gamma\sigma^2}{N(N-2)}\tilde{u}. \tag{1.4}$$

If $\tilde{u}$ is non-degenerate and $N = 2$ then a symmetric equilibrium does not exist.

Proof. Fix a strategy $(a, b) \in \mathbb{R} \times (0, \infty)$ to be played by all but one agent, and consider the optimal response problem faced by the remaining agent. If this agent plays strategy $(\alpha, \beta) \in \mathbb{R} \times (0, \infty)$, then his expected utility will be $\mathbb{E}[-e^{-\gamma q(\tilde{\mu}-p)}]$ with p and q given implicitly by

$$q = \alpha - \beta p \tag{1.5}$$

$$\frac{\tilde{u} - q}{N - 1} = a - bp. \tag{1.6}$$

Rearranging we obtain

$$p = F - \lambda\tilde{u} + \lambda q \tag{1.7}$$

$$q = \left[\frac{\alpha}{1+\lambda\beta} - \frac{\beta}{1+\lambda\beta}F\right] + \frac{\lambda\beta}{1+\lambda\beta}\tilde{u}, \tag{1.8}$$

where $F := \frac{a}{b}$ and $\lambda := \frac{1}{b(N-1)}$.

We note a few things about these equations. Firstly, the parameters F and λ characterize the symmetric profile given by (a, b) and are independent of the remaining agent's demand schedule. Secondly, equation (7) tells us that the price the remaining agent receives is uniquely determined by the number of shares he receives. Thus the agent is indifferent between schedules that lead to the same quantity of shares. Thirdly, from equation (8) we see that choosing $(\alpha, \beta) \in \mathbb{R} \times (0, \infty)$ amounts to choosing to receive the quantity $q \in V$ where

$$V := \{A + B\tilde{u} : (A, B) \in \mathbb{R} \times (0, 1)\}. \tag{1.9}$$

Fourthly, if the agent submits the schedule (a, b) then he receives the quantity $q = \frac{\tilde{u}}{N}$.

All this goes to show that in forming an optimal response, the remaining agent can maximize directly over $q \in V$, with the price given by (7), and in equilibrium the maximum must be attained at $\frac{\tilde{u}}{N}$. Thus symmetric equilibria correspond to $F \in \mathbb{R}$ and $\lambda > 0$ such that

$$\frac{\tilde{u}}{N} \in \arg\max_{q \in V} \mathbb{E}\left[-e^{-\gamma q\left(\tilde{\mu} - F + \lambda\tilde{u} - \lambda q\right)}\right], \tag{1.10}$$

and in equilibrium the price is

$$p = F - \frac{N-1}{N}\lambda\tilde{u}. \tag{1.11}$$

Now, note that $\tilde{u}$ is independent of $\tilde{\mu}$ and every $q \in V$ is $\tilde{u}$ measurable. Since the MGF of $\tilde{\mu}$ is $\mathbb{E}[e^{t\tilde{\mu}}] = e^{\mu t + \frac{\sigma^2}{2}t^2}$, we can compute the expectation in (10) for any $q \in V$ as

$$\mathbb{E}\left[-e^{-\gamma q\left(\mu - F + \lambda\tilde{u} - (\lambda + \frac{\gamma\sigma^2}{2})q\right)}\right]. \tag{1.12}$$

For any $F \in \mathbb{R}$ and $\lambda > 0$, the function of q in (12) is strictly concave over the convex set V. Thus the argmax in (10) consists of at most one point. Furthermore, for any $F \in \mathbb{R}$, $\lambda > 0$ and $u \in \mathbb{R}$, the much relaxed problem of maximizing the exponent in (12)

$$\max_{q \in \mathbb{R}} q(\mu - F + \lambda u - \left(\lambda + \frac{\gamma\sigma^2}{2}\right)q)$$

has unique solution $\hat{q} = \frac{\mu - F}{2\lambda + \gamma\sigma^2} + \frac{\lambda}{2\lambda + \gamma\sigma^2}u$.

It follows that (10) holds if and only if

$$\frac{\tilde{u}}{N} = \frac{\mu - F}{2\lambda + \gamma\sigma^2} + \frac{\lambda}{2\lambda + \gamma\sigma^2}\tilde{u}. \tag{1.13}$$

Thus symmetric equilibria correspond to $F \in \mathbb{R}$ and $\lambda > 0$ satisfying (13), and in equilibrium the price is given by (11).

Now, if $\tilde{u}$ is degenerate and equal to $u \in \mathbb{R}$ a.s., then (13) simply reads

$$F = \mu - \frac{\gamma\sigma^2}{N}u + \frac{N-2}{N}\lambda u. \tag{1.14}$$

Thus there a one-to-one correspondence between symmetric equilibria and $\lambda > 0$, with F given by (14). The first statement in the theorem follows by using (14) to substitute for F in (11).

Next suppose that $\tilde{u}$ is non-degenerate. Then (14) holds if and only if $F = \mu$ and $\lambda > 0$ satisfies

$$(N-2)\lambda = \gamma\sigma^2. \tag{1.15}$$

If $N = 2$ then no $\lambda > 0$ can satisfy (15), so a symmetric equilibrium does not exist, proving the last statement in the theorem. If $N \geq 3$ then the unique $\lambda > 0$ satisfying (15) is $\lambda = \frac{\gamma\sigma^2}{N-2}$, and so there is a unique symmetric equilibrium. The second statement in the theorem now follows by plugging in $F = \mu$ and $\lambda = \frac{\gamma\sigma^2}{N-2}$ in (11). □

Discussion

The idea in the proof is to characterize symmetric profiles in terms of the parameters F and λ of the induced the optimal response problem. The optimal response problem is to choose an expected utility maximizing quantity on the linear supply curve (7), which has intercept $F - \lambda\tilde{u}$ and slope $\lambda > 0$. This is a concave maximization problem with a unique solution. The symmetric equilibrium condition on F and λ is that this solution is $\frac{\tilde{u}}{N}$. In the non-degenerate case this condition uniquely determines F and λ whereas in the degenerate case it only specifies F as a function of λ.

What happens in the degenerate case is that the agents are able to form an agreement to misprice the asset and then take equal shares of the profit. Consider for example the case when $\tilde{u} = 1$, so a unit share is being cleared at the auction. Then (3) states that the equilibrium price can take on any value below $\mu - \frac{\gamma\sigma^2}{N}$, which is the price the asset would trade at in a competitive equilibrium. Thus we see that the asset is being priced relatively low, and since each agent takes $\frac{1}{N}$ shares, they split the profits equally.

Since the game is non-cooperative, in order to form an agreement the agents must have a way to prevent others from taking more than an equal share of the profits. The key point is that all agents submit entire demand schedules, which specify what the price must be contingent on the quantity the agent receives. So if one agent were to take more than $\frac{1}{N}$ shares, some other agents would receive less than $\frac{1}{N}$ shares, and this would cause the price to move, thus dissuading any one agent from trying to take extra shares in the first place.

The amount by which the price would move if an agent took more then $\frac{1}{N}$ shares is governed by the parameter λ, which corresponds to the quantity elasticity b of the equilibrium demand schedule. The lower the equilibrium price p, the more of an incentive an agent has to acquire more than $\frac{1}{N}$ shares, and thus the higher λ needs to be to prevent the agent from doing so. (3) says exactly that low equilibrium prices correspond to high values of λ.

The problem in the degenerate case is that agents suffer no cost from being quantity elastic, since there will be no surprise trades in equilibrium. Hence λ can take on any positive value in equilibrium. In the non-degenerate case, agents suffer costs from being quantity elastic in equilibrium, because there is uncertainty in the quantity to be cleared. These costs manifest in how the parameter λ effects the uncertainty of equilibrium prices. Since agents care about the uncertainty of prices, this pins down the unique equilibrium value of λ as $\frac{\gamma\sigma^2}{N-2}$.

The coefficient of $\tilde{u}$ in (4) is $\frac{N-1}{N(N-2)}\gamma\sigma^2$, which is the price impact of the noise traders' order. If the noise traders sell ϵ more shares, so the realization of $\tilde{u}$ is ϵ higher, then the equilibrium price is $\frac{N-1}{N(N-2)}\gamma\sigma^2\epsilon$ lower. Thus orders walk the book: the larger an order, the lower the transaction price.

The decomposition of price impact into the two terms is motivated by considering the competitive limit as $N \longrightarrow \infty$ and $\frac{\gamma}{N}$ is held fixed. In the limit the second term vanishes and only the first remains. Thus $\mu - \frac{\gamma\sigma^2}{N}\tilde{u}$ is the competitive benchmark, and the second term is the deviation due to imperfect competition. As in the competitive case, the term $\gamma\sigma^2\frac{\tilde{u}}{N}$ is the risk compensation each agent requires to take the equilibrium exposure of $\frac{\tilde{u}}{N}$.

The interpretation is that price impact arises for two reasons in the model. Firstly to make sure agents are appropriately compensated for bearing risk, and secondly because agents have market power. The first reason persists even in the competitive limit, and as result price impact does not vanish in the limit. This will be a recurring theme throughout the thesis.

1.2 The Auction Model with Inventories

The continuous time models considered in the subsequent chapters essentially consist of the auction above at each instant, with the addition of certain state variables that need to be carried from instant to instant. The state variables are the existing inventories of shares that agents have accumulated from trading in the past. This section introduces these state variables in a static setting as types, thus generalizing the model above to a Bayesian game.

In addition to forming a tighter connection with the continuous time models to follow, the rephrased model in this section has two other appealing features. Firstly, in the previous section the total number of outstanding shares plays no distinct role[4] from the noise traders' order. This is perhaps counterintuitive, as the noise traders' order should have price impact, whereas the total number of outstanding shares should be a fixed component of the price. In this section the total number of outstanding shares will show up as a fixed component of the price. Secondly, in the previous section all agents purchased the same number of shares in equilibrium. In this section their purchases will be heterogenous.

The model is exactly as before except that each of the N agents starts out with an existing inventory of $X^n \in \mathbb{R}$ shares. Thus if agent n purchases q^n shares in the auction for a price of p per share, then his expected utility is

$$\mathbb{E}\Big[- e^{-\gamma\big((X^n+q^n)\tilde{\mu}-pq^n\big)}\Big]. \tag{1.16}$$

The noise traders start out with zero shares, and the total number of outstanding shares is S, so $\sum_{n=1}^{N} X^n = S$.

More formally, we work on a probability space with a single objective probability measure. There are $N+2$ real-valued random variables defined on this probability space: $\tilde{\mu}$, $\tilde{u}$, and $X^1, \cdots, X^N$. There are exogenous constants $\mu, S \in \mathbb{R}$ and $\sigma > 0$ such that $\tilde{\mu} \sim \mathcal{N}(\mu, \sigma^2)$ and $\sum_{n=1}^{N} X^n = S$. Furthermore, $\tilde{\mu}$ and $\tilde{u}$ are independent of each other as well as $X^1, \cdots, X^N$.

The type (or private information) of agent n is X^n. A strategy is a measurable function mapping the realization of an agent's type to a choice of demand schedule. As before demand schedules are restricted to be affine and strictly decreasing, so a strategy for agent n is a measurable mapping $(a^n, b^n) : \mathbb{R} \to \mathbb{R} \times (0, \infty)$, $X \mapsto (a^n(X), b^n(X))$. Given a strategy profile and a realization of $(X^1, \cdots, X^N)$, prices and quantities are determined from (1) and (2) with $a^n = a^n(X^n)$ and $b^n = b^n(X^n)$.

This completes the description of the model as a Bayesian game. We will be interested in identifying Bayesian Nash equilibria[5] in this game, but we will focus on equilibria that have a very specific structure.

[4] In the previous section, the total number of outstanding shares was absorbed into the mean of $\tilde{u}$.

[5] What we call a Bayesian Nash equilibrium is sometimes called a strong Bayesian Nash

Definition 1.2.1. A strategy $s : \mathbb{R} \to \mathbb{R} \times (0, \infty)$ is called **linear** if there exist constants $a, \xi \in \mathbb{R}$ and $b \in (0, \infty)$ such that

$$s(X) = \big(aX + \xi, b\big) \quad \forall X \in \mathbb{R}.$$

Our focus will be on linear symmetric equilibria, that is on equilibria where all agents play the same linear strategy.

Theorem 1.2.2. *Fix exogenous parameters $N \geq 3$, $\gamma, \sigma > 0$ and $\mu, S \in \mathbb{R}$.*

If $\tilde{u}$ is non-degenerate then there is a unique linear symmetric equilibrium. In equilibrium, the price is

$$p = \mu - \frac{\gamma\sigma^2}{N} S - \frac{N-1}{N(N-2)} \gamma\sigma^2 \tilde{u} \tag{1.17}$$

and the quantities purchased by each agent are

$$q^n = -\frac{N-2}{N-1}(X^n - \frac{S}{N}) + \frac{\tilde{u}}{N}. \tag{1.18}$$

Proof. Fix a linear strategy given by $a, \xi \in \mathbb{R}$ and $b \in (0, \infty)$ to be played by all but one agent, and consider the optimal response problem faced by the remaining agent. Suppose the remaining agent plays the strategy $\mathbb{R} \to \mathbb{R} \times (0, \infty)$, $x \mapsto (\alpha(x), \beta(x))$. If the agent's initial inventory is X, then his expected utility is $\mathbb{E}[-e^{-\gamma((X+q)\tilde{\mu} - pq)}]$, where p and q are given implicitly by

$$\frac{a}{N-1}(S - X) + \xi - bp = \frac{\tilde{u} - q}{N-1} \tag{1.19}$$

$$\alpha(X) - \beta(X)p = q. \tag{1.20}$$

Rearranging we obtain

$$p = F + C(S - X) - \lambda\tilde{u} + \lambda q \tag{1.21}$$

$$q = \left[\frac{\alpha(X)}{1 + \lambda\beta(X)} - \frac{\beta(X)}{1 + \lambda\beta(X)}\Big(F + C(S - X)\Big)\right] + \frac{\lambda\beta(X)}{1 + \lambda\beta(X)}\tilde{u}, \tag{1.22}$$

equilibrium. We require agent n to choose a strategy that maximizes (16) *conditional* on X^n for *every realization* of X^n. This is in contrast to the weaker requirement of choosing a strategy that just maximizes (16), which *averages over realizations* of X^n.

where $F := \frac{\xi}{b}$, $C := \frac{a}{b(N-1)}$ and $\lambda := \frac{1}{b(N-1)}$.

We note a few things about theses equations. Firstly, the parameters F, C and λ characterize the symmetric profile given by a, ξ and b, and they are independent of the remaining agent's demand schedule. Secondly, equation (21) tells us that if we hold the remaining agent's initial inventory fixed, then the price the remaining agent receives is uniquely determined by the quantity he trades. Thus the agent is indifferent between demand schedules that lead to the same quantity of shares. Thirdly, from equation (22) we see that choosing a strategy (α, β) amounts to choosing functions $A : \mathbb{R} \longrightarrow \mathbb{R}$ and $B : \mathbb{R} \longrightarrow (0, 1)$ such that the agent's traded quantity is $q = A(X) + B(X)\tilde{u}$. Fourthly, if the agent uses the linear strategy $(aX + \xi, b)$ then his traded quantity is $q = \frac{C}{\lambda}\big(X - \frac{S}{N}\big) + \frac{\tilde{u}}{N}$.

All this goes to show that linear symmetric equilibria correspond to $F, C \in \mathbb{R}$ and $\lambda > 0$ such that

$$\frac{C}{\lambda}\big(x - \frac{S}{N}\big) + \frac{\tilde{u}}{N} \in \arg\max_{q \in V} \mathbb{E}\Big[- e^{-\gamma\big(x\tilde{\mu} + q\big(\tilde{\mu} - F - C(S-x) + \lambda\tilde{u} - \lambda q\big)\big)}\Big] \quad \forall x \in \mathbb{R}, \tag{1.23}$$

where V is as in (9). In equilibrium the price is

$$p = F + \frac{N-1}{N} CS - \frac{N-1}{N}\lambda\tilde{u}, \tag{1.24}$$

and agents' trades are

$$q^n = \frac{C}{\lambda}\big(X^n - \frac{S}{N}\big) + \frac{\tilde{u}}{N}. \tag{1.25}$$

Now, note that $\tilde{u}$ is independent of $\tilde{\mu}$ and every $q \in V$ is $\tilde{u}$ measurable. Since the MGF of $\tilde{\mu}$ is $\mathbb{E}[e^{t\tilde{\mu}}] = e^{\mu t + \frac{\sigma^2}{2}t^2}$, we can compute the expectation in (23) for any $q \in V$ and $x \in \mathbb{R}$ as

$$e^{-\gamma\mu x}\mathbb{E}\Big[- e^{-\gamma\big(q\big(\mu - F - C(S-x) + \lambda\tilde{u} - \lambda q\big) - \frac{\gamma\sigma^2}{2}(x+q)^2\big)}\Big]. \tag{1.26}$$

For any $F, C, x \in \mathbb{R}$ and $\lambda > 0$, the function of q in (26) is strictly concave over the convex set V. Thus the argmax in (23) consists of at most one point. Furthermore, for any $F, C, x \in \mathbb{R}$, $\lambda > 0$ and $u \in \mathbb{R}$, the much relaxed problem of maximizing the exponent inside the expectation in (26)

$$\max_{q\in\mathbb{R}} q\big(\mu - F - C(S-x) + \lambda\tilde{u} - \lambda q\big) - \frac{\gamma\sigma^2}{2}(x+q)^2$$

has unique solution $\hat{q} = \frac{\mu-F-CS}{2\lambda+\gamma\sigma^2} + \frac{C-\gamma\sigma^2}{2\lambda+\gamma\sigma^2}x + \frac{\lambda}{2\lambda+\gamma\sigma^2}u$.

It follows that (23) holds if and only if

$$\frac{C}{\lambda}\big(x - \frac{S}{N}\big) + \frac{\tilde{u}}{N} = \frac{\mu - F - CS}{2\lambda+\gamma\sigma^2} + \frac{C-\gamma\sigma^2}{2\lambda+\gamma\sigma^2}x + \frac{\lambda}{2\lambda+\gamma\sigma^2}\tilde{u} \quad \forall x \in \mathbb{R}. \qquad (1.27)$$

Since $\tilde{u}$ is non-degenerate, (27) holds if and only if $F = \mu, C = -\frac{\gamma\sigma^2}{N-1}$, and $\lambda = \frac{\gamma\sigma^2}{N-2}$. The theorem now follows by plugging these values is (24) and (25). □

Discussion

The proof is similar to the one in the previous section, with the idea being to characterize symmetric profiles in terms of the parameters F, C and λ of the induced optimal response problem. The additional parameter C governs how the intercept of the supply curve in the optimal response problem depends on the optimizing agent's initial inventory. More specifically, C captures the dependence of the intercept on the sum of all other agents' inventories, which the optimizing agent can compute by subtracting his own inventory from the total number of outstanding shares, i.e. $S - X$.

The supply curve represents the prices at which the other agents are willing to clear the joint order of the noise traders and the optimizing agent. These prices must depend on the preexisting exposures of the remaining agents, hence the presence of the parameter C. For symmetric profiles, the others' exposure can be aggregated instead of considering individual exposures, which greatly simplifies the problem.

The constant term in the equilibrium price is $\mu - \frac{\gamma\sigma^2}{N}S$, as opposed to just μ in the previous theorem. Thus there is a constant discount in the price reflecting the total number of outstanding shares. Intuitively this discount appears here because the agents are already in possession of S shares prior to the auction, whereas in the previous section they initially posses no shares. The coefficient of $\tilde{u}$ in the equilibrium price is $\frac{N-1}{N(N-2)}\gamma\sigma^2$, exactly as in the previous section.

A unified way to write the equilibrium price in the two theorems is in terms of the aggregate inventory of the agents *after* the auction, denoted S_{post}. In the first section's model we have $S_{post} = \tilde{u}$ and in the second section we have $S_{post} = S + \tilde{u}$. In both sections, the equilibrium price is

$$p = \mu - \frac{\gamma\sigma^2}{N} S_{post} - \frac{\gamma\sigma^2}{N(N-2)}\tilde{u}. \tag{1.28}$$

The first two terms here are the competitive benchmark, and the last term is the deviation due to imperfect competition.

At first glance it might seem surprising and counterintuitive that the deviation due to imperfect competition depends on $\tilde{u}$ and not S_{post}. For example, if $S_{post} > 0$ and $\tilde{u} < 0$, then the agents are in aggregate long the asset, but the price is high relative to the competitive benchmark. The deviation due to imperfect competition should always favor the agents, so one might expect it to make the price low when they are going long and high when they are going short. However this reasoning is flawed because the price in (28) is not the price at which the agents enter their aggregate position of S_{post}. It is merely the price at which the agents shift their aggregate position from S to S_{post}. Said another way, (28) is not the denominator in the agents' aggregate return, and as such the low/high long/short reasoning does not apply.

The logic behind (28) is that the price is low when the noise traders are selling, $\tilde{u} > 0$, and high when they are buying, $\tilde{u} < 0$. Thus, roughly speaking, the noise traders are always "getting ripped off." This can be made more precise by recalling the analogy between the auction and a limit order book. Based on this analogy, the price in (17) can be interpreted as saying that the mid-price is $\mu - \frac{\gamma\sigma^2}{N}S$, and orders walk the book at a rate of $\frac{\gamma\sigma^2}{N(N-2)}$ per share.

Now, suppose the auction is repeated an instant later (prior to the realization of payoffs). Since the aggregate inventory of the agents will be S_{post} an instant later, the mid-price will be $\mu - \frac{\gamma\sigma^2}{N}S_{post}$. So, (28) says that if the agents buy in the first auction, $\tilde{u} > 0$, then they do so at a price lower than the mid-price in the second auction. Similarly if they sell in the first auction, $\tilde{u} < 0$, then they do so at a price higher than the mid-price in the second auction. Thus agents always trade in the first auction at prices that are favorable relative to the mid-price in the second auction. In particular, if an agent were to unwind the position acquired in the first auction with a limit order[6] in the second auction, then the roundtrip trade would earn positive profits. This is the sense in which the deviation due to imperfect competition always favors the agents. This logic will reappear more clearly in future chapters, where we look at dynamic models and the auction truly is repeated instant after instant.

From (18) we see that the agents trade heterogenous quantities of shares in the auction, unlike in the previous section where they all traded $\frac{\tilde{u}}{N}$. Since the agents must clear $\tilde{u}$ in the auction, it follow that the average number of shares bought by each agent is $\frac{\tilde{u}}{N}$, but some agents might buy more and some less. The total initial inventory of the agents is S, so the average inventory held by each agent is $\frac{S}{N}$. (18) says that the agents with above average inventories buy less, and those with below average inventories buy more.

A more precise way to understand the traded quantities is in terms of the Pareto optimality of the inventory distribution before and after the auction. Since all agents are identical, it would be Pareto optimal for them to hold $\frac{S}{N}$ shares before the auction and $\frac{S_{post}}{N}$ after the auction. Individual inventories after the auction are $X^n_{post} := X^n + q^n$ and from (18) it follows that

$$X^n_{post} - \frac{S_{post}}{N} = \frac{1}{N-1}(X^n - \frac{S}{N}). \tag{1.29}$$

Thus (29) says that trading in the auction moves inventories closer to efficiency by a factor of $\frac{1}{N-1}$. When N takes its smallest value of 3, the agents only move halfway towards efficiency, whereas they move entirely towards efficiency in the limit as

[6] Technically a market order could also work in the second auction if it does not suffer too much price impact. The point is to unwind the position in the second auction at a favorable price relative to (28). Since the mid-price in the second auction will be $\mu - \frac{\gamma\sigma^2}{N} S_{post}$, the price will always be favorable if using a limit order, and for a market order it depends on the price impact. In the dynamic model, equilibrium price impact will be constant over time and using market orders will not work for the agents. However noise trader flow will mean revert, so the agents will eventually be able to unwind their positions using limit orders, and this will earn positive profits.

$N \to \infty$. Thus imperfect competition among the agents results in imperfect risk sharing.

Chapter 2

CONTINUOUS TIME MODEL

This chapter presents a continuous time model of trading on an exchange. The model is set on an infinite horizon, and each instant in time consists of the auction from chapter one. Traders hold cash and shares, and these fluctuate over time based on the outcomes of the auctions. As in Chapter 1 there are two types of traders: market makers, endogenous, and liquidity traders, exogenous.

In this chapter the exogenous dynamics of the liquidity traders orders are taken to be as simple as possible: an Ornstein-Uhlenbeck process. This allows us to focus on the mechanics of solving the model in a relatively simple setting. These liquidity trader dynamics have some drawbacks, however. In particular, there is no mean reversion in the liquidity traders' shareholdings, and thus there is also no clear notion of round trip trades by the market makers. Chapter 3 considers alternative dynamics to deal with these issues.

The first section of the chapter formulates the model as a stochastic differential game and defines the equilibrium concept to be considered. The second section phrases the optimal response problem for an individual as a standard stochastic control problem. The third section provides a complete closed-form characterization of equilibrium by solving this problem. The last section discusses the economics behind the endogenous price and trading processes that arise in the models equilibrium.

2.1 The Model

Fix a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ equipped with two independent Brownian motions, $\{B_t^D\}$ and $\{B_t^N\}$, and satisfying the usual conditions. We consider a market on an infinite horizon where shares of a zero net supply risky asset are traded for cash. Cash is in infinitely elastic supply, earns no interest, and is the numeraire. The market is populated by two types of traders: $N \in \mathbb{N}$ market makers and a collection of liquidity traders. Each market maker starts out at time 0 holding X_0^n shares of the asset, $n = 1, \cdots, N$. The liquidity traders start out with a collective shareholding of $-S_0$, which must satisfy $S_0 = X_0^1 + \cdots + X_0^N$ since the asset is in zero net supply.

At each instant in time the traders transact with one another at a uniform price p_t. Trading occurs smoothly, meaning that each trader has a trading rate, which is the time derivative of his current shareholdings.[1] Trading rates and the trading price at each instant in time are determined by a demand schedule auction between the traders. Thus at time t each trader submits an affine[2] demand schedule of the form $q = u_t - v_t p$. This is a commitment to trade at rate $u_t - v_t p$ at time t if the trading price is p. We denote this demand schedule by (u_t, v_t). The process $\{(u_t, v_t)\}$ is required to be progressively measurable, though we will place more stringent conditions on it below.

The liquidity traders' collective demand schedule at time t is exogenously give as $(-\mathcal{N}_t, 0)$, where

$$d\mathcal{N}_t = -\psi \mathcal{N}_t dt + \sigma_{\mathcal{N}} dB_t^{\mathcal{N}} \tag{2.1}$$

and $\psi, \sigma_{\mathcal{N}} > 0$. Thus the liquidty traders' demand schedule is a vertical line through the point $-\mathcal{N}_t$, which we interpret as a market order to sell $\mathcal{N}_t dt$ shares over the time interval $[t, t + dt]$. The liquidity traders' motives to trade are not modeled; it is simply assumed that they results in the trajectory of market order trading rates $\{-\mathcal{N}_t\}$.

The market makers, on the other hand, submit demand schedules in order to maximize certain objectives. Denote the market makers' trading rates by q_t^n, so that their inventories evolve according to

$$dX_t^n = q_t^n dt. \tag{2.2}$$

[1] Shareholdings will also be referred to as inventories in what follows.

[2] It is actually not necessary to assume that agents can only submit affine schedules, as we will see below. For the class of equilibria we consider, when forming an optimal response an agent can achieve any trading rate via an affine demand schedule. Thus even if agents could submit arbitrary schedules, there would still be an equilibrium where they all submit affine schedules. Of course, there may also be other equilibria where agents submit more exotic schedules.

If the market makers submit the demand schedule processes $\{(\frac{\alpha_t^n}{\beta_t^n}, \frac{1}{\beta_t^n})\}$, then their trading rates q_t^n and the trading price p_t are determined implicitly by

$$\alpha_t^n - \beta_t^n q_t^n = p_t \qquad \forall n = 1, \cdots, N \tag{2.3}$$

$$q_t^1 + \cdots + q_t^N = \mathcal{N}_t. \tag{2.4}$$

Note that the price, trading rate, and inventory processes all depend on the choice of demand schedule processes. This dependence is suppressed in the notation.

Each market maker has cash holding M_t^n which evolve as a result of trading according to $dM_t^n = -q_t^n p_t dt$. The market makers are also assumed at time t to have a common exogenous valuation of the asset as D_t, where

$$dD_t = \mu dt + \sigma_D dB_t^D \tag{2.5}$$

and μ, $\sigma_D > 0$ are fixed constants. Thus each marker maker values his book, consisting of joint holdings in cash and the asset, as $W_t^n = X_t^n D_t + M_t^n$. This is the market maker's wealth, computed by valuing shareholdings at D_t.

Each market maker chooses his demand schedule to maximize the objective

$$\mathbb{E}\Big[\int_0^\infty e^{-\rho t}\Big(dW_t^n - \frac{\gamma}{2} d\langle W^n\rangle_t\Big)\Big], \tag{2.6}$$

where ρ, $\gamma > 0$. Here $\langle W^n\rangle_t$ is the quadratic variation of the market makers wealth, and so $d\langle W^n\rangle_t$ can be thought of as the variance of instantaneous wealth changes. Thus the integrand in (6) can be interpreted as a mean-variance utility flow from instantaneous returns, which means that (6) embodies myopic[3] mean-variance preferences over returns.

One can compute the objective function in (6) as

$$\mathbb{E}\Big[\int_0^\infty e^{-\rho t}\Big(- q_t^n(p_t - D_t) + \mu X_t^n - \frac{\gamma\sigma_D^2}{2}(X_t^n)^2\Big)dt\Big]. \tag{2.7}$$

Thus the market makers want to buy, $q_t^n > 0$, when the price is below their valuation, $p_t - D_t < 0$, and vice versa. Furthermore, they enjoy holding inventory to the extent

[3]Myopic because the utility flow comes from instantaneous returns.

that valuations grow on average, $\mu > 0$, and they are averse to holding inventory to the extent that valuations are volatile, $\sigma_D^2 > 0$.

Barring technicalities, this completes the description of the model as a stochastic differential game between the N market makers. Indeed, the control process[4] of each market maker is $\{(\alpha_t^n, \beta_t^n)\}$ and the controlled dynamics are described by (1) - (5). The coupled objectives of the market makers are given by (7), the coupling being induced by (3) and (4). The initial conditions for the game are the initial conditions for equations (1), (2), and (5). We are interested in studying the Nash equilibria of this game.

What remains is to formally specify the equilibrium concept that we will consider for this game. This includes any measurability and integrability conditions that the admissible controls must satisfy, as well as any restrictions on the class of equilibria that will be studied. This is carried out in the subsection below. The rest of the chapter is dedicated to characterizing and analyzing the model's equilibrium.

Equilibrium Concept

We begin by specifying the admissibility conditions that the market makers' demand schedules must satisfy. Firstly, we need to make sure that the system (3) - (4) can be solved uniquely to define progressively measurable trading rates and prices. Secondly, prices and trading rates must be sufficiently well-behaved so that the (implicit) integrals in (2) and the double integral in (7) converge absolutely. Finally, we will want to place some measurability restrictions on the demand schedule processes in order to reflect the type of information that market makers have access to. This gives rise to the definition of admissible profiles of demand schedules.

Definition 2.1.1. Given initial conditions $(\vec{x}, \eta, d) \in \mathbb{R}^N \times \mathbb{R} \times \mathbb{R}$ for $(X_0^1, \cdots, X_0^N)$, $\mathcal{N}_0$, and D_0, we say that the profile of progressively measurable demand schedules $\{(\alpha_t^1, \beta_t^1)\}, \cdots, \{(\alpha_t^N, \beta_t^N)\}$ is **admissible starting from** $(\vec{x}, \eta, d)$ if:

[4] We take as controls the parameters of the inverse demand as opposed to the demand. This simplifies much of the algebra below. We will continue to refer to the controls as demand schedules.

1. $\beta_t^n > 0 \ \forall t \geq 0, \ \forall n = 1, \cdots, N$ almost surely

2. $\int_0^T |q_t^n| dt < \infty \ \forall T \geq 0, \ \forall n = 1, \cdots, N$ almost surely

3. The double integral (7) converges absolutely

4. $\alpha_t^n, \beta_t^n \in \sigma\Big(\{D_s\}_{0 \leq s \leq t}, \{p_s\}_{0 \leq s < t}, \{X_s^n\}_{0 \leq s \leq t}, S_0\Big) \ \forall t \geq 0, \ \forall n = 1, \cdots N.$

The first condition says that market makers may only submit strictly decreasing demand schedules, and it guarantees that the system (3) - (4) can be solved uniquely to define progressively measurable trading rates and prices. The second and third conditions ensure that the integrals in (2) and (7) converge. The fourth condition states that the information a market maker has access to at any moment in time consists of the history of valuations, the history of prices, the history of his own shareholdings, and the initial level of the liquidity traders' shareholdings. This is essentially the information that exchanges provide to traders in reality.

We are interested in identifying admissible profiles of demand schedules that are Nash equilibria. Given a profile of demand schedules $\{(\alpha_t^1, \beta_t^1)\}, \cdots, \{(\alpha_t^N, \beta_t^N)\}$ that's admissible starting from $(\vec{x}, \eta, d)$, denote by $J^n\big(\vec{x}, \eta, d, \{(\alpha_t^1, \beta_t^1)\}, \cdots, \{(\alpha_t^N, \beta_t^N)\}\big)$ the value of the double integral (7). This is the payoff that market maker n receives when everyone's strategies are $\{(\alpha_t^1, \beta_t^1)\}, \cdots, \{(\alpha_t^N, \beta_t^N)\}$ and the initial conditions are $(X_0^1, \cdots, X_0^N) = \vec{x}, \ \mathcal{N}_0 = \eta$, and $D_0 = d$.

Definition 2.1.2. Given initial conditions $(\vec{x}, \eta, d)$, we say that a profile of demand schedules $\{(\alpha_t^1, \beta_t^1)\}, \cdots, \{(\alpha_t^N, \beta_t^N)\}$ is a **Nash Equilibrium starting from** $(\vec{x}, \eta, d)$ if:

1. The profile is admissible starting from $(\vec{x}, \eta, d)$

2. For any $n = 1, \cdots, N$, and for any demand schedule process $\{(\alpha_t, \beta_t)\}$ such that
   $\{(\alpha_t^1, \beta_t^1)\}, \cdots, \{(\alpha_t^{n-1}, \beta_t^{n-1})\}, \{(\alpha_t, \beta_t)\}, \{(\alpha_t^{n+1}, \beta_t^{n+1})\}, \cdots, \{(\alpha_t^N, \beta_t^N)\}\big)$ is ad-

missible starting from $(\vec{x}, \eta, d)$, we have that

$$J^n\big(\vec{x}, \eta, d, \{(\alpha_t^1, \beta_t^1)\}, \cdots, \{(\alpha_t^N, \beta_t^N)\}\big) \geq$$
$$J^n\big(\vec{x}, \eta, d, \{(\alpha_t^1, \beta_t^1)\}, \cdots, \{(\alpha_t^{n-1}, \beta_t^{n-1})\}, \{(\alpha_t, \beta_t)\}, \{(\alpha_t^{n+1}, \beta_t^{n+1})\}, \cdots, \{(\alpha_t^N, \beta_t^N)\}\big).$$

The second condition is the standard condition for a Nash equilibrium. It states that when all the market makers but market maker n play their equilibrium demand schedules, the maximum payoff the n^{th} market maker can earn is if he also plays his equilibrium demand schedule. In other words, when considering the optimal response problem against a profile of equilibrium demand schedules, each market maker finds it optimal to also use his equilibrium demand schedule.

Unfortunately, identifying all the Nash equilibria in this model is intractable and beyond the scope of this thesis. Instead we will focus on a special class of equilibria where all the market makers' demand schedules have a linear and symmetric structure. While this is fairly restrictive, the equilibria seem quite realistic and exhibit interesting dynamics.

Denote by $-S_t$ the collective share holdings of the liquidity traders at time t. Since the asset is in zero net supply, S_t must satisfy

$$S_t = X_t^1 + \cdots + X_t^N = X_0^1 + \int_0^t q_s^1 ds + \cdots + X_0^N + \int_0^t q_s^N ds = S_0 + \int_0^t \mathcal{N}_s ds \quad (2.8)$$

$\forall t \geq 0$. We will only consider equilibria where all market makers use demand schedules with the same constant slope and with an intercept that is the same linear function of individual state variables. The individual state variables will be D_t, X_t^n, and S_t. The precise formulation is given in the next definitions.

Definition 2.1.3. A profile of demand schedules $\{(\alpha_t^1, \beta_t^1)\}, \cdots, \{(\alpha_t^N, \beta_t^N)\}$ is said to be **linear symmetric** if $\exists\ a, \lambda, b, c, \xi \in \mathbb{R}$ s.t.

$$\alpha_t^n = aX_t^n + bD_t + cS_t + \xi \quad (2.9)$$

$$\beta_t^n = \lambda \quad (2.10)$$

$\forall t \geq 0,\ \forall n = 1, \cdots, N$.

Definition 2.1.4. We say that $a, \lambda, b, c, \xi \in \mathbb{R}$ are a **linear symmetric Nash equilibrium** if the linear symmetric profile defined by (9) and (10) is a Nash equilibrium starting from any set of initial conditions.

For the rest of the chapter we focus only on linear symmetric Nash equilibria. Our goal will be to classify all such equilibria and to study the dynamics of prices and trading rates in these equilibria. Before proceeding we prove the following lemma, which states the for linear symmetric profiles the admissibility[5] conditions manifest in simple constraints on the parameters a and λ. One of the key points of the lemma is that under a linear symmetric profile market makers can infer S_t from the history of prices and valuations. The market makers are thus able to implement the demand schedules (9) and (10) given their individual information sets.

Lemma 2.1.5. *A linear symmetric profile is admissible if and only if* $\lambda > 0$ *and* $\frac{a}{\lambda} < \frac{\rho}{2}$.

Proof. Fix a linear symmetric profile given by $a, \lambda, b, c, \xi \in \mathbb{R}$ as in (9) and (10). We need to show that the four conditions for admissibility in Definition 1.1 are satisfied if and only if $\lambda > 0$ and $\frac{a}{\lambda} < \frac{\rho}{2}$. Clearly the first condition holds if and only if $\lambda > 0$. Next we will show that a profile satisfying (9) and (10) always satisfies the second and fourth conditions. Finally we will show that the third condition holds if and only if $\frac{2a}{\lambda} < \rho$, thus completing the proof of the lemma.

Note that by combining equations (9) and (10) with equations (2), (3) and (4) we can conclude that prices, trading rates, and inventories under a linear symmetric profile

[5] Admissibility for a linear symmetric profile means that (9) and (10) are admissible given *any* initial conditions.

must satisfy

$$p_t = \left(\frac{a}{N} + c\right)S_t + bD_t + \xi - \frac{\lambda}{N}\mathcal{N}_t \tag{2.11}$$

$$X_t^n = e^{\frac{a}{\lambda}t}\left(X_0^n - \frac{S_0}{N}\right) + \frac{S_t}{N} \tag{2.12}$$

$$q_t^n = \frac{a}{\lambda}\left(X_t^n - \frac{S_t}{N}\right) + \frac{1}{N}\mathcal{N}_t \tag{2.13}$$

$\forall t \geq 0$ and $\forall n = 1, \cdots, N$. These formulas imply that trading rates are almost surely continuous and thus the second condition holds.

To prove the fourth condition it suffices to prove that $S_t \in \sigma\Big(\{D_s\}_{0\leq s\leq t}, \{p_s\}_{0\leq s<t}, S_0\Big)$ $\forall t \geq 0$. Equation (11) implies that the following ODE holds path by path for the process $\{S_t\}$:

$$\frac{dS_t}{dt} = \frac{N}{\lambda}\left(\frac{a}{N} + c\right)S_t + \mathcal{A}_t,$$

where the process $\{\mathcal{A}_t\}$ is defined by

$$\mathcal{A}_t = \frac{N}{\lambda}(bD_t + \xi - p_t).$$

It's clear from this definition that $\{\mathcal{A}_s\}_{0\leq s<t} \in \sigma\Big(\{D_s\}_{0\leq s\leq t}, \{p_s\}_{0\leq s<t}, S_0\Big)$ $\forall t \geq 0$. Furthermore the ODE above implies that

$$S_t = e^{\frac{N}{\lambda}\left(\frac{a}{N}+c\right)t}S_0 + \int_0^t e^{\frac{N}{\lambda}\left(\frac{a}{N}+c\right)(t-s)}\mathcal{A}_s ds$$

from which it follows that $S_t \in \sigma\Big(\{D_s\}_{0\leq s\leq t}, \{p_s\}_{0\leq s<t}, S_0\Big)$ $\forall t \geq 0$.

Turning to the third condition for admissibility, note that (11) - (13) imply that under a linear symmetric profile the integrand in (7) is of the form $e^{-\rho t}Q(X_t^n, D_t, \mathcal{N}_t, S_t)$, where Q is a second order polynomial with an $(X_t^n)^2$ coefficient of $-\frac{\gamma\sigma_D^2}{2} \neq 0$. Thus $\exists$ constants $M_0, M_1, M_2, M_3, M_4 > 0$ such that

$$|Q(X_t^n, D_t, \mathcal{N}_t, S_t)| \leq M_0 + M_1\big((X_t^n)^2 + D_t^2 + \mathcal{N}_t^2 + S_t^2\big)$$

$$(X_t^n)^2 \leq M_2 + M_3|Q(X_t^n, D_t, \mathcal{N}_t, S_t)| + M_4(D_t^2 + \mathcal{N}_t^2 + S_t^2)$$

$\forall t \geq 0$ almost surely. Equations (1), (5), and (8) imply that $e^{-\rho t}D_t^2$, $e^{-\rho_t}\mathcal{N}_t^2$, and $e^{-\rho t}S_t^2$ are integrable for any initial conditions. Thus from these bounds it follows that (7) converges absolutely for any initial conditions and for any $n = 1, \cdots, N$ if and only if $e^{-\rho t}(X_t^n)^2$ is integrable for any initial conditions and for any $n = 1, \cdots, N$. From (12) it follows that this latter condition holds if and only if $\frac{a}{\lambda} < \frac{\rho}{2}$. □

2.2 The Optimal Response Problem

This section formulates the optimal response problem for an individual market maker as a standard stochastic control problem. A key point is that in forming an optimal response, a market maker can optimize directly over his trading rate. This is analogous to the idea in Chapter 1 of formulating the optimal response problem as maximization against a linear supply curve.

Proposition 2.2.1. *$a, \lambda, b, c, \xi \in \mathbb{R}$ are a linear symmetric Nash equilibrium if and only if $\lambda > 0$, $\frac{a}{\lambda} < \frac{\rho}{2}$, and for any initial conditions we have that*

$$\frac{a}{\lambda}\Big(X_t - \frac{S_t}{N}\Big) + \frac{\mathcal{N}_t}{N} \in \arg\max_{\{q_t\}} \mathbb{E}\Bigg[\int_0^\infty e^{-\rho t}\Big(-q_t(p_t - D_t) + \mu X_t - \frac{\gamma\sigma_D^2}{2}(X_t)^2\Big)dt\Bigg], \quad (2.14)$$

where p_t is given by equation (15) below, the relevant dynamics are

$$\begin{aligned} dX_t &= q_t dt \\ dD_t &= \mu dt + \sigma_D dB_t^D \\ d\mathcal{N}_t &= -\psi \mathcal{N}_t dt + \sigma_{\mathcal{N}} dB_t^{\mathcal{N}} \\ dS_t &= \mathcal{N}_t dt \end{aligned}$$

and the optimization is constrained to those processes $\{q_t\}$ such that

1. $\int_0^T |q_t| dt < \infty \ \forall T \geq 0$
2. *The double integral in (14) converges absolutely*
3. $q_t \in \sigma(\{D_u\}_{0\leq u\leq t}, \{S_u\}_{0\leq u\leq t}, \{X_u\}_{0\leq u\leq t}, \{\mathcal{N}_u\}_{0\leq u\leq t}) \ \forall t \geq 0.$

Proof. To have a Nash equilibrium we must have a profile that satisfies the admissibility and optimality conditions in Definition 1.2. Lemma 1.5 states that for a linearly symmetric profile the admissibility condition holds if and only if $\lambda > 0$ and $\frac{a}{\lambda} < \frac{\rho}{2}$.

Thus to prove the proposition we need only show that the optimality condition holds if and only if (14) holds.

Consider the optimal response problem of an individual market maker when facing a linear symmetric profile of demand schedules given by $a, \lambda, b, c, \xi \in \mathbb{R}$. Denote the inventory process and trading rate process for the remaining market maker by $\{X_t\}$ and $\{q_t\}$. If the remaining market maker chooses the demand schedule process $\{(\alpha_t, \beta_t)\}$, then his trading rate and the price process $\{p_t\}$ are given implicitly by

$$p_t = \left(\frac{a}{N-1} + c\right) S_t + bD_t + \xi - \frac{\lambda}{N-1} \mathcal{N}_t - \frac{a}{N-1} X_t + \frac{\lambda}{N-1} q_t \tag{2.15}$$

$$p_t = \alpha_t - \beta_t q_t. \tag{2.16}$$

The optimal response problem for the remaining market maker is to choose the demand schedule process $\{(\alpha_t, \beta_t)\}$ such that the profile of all agents' schedules is admissible[6], and the objective in (14) is maximized over all such processes. In computing the objective the relevant equations are (15), (16), and the differential equations in the statement of the proposition. The optimality condition for Nash equilibrium is satisfied if and only if for any set of initial conditions a maximizing choice of demand schedule process is $(\alpha_t, \beta_t) = (aX_t + bD_t + cS_t + \xi, \lambda)$. We now make three points regarding this optimal response problem.

The first point is that the remaining market maker will be indifferent between any two demand schedule processes leading to the same trading rate. Indeed for fixed initial conditions, the objective in (14) depends on the choice of demand schedule process only through the trading rate and price processes. Furthermore, from (15) we see that the price process depends on the demand schedule processes only through the trading rate process. Hence the objective is constant over demand schedule processes leading to the same trading rate process.

The second point is that by choosing an appropriate admissible demand schedule, the remaining market maker can achieve any trading rate satisfying the three conditions

[6] In the rest of the proof we will say that the remaining market maker's demand schedule process is admissible if the corresponding profile of all market makers' demand schedules is admissible.

in the proposition. To prove this it suffice to show that

$$S_t,\ \mathcal{N}_t \in \sigma\Big(\{D_s\}_{0\le s\le t}, \{p_s\}_{0\le s<t}, \{X_s\}_{0\le s\le t}, \{q_s\}_{0\le s<t}, S_0\Big) \quad \forall t \ge 0 \qquad (2.17)$$

for any choice of the remaining market maker's demand schedule. If this is the case, then for any $\{\tilde{q}_t\}$ satisfying the three conditions in the proposition we have that the demand schedule process

$$\left\{\left(\Big(\frac{a}{N-1}+c\Big)S_t + bD_t + \xi - \frac{\lambda}{N-1}\mathcal{N}_t - \frac{a}{N-1}X_t + \Big(1+\frac{\lambda}{N-1}\Big)\tilde{q}_t, 1\right)\right\}$$

is admissible, and by (15) and (16) it gives the remaining market maker the trading rate process $\{q_t\} = \{\tilde{q}_t\}$.

To prove (17) it suffices to prove that

$$\{S_u\}_{0\le u\le t} \in \sigma\Big(\{D_s\}_{0\le s\le t}, \{p_s\}_{0\le s<t}, \{X_s\}_{0\le s\le t}, \{q_s\}_{0\le s<t}, S_0\Big) \quad \forall t \ge 0 \qquad (2.18)$$

for any choice of the remaining market maker's demand schedule. This is because by differentiating $\{S_u\}_{0\le u\le t}$ we can recover $\{\mathcal{N}\}_{0\le u\le t}$. To prove (18) note that (15) implies that for any choice of the remaining market maker's demand schedule we have

$$\begin{aligned}\frac{dS_t}{dt} &= \frac{N-1}{\lambda}\Big(\frac{a}{N-1}+c\Big)S_t + \mathcal{A}_t\\ \mathcal{A}_t &= \frac{N-1}{\lambda}\Big(bD_t + \xi - \frac{a}{N-1}X_t + \frac{\lambda}{N-1}q_t - p_t\Big).\end{aligned}$$

From this (18) follows exactly as in Lemma 1.5.

The final point is that if the remaining market maker chooses the demand schedule process $(\alpha_t, \beta_t) = (aX_t + bD_t + cS_t + \xi, \lambda)$, then his trading rate is

$$q_t = \frac{a}{\lambda}\Big(X_t - \frac{S_t}{N}\Big) + \frac{\mathcal{N}_t}{N} \qquad (2.19)$$

$\forall t \ge 0$. Indeed (19) follows simply by plugging (α_t, β_t) in (15) - (16) and solving for q_t.

This third point shows that if the remaining market maker follows the linear symmetric profile, then his trading rate is the maximizer in (14). The first two points imply that solving the optimal response problem against a linear symmetric profile is the same as solving the constrained optimization problem in the proposition. Hence the proposition follows. □

This proposition associates to each linear symmetric profile $a, \lambda, b, c, \xi \in \mathbb{R}$ an optimization problem, as well as a candidate solution of the problem. The parameters provide an equilibrium if and only if the candidate is truly a solution. In the next section we will use the theory of stochastic control to derive first order conditions for the optimization problem. Using these we will demonstrate that there is a unique profile for which the candidate is a true solution, and thus there is a unique linear symmetric equilibrium.

Before proceeding, we establish some notation relevant to the proposition and it's use below. Given a set of parameters $a, \lambda, b, c, \xi \in \mathbb{R}$, the optimization in the proposition is a standard stochastic control problem on an infinite horizon, with state space $(x, d, \eta, s) \in \mathbb{R}^4$ and control space $q \in \mathbb{R}$. Denote the covariables for the problem by $y = (y_x, y_d, y_\eta, y_s) \in \mathbb{R}^4$ and

$$z = \begin{pmatrix} z_{xx} & z_{xd} & z_{x\eta} & z_{xs} \\ z_{dx} & z_{dd} & z_{d\eta} & z_{ds} \\ z_{\eta x} & z_{\eta d} & z_{\eta\eta} & z_{\eta s} \\ z_{sx} & z_{sd} & z_{s\eta} & z_{ss} \end{pmatrix} \in \mathbb{R}^{4\times 4}.$$

The Hamiltonian for the problem is $H : \mathbb{R}^4 \times \mathbb{R}^4 \times \mathbb{R}^{4\times 4} \times \mathbb{R} \to \mathbb{R}$ given by

$$\begin{aligned} H(x, d, \eta, s, y, z, q) = q y_x + \mu y_d - \psi \eta y_\eta + \eta y_s + \frac{\sigma_D^2}{2} z_{dd} + \frac{\sigma_N^2}{2} z_{\eta\eta} + \mu x \\ - \frac{\gamma \sigma_D^2}{2} x^2 - q\big(P(x, d, \eta, s, q) - d\big), \quad (2.20) \end{aligned}$$

where $P : \mathbb{R}^4 \times \mathbb{R} \to \mathbb{R}$ is the function specifying prices as a function of state and control from (15) above, i.e.

$$P(x, d, \eta, s, q) = \Big(\frac{a}{N-1} + c\Big)s + bd + \xi - \frac{\lambda}{N-1}\eta - \frac{a}{N-1}x + \frac{\lambda}{N-1}q. \quad (2.21)$$

Also denote by $Q : \mathbb{R}^4 \to \mathbb{R}$ the mapping corresponding to the Markov control in (14), i.e.

$$Q(x, d, \eta, s) = \frac{a}{\lambda}\big(x - \frac{s}{N}\big) + \frac{\eta}{N}. \tag{2.22}$$

H and P describe the optimization problem, and Q is the candidate solution. The functions H, P, and Q all depend on the parameters a, λ, b, c, and ξ, but this dependence is suppressed in the notation.

2.3 Equilibrium Characterization

Theorem 2.3.1. *Fix exogenous parameters $N \geq 3$, $\rho, \gamma, \sigma_D, \sigma_N, \psi > 0$, and $\mu \in \mathbb{R}$. There is a uique linear symmetric Nash equilibrium with price*

$$p_t = D_t + \frac{\mu}{\rho} - \frac{1}{\rho}\frac{\gamma}{N}\sigma_D^2 S_t - \frac{\gamma}{N}\frac{N-1}{2}\frac{\sigma_D^2}{\delta(\delta - \frac{\rho}{2})}\mathcal{N}_t$$

and trading rates

$$q_t^n = -(\delta - \frac{\rho}{2})\Big(X_t^n - \frac{S_t}{N}\Big) + \frac{1}{N}\mathcal{N}_t,$$

where

$$\delta := \frac{1}{2}\sqrt{\rho^2 + 2\rho(\rho + \psi)(N-2)}.$$

Proof. We will prove that $a, \lambda, b, c, \xi \in \mathbb{R}$ are a linear symmetric Nash equilibrium if and only if

$$a = -\frac{N-1}{\delta}\gamma\sigma_D^2 \tag{2.23}$$

$$\lambda = \frac{N-1}{N-2}\frac{\gamma\sigma_D^2}{\rho+\psi}\Big(\frac{1}{\delta} + \frac{1}{\rho}\Big) \tag{2.24}$$

$$b = 1 \tag{2.25}$$

$$c = -\frac{1}{\rho}\frac{\gamma\sigma_D^2}{N} + \frac{N-1}{\delta}\frac{\gamma\sigma_D^2}{N} \tag{2.26}$$

$$\xi = \frac{\mu}{\rho} \tag{2.27}$$

Equilibrium prices and trading rates are given by (11) - (13), and plugging in these values gives the formulas in the statement of the theorem.

We begin by proving the only if part of the statement. To this end, suppose the parameters $a, \lambda, b, c, \xi \in \mathbb{R}$ are a linear symmetric Nash equilibrium. Denote by $V(x, d, \eta, s)$ the value function of the optimization problem in Proposition 2.1, i.e. the value of the supremum in (14) when the initial conditions are $(X_0, D_0, \mathcal{N}_0, S_0) = (x, d, \eta, s)$.

Note that the value function is smooth. Indeed because the parameters provide an equilibrium, (14) must hold, and therefore $V(x, d, \eta, s)$ can be computed by evaluating the objective in (14) along the control process $\frac{a}{\lambda}\left(X_t - \frac{S_t}{N}\right) + \frac{\mathcal{N}_t}{N}$. This provides us with an explicit expression for V, and by direct inspection it follows that V is smooth.

It follows that V satisfies the following HJB equation:

$$\rho V(x, d, \eta, s) = \sup_{q\in\mathbb{R}} H(x, d, \eta, s, \nabla V, \nabla^2 V, q) \quad \forall (x, d, \eta, s) \in \mathbb{R}^4. \tag{2.28}$$

Furthermore, since Q is an optimal Markov control, it also follows that

$$Q(x, d, \eta, s) \in \arg\max_{q\in\mathbb{R}} H(x, d, \eta, s, \nabla V, \nabla^2 V, q) \quad \forall (x, d, \eta, s) \in \mathbb{R}^4. \tag{2.29}$$

(28) is the classical result that if the value function is smooth then it satisfies the HJB equation [Tou13]. When an optimal Markov control is known to exist, one way to prove (28) is to first prove (29) [Car16]. A lemma explicitly proving (29) is included in the appendix to this chapter for completeness.

The first order condition for (29) is

$$V_x(x, d, \eta, s) = P\big(x, d, \eta, s, Q(x, d, \eta, s)\big) - d + \frac{\lambda}{N-1} Q(x, d, \eta, s) \quad \forall (x, d, \eta, s) \in \mathbb{R}^4. \tag{2.30}$$

Anti-differentiating (30) it follows that $\exists$ a smooth function $w : \mathbb{R}^3 \to \mathbb{R}$ such that

$$V(x, d, \eta, s) = \frac{1}{2}\frac{a}{N-1}x^2 + (b-1)xd - \Big(\frac{N-2}{N(N-1)}\lambda\Big)x\eta + \Big(\frac{N-2}{N(N-1)}a + c\Big)xs + \xi x + w(d, \eta, s) \tag{2.31}$$

$\forall (x, d, \eta, s) \in \mathbb{R}^4$.

In summary, we've shown that if $a, \lambda, b, c, \xi \in \mathbb{R}$ are a linear symmetric Nash equilibrium then (28) - (31) hold. Combining these equations, we conclude that $\exists$ a smooth function $w : \mathbb{R}^3 \to \mathbb{R}$ such that

$$\frac{\rho}{2}\frac{a}{N-1}x^2+\rho(b-1)xd-\rho\frac{N-2}{N(N-1)}\lambda x\eta+\rho\Big(\frac{N-2}{N(N-1)}a+c\Big)xs+\rho\xi x+\rho w(d,\eta,s)$$
$$= \Big(\frac{a^2}{\lambda(N-1)} - \frac{\gamma\sigma_D^2}{2}\Big)x^2 + \Big(\frac{a}{N-1} + \frac{N-2}{N(N-1)}\psi\lambda + c\Big)x\eta - \frac{2a^2}{N(N-1)\lambda}xs + \mu bx$$
$$+ \mu w_d - \psi\eta w_\eta + \eta w_s + \frac{\sigma_D^2}{2}w_{dd} + \frac{\sigma_N^2}{2}w_{\eta\eta} + \frac{\lambda}{(N-1)N^2}\Big(\eta - \frac{a}{\lambda}s\Big)^2 \quad (2.32)$$

$\forall(x, d, \eta, s) \in \mathbb{R}^4$.

It remains to be shown that (32) implies (23) - (27). Note that since the function w is independent of x, the coefficients of x^2, xd, $x\eta$, xs, and x must be equal on the left and right hand sides of (32). Equating the coefficients of xd and x immediately gives (25) and (27). Equating the remaining coefficients yields the algebraic system

$$\frac{\rho}{2}\frac{a}{N-1} = \frac{a^2}{\lambda(N-1)} - \frac{\gamma\sigma_D^2}{2} \quad (2.33)$$
$$-\rho\frac{N-2}{N(N-1)}\lambda = \frac{a}{N-1} + \frac{N-2}{N(N-1)}\psi\lambda + c \quad (2.34)$$
$$\rho\Big(\frac{N-2}{N(N-1)}a + c\Big) = -\frac{2a^2}{N(N-1)\lambda}. \quad (2.35)$$

Using equations (33) and (35) we can solve for c in terms of a to get $c = -\frac{a}{N} - \frac{1}{\rho}\frac{\gamma}{N}\sigma_D^2$. From this equation it follows that if (23) and (24) hold then so too does (26). Next we plug this expression for c into (34) and solve for λ in terms of a to get $\lambda = -\frac{1}{\rho+\psi}\Big(\frac{a}{N-2} - \frac{N-1}{N-2}\frac{\gamma\sigma_D^2}{\rho}\Big)$. From this equation it follows that if (23) holds then so too does (24). Hence it only remains to prove that (23) holds. Plugging this expression for λ into (33), we see that a must satisfy $a^2 = \frac{(N-1)^2\gamma^2\sigma_D^4}{\delta^2}$. Now, because the parameters provide an equilibrium, the constraint $\frac{2a}{\lambda} < \rho$ must hold. The positive root for a violates the constraint and the negative root satisfies it, so it follows that (23) holds.

We now prove the if part of the statement. To this end, suppose that $a,\ b,\ c,\ \xi$, and λ are given by equations (23) - (27). We need to show (14) holds, i.e. that the mapping

Q in (22) provides an optimal Markov control for the stochastic control problem in Proposition 2.1. This will be done by by using the verification theorem for the HJB equation [Pha09].

Given a set of initial conditions, denote by $\{\hat{q}_t\}$ and $\{\hat{X}_t\}$ the trading rate and inventory processes arising from using the Markov control given by Q, i.e. $d\hat{X}_t = Q(\hat{X}_t, D_t, \mathcal{N}_t, S_t)dt$ and $\hat{q}_t = Q(\hat{X}_t, D_t, \mathcal{N}_t, S_t)$. These processes depend on the choice of initial conditions, but this is suppressed in the notation. We need to show that $\exists$ a smooth function $V(x, d, \eta, s)$ such that

1. (28) holds
2. (29) holds
3. $e^{-\rho t}\mathbb{E}[V(\hat{X}_t, D_t, \mathcal{N}_t, S_t)] \longrightarrow 0$ as $t \longrightarrow \infty$ for any choice of initial conditions
4. $\{\hat{q}_t\}$ satisfies the the constraints in Proposition 3.1 for any choice of initial conditions.

If this can be done then it follows by the verification theorem that Q is an optimal Markov control for the stochastic control problem in Proposition 2.1.

Note that we can find a second order polynomial $w(d, \eta, s)$ that satisfies the equation

$$\rho w = \mu w_d - \psi\eta w_\eta + \eta w_s + \frac{\sigma_D^2}{2} w_{dd} + \frac{\sigma_{\mathcal{N}}^2}{2} w_{\eta\eta} + \frac{\lambda}{(N-1)N^2}\left(\eta - \frac{a}{\lambda}s\right)^2$$

on all of $\mathbb{R}^3$. Now define the function $V(x, d, \eta, s)$ by equation (31). Then the function V is smooth and by construction equations (30) and (32) hold. Notice that as a function of q the Hamiltonian is a quadratic polynomial with leading coefficient $-\frac{\lambda}{N-1}$. Since $\lambda > 0$ it follows that (30) is not only a necessary condition for (29) but also a sufficient one. Thus (29) holds. This implies that the equation (28) is precisely the equation (32), and so (28) also holds.

Since V is a second order polynomial, in order to prove the third condition it suffices to show that $\mathbb{E}[e^{-\rho t}\hat{X}_t^2]$, $\mathbb{E}[e^{-\rho t}D_t^2]$, $\mathbb{E}[e^{-\rho t}\mathcal{N}_t^2]$, and $\mathbb{E}[e^{-\rho t}S_t^2]$ converge to 0 as $t \longrightarrow \infty$ for any set of initial conditions. This follows from the fact that $\frac{2a}{\lambda} < \rho$ and $\psi > 0$.

Finally, we need to check that $\{\hat{q}_t\}$ satisfies the constraints in Proposition 3.1. The third constraint is trivial since $\hat{q}_t = Q(\hat{X}_t, D_t, \mathcal{N}_t, S_t)$. This formula also implies $\{\hat{q}_t\}$ is almost surely continuous, and thus the first condition holds. Lastly, the second condition holds because $\frac{2a}{\lambda} < \rho$ and $\psi > 0$. □

2.4 Equilibrium Analysis

Price Analysis

Consider the equilibrium price process given in Theorem 3.1. The four terms admit intuitive economic interpretations. The first term D_t is simply the market makers' current valuation for the asset. The second term $\frac{\mu}{\rho}$ is a premium for expected valuation growth. μ is the drift of valuations, so on average valuations increase by $\mu(T-t)$ over a time interval $[t, T]$. As this is common knowledge among the market makers, this must be reflected in the price at time t. Otherwise, a profitable deviation from equilibrium would be to buy the asset at time t and sell it at time T. Since market makers discount payoffs from time T to time t by $e^{-\rho(T-t)}$, the appropriate premium in the price to prevent this deviation is $\frac{\mu}{\rho}$. Said another way, we have[7]

$$\frac{\mu}{\rho} = \mathbb{E}_t\left[\int_t^{\infty} e^{-\rho(T-t)} dD_T\right].$$

Thus the second term is the (risk-neutral) present value of expected future changes in the asset's value.

Since the market makers are not risk neutral, they also require risk compensations to take exposures to the asset. This is the role of the third term in the price $-\frac{1}{\rho}\frac{\gamma\sigma_D^2}{N}S_t$. When the market makers are in aggregate long the asset, so S_t is positive, this term is negative and thus the asset is trading at a relatively low price. Because the asset is trading at a low price, market makers don't find it profitable to deviate from equilibrium by selling the asset to reduce their exposures. Similarly, when the market makers are in aggregate short the asset, this term causes the asset to trade at a high price, and thus market makers don't find it profitable to buy the asset to reduce their exposures.

[7]The notation here and below is $\mathbb{E}_t[\cdot] := \mathbb{E}[\cdot|\mathcal{F}_t]$.

The magnitude of the compensation per unit of exposure is given by $\frac{1}{\rho}\frac{\gamma\sigma_D^2}{N}$. σ_D^2 is the volatility of valuations, so this is the amount of risk per unit of exposure to the asset. $\frac{\gamma}{N}$ is the aggregate risk aversion of the market makers, so this is the dollar compensation they require to hold a unit of risk. Thus $\frac{\gamma\sigma_D^2}{N}$ is the dollar compensation the market makers require to hold a unit of exposure to the asset. The presence of $\frac{1}{\rho}$ suggests a discounted present value interpretation, but it is not straightforward to write this term as a discounted present value like the second term. This is partly because this term is not the full risk compensation market makers require, as discussed a few paragraphs below.

The final term in the price process is $-\frac{\gamma}{N}\frac{N-1}{2}\frac{\sigma_D^2}{\delta(\delta-\frac{\rho}{2})}\mathcal{N}_t$, which is the price impact the liquidity traders face on their trades, or equivalently the slope of the supply curve they face when trading. The liquidity traders submit market orders to trade at rate $-\mathcal{N}_t$, so they are buying when $\mathcal{N}_t < 0$ and selling when $\mathcal{N}_t > 0$. When the liquidity traders are buying, price impact causes the trading price to be high, and when the liquidity traders are selling, price impact causes the trading price to be low. This is the model's analogue of liquidity traders' market orders walking the book. As suggested by Kyle, we define liquidity in the model as the reciprocal of price impact and study it's comparative statics.

Definition 2.4.1. $Price\ Impact := \frac{\gamma}{N}\frac{N-1}{2}\frac{\sigma_D^2}{\delta(\delta-\frac{\rho}{2})}$

Definition 2.4.2. $Liquidity := \frac{1}{Price\ Impact}$

Proposition 2.4.3.

1. $\frac{\partial}{\partial\gamma}Liquidity < 0$.

   Liquidity is decreasing in market makers' risk aversion.

2. $\frac{\partial}{\partial\sigma_D}Liquidity < 0$.

   Liquidity is decreasing in fundamental volatility.

3. *If* $\frac{\gamma}{N}$ *is held fixed then* $\frac{\partial}{\partial N}Liquidity > 0$.

   Liquidity is increasing in market maker competition.

4. $\frac{\partial}{\partial \psi} Liquidity > 0$.

   Liquidity is decreasing in order flow uncertainty.

To understand the fourth point, recall that the liquidity traders' order flow is given by $\{\mathcal{N}_t\}$, which is an Ornstein-Uhlenbeck process whose stationary distribution has variance $\frac{\sigma_N^2}{2\psi}$. Thus as ψ increases, the liquidity traders' orders arrive with less uncertainty.

These comparative statics agree with real word intuition about liquidity, and thus they justify the theoretical definition of liquidity. Put simply, liquidity should reflect market makers' willingness to absorb temporary flow. Market makers are less willing to absorb temporary flow when they are more risk averse, when the asset is riskier, when there are not many of them, or when order flow is more uncertain. By the proposition, liquidity is also lower in the model in these situations.

Next we study what happens to price impact in the competitive limit of the model. This is the limit when $N \to \infty$ and $\frac{\gamma}{N} \to \gamma_0$. Essentially one considers the sequence of models with N market makers each having risk aversion $\gamma_N := N\gamma_0$. Thus for any model in the sequence, the aggregate risk aversion of the market makers is $\frac{\gamma_N}{N} = \gamma_0$. Hence the sequence considers increasingly competitive market making sectors that in aggregate have the same risk bearing capacity. Taking the limit of the sequence provides a perfectly competitive benchmark for the model.

Proposition 2.4.4. *In the competitive limit, we have that*

$$Price\ Impact \to \frac{\gamma_0}{\rho} \frac{\sigma_D^2}{\rho + \psi}$$

and

$$p_t \to D_t + \frac{\mu}{\rho} - \mathbb{E}_t\left[\int_t^{\infty} e^{-\rho(T-t)} \gamma_0 \sigma_D^2 S_T dT \right] \quad \forall (t, \omega) \in [0, \infty) \times \Omega$$

That price impact does not vanish in the competitive limit is somewhat surprising. Naive intuition would suggest that market makers exercise their market power by charging price impact. So in the competitive limit, when individual market makers no longer have market power, price impact should vanish. The reason this intuition does not hold is given by the limiting value of the equilibrium price.

In the competitive limit, market makers have an aggregate risk aversion of $\gamma_0 > 0$, so the equilibrium price should consist of a risk neutral present value as well as a discount based on the market makers' aggregate exposure. The first two terms in the limiting price, $D_t + \frac{\mu}{\rho}$, are a risk neutral present value as discussed above. Thus the third term $-\mathbb{E}_t\left[\int_t^\infty e^{-\rho(T-t)}\gamma_0\sigma_D^2 S_T dT\right]$ should be the appropriate risk discount.

As discussed above, S_T is the aggregate exposure of the market makers at time T, and $-\gamma_0\sigma_D^2 S_T$ is the dollar compensation they require in order to maintain this exposure. However, because S_T evolves with T, this is only the exposure over the infinitesimal time interval $[T, T+dt]$. Furthermore, standing at time t, market makers require a compensation for the entire path of exposures they will be taking over $[t, \infty)$. The appropriate compensation for the exposure over $[T, T+dt]$ is $\gamma_0\sigma_D^2 S_T$, and market makers discount payoffs from time T to time t by $e^{-\rho(T-t)}$, so the appropriate compensation for the entire path of exposures over $[t, \infty)$ is $-\mathbb{E}_t\left[\int_t^\infty e^{-\rho(T-t)}\gamma_0\sigma_D^2 S_T dT\right]$.

Said another way, the third term is exactly what is needed to prevent market makers deviating from equilibrium by pursuing a strategy that buys/sells based on the current level of the aggregate exposure. If S_T does not evolve with T and has a constant value equal to S, then this term reads $-\frac{1}{\rho}\gamma_0\sigma_D^2 S$. This is exactly the risk discount a representative CARA investor with risk aversion γ_0 and time discount rate ρ would require to hold S shares of an asset with volatility σ_D^2.

Based on this analysis, we can conclude, as suggested previously, that $-\frac{1}{\rho}\frac{\gamma\sigma_D^2}{N}S_t$ is not the full risk compensation component of the equilibrium price, but instead that $-\mathbb{E}_t\left[\int_t^\infty e^{-\rho(T-t)}\gamma_0\sigma_D^2 S_T dT\right]$ is. Thus part of the price impact component of prices compensates market makers for risk, and so it does not vanish in the competitive limit. The rest of the price impact component of prices is a manifestation of market makers' market power, and so it does vanish in the competitive limit. The interpretation is that market makers charge price impact for (at least) two reasons. The first reason is simply because they can, since they have market power, and doing so is profitable. The second reason is that incoming trades change the entire path of exposures that market makers will be taking going forward. In order for the market makers to find it utility maximizing to clear the incoming trade, this must be reflected in the trading price.

Inventory Analysis

Next consider the equilibrium trading rates given in Theorem 3.1. The two terms encode two important properties of the market makers' equilibrium trading behavior. The first is that in aggregate the market makers must buy at rate $\mathcal{N}_t$, or equivalently the average trading rate of the market makers must be $\frac{1}{N}\mathcal{N}_t$. This is simply a consequence of market clearing, since the liquidity traders are selling at rate $\mathcal{N}_t$. That this is indeed the case is guaranteed by the second term in the formula for trading rates; the first terms all cancel out when aggregating/averaging.

The first term in the trading rates dictates which market makers buy more/less than the average. Note that $\delta > \frac{\rho}{2}$, so the market makers with inventories larger than $\frac{S_t}{N}$ buy less, and vice versa. This brings us to the second important property of the market makers' trading behavior: they are continually moving towards a Pareto optimal allocation amongst themselves. The market makers must in aggregate hold S_t shares, simply by market clearing. Since they are all identical, it would be Pareto optimal for everyone to hold $\frac{S_t}{N}$ shares. However, this certainly can't hold at time 0, as initial inventories are exogenous and arbitrary. Beyond time 0 efficiency of the allocations depends on endogenous trading behavior.

Using the formula for trading rates, we can compute that in equilibrium the trajectory of each market maker's inventory is

$$X_t^n = e^{-(\delta-\frac{\rho}{2})t}(X_0^n - \frac{S_0}{N}) + \frac{S_t}{N}.$$

Thus each market maker deviates from the efficient allocation by the first term. This term is non-zero if and only if $X_0^n \neq \frac{S_0}{N}$, and in this case it converges monotonically towards 0 over time. Thus the market makers take as given the inefficiency in their initial allocations, and then trade amongst themselves to make allocations more efficient. Allocations are inefficient at time t is because they were inefficient at time 0; the endogenous trading behavior of the market makers does not in any way create allocational inefficiencies.

But the market makers' trading behavior also does not perfectly remove the initial inefficiency in allocations. Indeed, inventories converge to efficiency at an exponential rate of $\delta - \frac{\rho}{2}$, and this rate is not infinite. Based on the formula for δ, we see that the two main drivers of this rate are ψ, governing the uncertainty in order flow, and N, governing the degree of competition. As $\psi \to \infty$, so order flow uncertainty vanishes, or as $N \to \infty$, so the market is perfectly competitive, this rate of convergence goes to infinity.

The story here is essentially one of individual market makers behaving strategically in an attempt to earn profits. Note that the price in the absence of liquidity trades ($\mathcal{N}_t = 0$), interpreted as the mid-price, is $D_t + \frac{\mu}{\rho} - \frac{1}{\rho}\frac{\gamma}{N}\sigma_D^2 S_t$. Any market maker holding more than $\frac{S_t}{N}$ shares should find this price high relative to his exposure and should want to sell, and vice versa. However, market makers take into account price impact, so they know they can't actually trade to $\frac{S_t}{N}$ at this price. The market makers also know that at any moment a random liquidity trade might come in and move their exposure in the right direction, but now without price impact[8]. Thus instead of trading all the way to $\frac{S_t}{N}$, market makers only move partially and take the chance that the liquidity traders' flow will move them the rest of the way. In the absence of order flow uncertainty, or under perfect competition, the market makers no longer go through this calculation and simply trade all the way to $\frac{S_t}{N}$, converging instantly to efficiency (i.e. at an infinite rate).

2.5 Appendix: A Stochastic Control Lemma

Fix a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ equipped with a d-dimensional Brownian motion $\{B_t\}$ and satisfying the usual conditions. Consider a standard infinite

[8] More specifically, price impact *favors* the market maker in this situation, as his limit order is getting hit by an incoming liquidity trader market order. If the market maker was insistent on moving to $\frac{S_t}{N}$, then he would have to place a market order, so price impact would go *against* him.

horizon stochastic control problem with state space $\mathbb{R}^S$, control space $\mathbb{R}^A$, and randomness coming from $\{B_t\}$. Thus we take as given mappings $b : \mathbb{R}^S \times \mathbb{R}^A \to \mathbb{R}^S$, $\sigma : \mathbb{R}^S \times \mathbb{R}^A \to \mathbb{R}^{S\times d}$, and $f : \mathbb{R}^S \times \mathbb{R}^A \to \mathbb{R}$. We assume that b and σ are uniformly Lipschitz in their first variables and that f is Borel measurable. Also fix a constant $\beta > 0$.

Denote by $\mathcal{A}$ the set of progressively measurable $\mathbb{R}^A$ valued processes $\alpha = \{\alpha_t\}$ such that

$$\mathbb{E}\bigg[\int_0^T |b(0,\alpha_t)|^2 + |\sigma(0,\alpha_t)|^2 dt\bigg] \quad \forall T > 0.$$

Given $x \in \mathbb{R}^S$ and $\alpha = \{\alpha_t\} \in \mathcal{A}$, denote by $X^{x,\alpha} = \{X_t^{x,\alpha}\}$ the unique strong solution[9] to the SDE

$$\begin{aligned} dX_t &= b(X_t,\alpha_t)dt + \sigma(X_t,\alpha_t)dB_t \\ X_0 &= x. \end{aligned}$$

For each $x \in \mathbb{R}^S$ denote by $\mathcal{A}(x)$ the subset of $\alpha = \{\alpha_t\} \in \mathcal{A}$ such that

$$\mathbb{E}\bigg[\int_0^\infty e^{-\beta t}|f(X_t^{x,\alpha},\alpha_t)|dt\bigg] < \infty$$

and assume that $\mathcal{A}(x)$ is nonempty $\forall x \in \mathbb{R}^S$.

For $x \in \mathbb{R}^S$ and $\alpha = \{\alpha_t\} \in \mathcal{A}(x)$ define the cost functional by

$$J(x,\alpha) = \mathbb{E}\bigg[\int_0^\infty e^{-\beta t} f(X_t^{x,\alpha},\alpha_t)dt\bigg].$$

The value function is defined for $x \in \mathbb{R}^S$ by

$$V(x) = \sup_{\alpha\in\mathcal{A}(x)} J(x,\alpha).$$

We also define the Hamiltonian $H : \mathbb{R}^S \times \mathbb{R}^S \times \mathbb{R}^{S\times S} \times \mathbb{R}^A \to \mathbb{R}$ by

$$H(x,y,z,a) = b(x,a)\cdot y + \frac{1}{2} trace\big(\sigma(x,a))\sigma^T(x,a)z\big) + f(x,a).$$

[9] The assumptions on b and σ and the definition of $\mathcal{A}$ were made precisely so that this equation admits a unique strong solution [Pha09].

The significance of the Hamiltonian is that it describes how functions of a controlled state process evolve over time. That is, by Ito's formula, for any $\alpha = \{\alpha_t\} \in \mathcal{A}$, $h > 0$ and $g \in C^{1,2}\big([0,\infty) \times \mathbb{R}^S\big)$ we have that

$$g(t+h, X^{x,\alpha}_{t+h}) - g(t, X^{x,\alpha}_t) = \int_t^{t+h} \frac{\partial g}{\partial t}(s, X^{x,\alpha}_s) + H\big(X^{x,\alpha}_s, \nabla g(s, X^{x,\alpha}_s), \nabla^2 g(s, X^{x,\alpha}_s), \alpha_s\big) - f(X^{x,\alpha}_s, \alpha_s)\, ds + Martingale.$$

Definition 2.5.1. We say that $a : \mathbb{R}^S \to \mathbb{R}^A$ is an **optimal Markov control** if $\forall x \in \mathbb{R}^S$, $\exists \hat{\alpha}^x = \{\hat{\alpha}^x_t\} \in \arg\max_{\alpha \in \mathcal{A}(x)} J(x,\alpha)$ such that $\hat{\alpha}^x_t = a\big(X^{x,\hat{\alpha}^x}_t\big)$ $\forall t \geq 0$ almost surely.

Lemma 2.5.2. *Suppose that $V \in C^2(\mathbb{R}^S)$ and H is continuous. If $a : \mathbb{R}^S \to \mathbb{R}^A$ is continuous and provides an optimal Markov control then*

$$a(x) \in \arg\max_{\tilde{a} \in \mathbb{R}^A} H(x, \nabla V(x), \nabla^2 V(x), \tilde{a}) \quad \forall x \in \mathbb{R}^S.$$

Proof. We proceed by contradiction. Suppose that the conclusion of the theorem does not hold. Then $\exists x_0 \in \mathbb{R}^S$, $\tilde{a}_0 \in \mathbb{R}^A$ and $\epsilon > 0$ such that

$$H(x_0, \nabla V(x_0), \nabla^2 V(x_0), a(x_0)) < H(x_0, \nabla V(x_0), \nabla^2 V(x_0), \tilde{a}_0) - 4\epsilon.$$

By continuity of a, H and V it follows that $\exists$ a neighborhood U_1 of x_0 such that

$$H(x, \nabla V(x), \nabla^2 V(x), a(x)) < H(x, \nabla V(x), \nabla^2 V(x), \tilde{a}_0) - 3\epsilon \quad \forall x \in U_1. \tag{2.36}$$

Let $\{X_t\}$ be the unique strong solution to the SDE

$$\begin{aligned} dX_t &= b(X_t, a(X_t))dt + \sigma(X_t, a(X_t))dBt \\ X_0 &= x_0 \end{aligned}$$

and let $\{\tilde{X}_t\}$ be the unique strong solution to the SDE

$$\begin{aligned} d\tilde{X}_t &= b(\tilde{X}_t, \tilde{a}_0)dt + \sigma(\tilde{X}_t, \tilde{a}_0)dBt \\ \tilde{X}_0 &= x_0. \end{aligned}$$

Note that by smoothness of V and continuity of H we can find neighborhoods U_2 and U_3 of x_0 such that

$$|V(x) - V(y)| < \frac{\epsilon}{\beta} \quad \forall x, y \in U_2 \tag{2.37}$$

$$|H(x, \nabla V(x), \nabla^2 V(x), \tilde{a}_0) - H(y, \nabla V(y), \nabla^2 V(y), \tilde{a}_0)| < \epsilon \quad \forall x, y \in U_3. \tag{2.38}$$

Define the stopping time $\tau = \inf\{t \geq 0 : (X_t, \tilde{X}_t) \notin U_1 \cap U_2 \cap U_3 \times U_1 \cap U_2 \cap U_3\}$. Since the processes $\{X_t\}$ and $\{\tilde{X}_t\}$ have a.s. continuous paths it follows that

$$\tau > 0 \tag{2.39}$$

almost surely.

Now, we can estimate

$$\begin{aligned}
V(x_0) &= \mathbb{E}\Big[\int_0^\tau e^{-\beta t} f(X_t, a(X_t))dt + e^{-\beta\tau} V(X_\tau)\Big] \\
&= V(x_0) + \mathbb{E}\Big[\int_0^\tau e^{-\beta t}\Big(H\big(X_t, \nabla V(X_t), \nabla^2 V(X_t), a(X_t)\big) - \beta V(X_t)\Big)dt\Big] \\
&\leq V(x_0) + \mathbb{E}\Big[\int_0^\tau e^{-\beta t}\Big(H\big(X_t, \nabla V(X_t), \nabla^2 V(X_t), \tilde{a}_0\big) - 3\epsilon - \beta V(X_t)\Big)dt\Big] \\
&\leq V(x_0) + \mathbb{E}\Big[\int_0^\tau e^{-\beta t}\Big(H\big(\tilde{X}_t, \nabla V(\tilde{X}_t), \nabla^2 V(\tilde{X}_t), \tilde{a}_0\big) + \epsilon - 3\epsilon - \beta V(\tilde{X}_t) + \epsilon\Big)dt\Big] \\
&= \mathbb{E}\Big[\int_0^\tau e^{-\beta t} f(\tilde{X}_t, \tilde{a}_0)dt + e^{-\beta\tau} V(\tilde{X}_\tau)\Big] - \frac{\epsilon}{\beta}\mathbb{E}[1 - e^{-\beta\tau}] \\
&\leq V(x_0) - \frac{\epsilon}{\beta}\mathbb{E}[1 - e^{-\beta\tau}].
\end{aligned}$$

The first equality uses the dynamic programing principle [Pha09] and the optimality of the Markov control given by a. The second and last equality use Ito's lemma. The last inequality uses the dynamic programming principle. The inequalities in the middle follow from the definition of τ and inequalities (36) - (38). Since $\frac{\epsilon}{\beta} > 0$, it follows from this estimate that $\mathbb{E}[e^{-\beta\tau}] \geq 1$. However, this contradicts (39). □

Chapter 3

MEAN REVERTING LIQUIDITY TRADERS

In this chapter we consider the same model as in chapter two, but we modify the dynamics of the liquidity traders' market orders so that their inventories mean revert about zero. As before, market makers submit demand schedules at each instant in time, and then their trading rates and the trading price are determined by

$$\alpha_t^n - \beta_t^n q_t^n = p_t \qquad \forall n = 1, \cdots, N$$
$$q_t^1 + \cdots + q_t^N = \mathcal{N}_t.$$

Here $-\mathcal{N}_t$ is the current trading rate of the liquidity traders, interpreted as a market order to sell $\mathcal{N}_t dt$ shares over the time interval $[t, t + dt]$. The liquidity traders' time t inventory is then $-S_t$, where $S_t = S_0 + \int_0^t \mathcal{N}_s ds$ and $-S_0$ is their initial inventory. Since the asset is in zero net supply, $\{S_t\}$ is the supply process faced by the market making sector, i.e. $S_t = X_t^1 + \cdots + X_t^N$ $\forall t \geq 0$, where X_t^n is the time t inventory of each market maker.

In chapter two we took $\{\mathcal{N}_t\}$ to be an Ornstein-Uhlenbeck process, which meant that the long run expectation of $\{S_t\}$, conditional on the present, was not necessarily zero. Here we instead take

$$\mathcal{N}_t = -\phi(S_t - \tilde{S}_t)$$
$$d\tilde{S}_t = -\psi \tilde{S}_t dt + \sigma_{\tilde{S}} dB_t^{\tilde{S}},$$

where $\phi, \psi, \sigma_{\tilde{S}} > 0$.

The interpretation is that $-\tilde{S}_t$ is the current inventory target of the liquidity traders, and ϕ governs the speed with which they trade towards their target. Thus when their current inventory is above the target, $-S_t > -\tilde{S}_t$, the liquidity traders submit market orders to sell, $\mathcal{N}_t > 0$, and vice versa. If ϕ is large, the liquidity traders are impatient and submit large market orders to quickly reach their target. If ϕ is small, the liquidity traders are patient, submit smaller market orders, and only move slowly towards their target. One can think of $-\tilde{S}_t$ as the inventory the liquidity traders would want to hold if markets were perfectly liquid. Due to illiquidity they cannot instantly acquire this inventory, and instead do so gradually. Note that since $\{\tilde{S}_t\}$ mean reverts about zero, so too does $\{S_t\}$, i.e. we have $\mathbb{E}[S_T|S_t, \tilde{S}_t] \to 0$ as $T \to \infty$ almost surely $\forall t \geq 0$.

The rest of the model is exactly as before. We work on a filtered probability space equipped with two independent Brownian motions, one driving the inventory target process, $\{B_t^{\tilde{S}}\}$, and the other driving the valuation process, $\{B_t^D\}$. Valuations follow an arithmetic Brownian motion, market makers' trading rates are the time derivatives of their inventories, and market makers' objectives are exactly as before. The initial conditions for the game are $X_0^1, \cdots, X_0^N$, D_0 and $\tilde{S}_0$. The definitions of admissible profiles, Nash equilibrium, linear symmetric profiles, and linear symmetric Nash equilibrium are precisely as before.

Lemma 1.5. carries through verbatim and the analogue to the optimal response proposition is below. The proofs are omitted as they are essentially identical to those in chapter two. The rest of the chapter is dedicated to characterizing and analyzing linear symmetric Nash equilibria of the model.

Proposition 3.0.1. *$a, \lambda, b, c, \xi \in \mathbb{R}$ are a linear symmetric Nash equilibrium if and only if $\lambda > 0$, $\frac{a}{\lambda} < \frac{\rho}{2}$, and for any initial conditions we have that*

$$\frac{a}{\lambda}\Big(X_t - \frac{S_t}{N}\Big) - \frac{\phi}{N}(S_t - \tilde{S}_t) \in \arg\max_{\{q_t\}} \mathbb{E}\Big[\int_0^\infty e^{-\rho t}\Big(-q_t(p_t - D_t) + \mu X_t - \frac{\gamma\sigma_D^2}{2}(X_t)^2\Big)dt\Big], \tag{3.1}$$

where

$$p_t = \Big(\frac{a}{N-1} + c + \frac{\lambda\phi}{N-1}\Big)S_t + bD_t + \xi - \frac{\lambda\phi}{N-1}\tilde{S}_t - \frac{a}{N-1}X_t + \frac{\lambda}{N-1}q_t.$$

The relevant dynamics are

$$\begin{aligned} dX_t &= q_t dt \\ dD_t &= \mu dt + \sigma_D dB_t^D \\ d\tilde{S}_t &= -\psi \tilde{S}_t dt + \sigma_{\tilde{S}} dB_t^{\tilde{S}} \\ dS_t &= -\phi(S_t - \tilde{S}_t)_t dt \end{aligned}$$

and the optimization is constrained to those processes $\{q_t\}$ *such that*

1. $\int_0^T |q_t| dt < \infty \quad \forall T \geq 0$
2. *The double integral on the right side of (1) converges absolutely*
3. $q_t \in \sigma(\{D_u\}_{0\leq u\leq t}, \{S_u\}_{0\leq u\leq t}, \{X_u\}_{0\leq u\leq t}, \{\tilde{S}_u\}_{0\leq u\leq t}) \quad \forall t \geq 0.$

3.1 Equilibrium Characterization

Theorem 3.1.1. *Fix exogenous parameters* $N \geq 3$, $\rho, \gamma, \sigma_D, \phi, \psi, \sigma_{\tilde{S}} > 0$ *and* $\mu \in \mathbb{R}$. *There is a unique linear symmetric Nash equilibrium with price*

$$p_t = D_t + \frac{\mu}{\rho} - \theta \frac{\gamma \sigma_D^2}{N} S_t - \frac{\gamma}{N} \frac{N-1}{N-2} \frac{\rho \sigma_D^2}{(\rho+\psi)(\rho+\phi)} \Big(\frac{1}{\delta} + \frac{1}{\rho}\Big) \mathcal{N}_t$$

and trading rates

$$q_t^n = -\kappa\Big(X_t^n - \frac{S_t}{N}\Big) + \frac{1}{N}\mathcal{N}_t,$$

where

$$\begin{aligned} \kappa &:= \rho(N-2)\frac{\rho+\psi}{\rho+\delta} \\ \delta &:= \sqrt{\rho^2 + 2(N-2)(\rho+\psi)(\rho+\phi)} \\ \theta &:= \frac{(\rho+\psi+\phi)\delta - \psi\phi}{(\rho+\psi)(\rho+\phi)}. \end{aligned}$$

Proof. The proof uses Proposition 1 and stochastic control theory analogously to the proof in Chapter 2. The state space for the optimization in Proposition 1 is

$(x, d, \tilde{s}, s) \in \mathbb{R}^4$, the control space is $q \in \mathbb{R}$ and the Hamiltonian is

$$H(x, d, \tilde{s}, s, y, z, q) = qy_x + \mu y_d - \psi \tilde{s} y_{\tilde{s}} - \phi(s - \tilde{s})y_s + \frac{\sigma_D^2}{2} z_{dd} + \frac{\sigma_{\tilde{S}}^2}{2} z_{\tilde{s}\tilde{s}} + \mu x - \frac{\gamma\sigma_D^2}{2} x^2 - q\big(P(x, d, \tilde{s}, s, q) - d\big).$$

Here $P : \mathbb{R}^4 \times \mathbb{R} \to \mathbb{R}$ is given by

$$P(x, d, \tilde{s}, s, q) = \Big(\frac{a}{N-1} + c + \frac{\lambda\phi}{N-1}\Big)s + bd + \xi - \frac{\lambda\phi}{N-1}\tilde{s} - \frac{a}{N-1}x + \frac{\lambda}{N-1}q.$$

Denote by $Q : \mathbb{R}^4 \to \mathbb{R}$ the candidate optimal Markov control given in in the proposition, i.e.

$$Q(x, d, \tilde{s}, s) = \frac{a}{\lambda}\big(x - \frac{s}{N}\big) - \frac{\phi}{N}(s - \tilde{s}).$$

We will prove that $a, \lambda, b, c, \xi \in \mathbb{R}$ are a linear symmetric Nash equilibrium if and only if

$$a = -\frac{N-1}{\delta}\gamma\sigma_D^2 \tag{3.2}$$

$$\lambda = \frac{N-1}{N-2}\frac{\rho\gamma\sigma_D^2}{(\rho+\psi)(\rho+\phi)}\Big(\frac{1}{\delta} + \frac{1}{\rho}\Big) \tag{3.3}$$

$$b = 1 \tag{3.4}$$

$$c = -\frac{\rho(\rho+\psi+\phi)}{(\rho+\psi)(\rho+\phi)}\frac{\gamma\sigma_D^2}{N}\Big(\frac{1}{\delta} + \frac{1}{\rho}\Big) + \frac{\gamma\sigma_D^2}{\delta} \tag{3.5}$$

$$\xi = \frac{\mu}{\rho}. \tag{3.6}$$

The equations from Chapter 2 relating equilibrium prices and trading rates to equilibrium values of a, λ, b, c, ξ continue to hold. Plugging in (2)-(6) gives the formulas in the theorem.

Beginning with the if part of the statement, suppose the parameters $a, \lambda, b, c, \xi \in \mathbb{R}$ are a linear symmetric Nash equilibrium. As in chapter two, it follows that the Hamiltonian H, the optimal Markov control Q and the value function V are related by

$$\rho V(x, d, \tilde{s}, s) = \sup_{q\in\mathbb{R}} H(x, d, \tilde{s}, s, \nabla V, \nabla^2 V, q) \quad \forall (x, d, \tilde{s}, s) \in \mathbb{R}^4 \tag{3.7}$$

and

$$Q(x,d,\tilde{s},s) \in \arg\max_{q\in\mathbb{R}} H(x,d,\tilde{s},s,\nabla V,\nabla^2 V,q) \quad \forall (x,d,\tilde{s},s) \in \mathbb{R}^4. \tag{3.8}$$

As before, we can anti-differentiate the first order condition for (8) to obtain an expression for the value function in terms of an auxiliary smooth function $w(d,\tilde{s},s)$. Once again we feed this expression for V back in to the HJB (7) and equate like terms, using the fact that w is independent of x. This immediately gives (4) and (6), as well as the following system for a, λ and c:

$$\frac{\rho}{2}\frac{a}{N-1} = \frac{a^2}{\lambda(N-1)} - \frac{\gamma\sigma_D^2}{2} \tag{3.9}$$

$$-\rho\frac{N-2}{N(N-1)}\lambda = \frac{a}{N-1} + \frac{N-2}{N(N-1)}(\psi+\phi)\lambda + c \tag{3.10}$$

$$\rho\Big(\frac{N-2}{N(N-1)}(a+\lambda\phi)+c\Big) = -\frac{2a^2}{N(N-1)\lambda} - \frac{N-2}{N(N-1)}\phi^2\lambda - \phi\Big(\frac{a}{N-1}+c\Big). \tag{3.11}$$

This system can be solved for a, λ and c by following the exact same procedure used in the proof of Chapter 2. This yields two possible values for a, but only the one given in (2) satisfies the constraint $\frac{a}{\lambda} < \frac{\rho}{2}$. This proves the if portion of the statement.

Turning to the only if part of the statement, suppose that a, b, c, ξ, and λ are given by equations (2) - (6). As in chapter 2, we will use the verification theorem to show that (1) holds. Note that we can find a second order polynomial $w(d,\tilde{s},s)$ that satisfies the equation

$$\rho w = \mu w_d - \psi\tilde{s}w_{\tilde{s}} - \phi(s-\tilde{s})w_s + \frac{\sigma_D^2}{2}w_{dd} + \frac{\sigma_{\tilde{S}}^2}{2}w_{\tilde{s}\tilde{s}} + \left(\frac{\phi}{N}\sqrt{\frac{\lambda}{N-1}}(s-\tilde{s}) + \frac{a}{N}\frac{1}{\sqrt{\lambda(N-1)}}s\right)^2$$

on all of $\mathbb{R}^3$. Consider the expression for the value function derived in the if portion of the proof above. Insert this w as the auxiliary function there, and then define V using that expression. Then by construction V satisfies (7) and (8), with the first order condition for (8) being sufficient because $\lambda > 0$. The admissibility and transversality conditions hold because $\frac{a}{\lambda} < \frac{\rho}{2}$. Thus (1) follows by the verification theorem. □

Below is a simple corollary that will be used in the next chapter.

Corollary 3.1.2. *Suppose that we instead take* $\mathcal{N}_t = -\phi(S_t - \pi\tilde{S}_t)$ *for some* $\pi \neq 0$. *Then the characterization of linear symmetric Nash equilibria is unchanged.*

Proof. Define $\tilde{\tilde{S}}_t := \pi\tilde{S}_t$. Then we have

$$\begin{aligned}\mathcal{N}_t &= -\phi(S_t - \tilde{\tilde{S}}_t)\\ d\tilde{\tilde{S}}_t &= -\psi\tilde{\tilde{S}}_t + \pi\sigma_{\tilde{S}}dB_t^{\tilde{S}}.\end{aligned}$$

Thus the model is exactly as above except with $\sigma_{\tilde{S}}$ replaced with $\pi\sigma_{\tilde{S}}$. Since the equilibrium parameters found in the theorem above don't depend on $\sigma_{\tilde{S}}$, it follows that the characterization of linear symmetric equilibria is unchanged. □

Chapter 4

ENDOGENOUS LIQUIDITY TRADERS

In this chapter we endogenize the liquidity traders' market order process and derive it as the result of an optimization problem. As in Chapters 2 and 3, the central equations determining the trading price and trading rates at each instant in time are

$$\alpha_t^n - \beta_t^n q_t^n = p_t \qquad \forall n = 1, \cdots, N$$
$$q_t^1 + \cdots + q_t^N = \mathcal{N}_t.$$

The demand schedules $\{(\alpha_t^n, \beta_t^n)\}$ are the control processes of each market maker, and are chosen to maximize

$$\mathbb{E}\Big[\int_0^\infty e^{-\rho t}\Big(- q_t^n(p_t - D_t) - \frac{\gamma\sigma_D^2}{2}(X_t^n)^2\Big)dt\Big]. \tag{4.1}$$

The market orders to sell $\{\mathcal{N}_t\}$ are the control process of the liquidity traders, and are chosen to maximize

$$\mathbb{E}\Big[\int_0^\infty e^{-\rho t}\Big(\mathcal{N}_t(p_t - D_t) - \frac{\theta\sigma_D^2}{2}(S_t - \tilde{S}_t)^2\Big)dt\Big]. \tag{4.2}$$

Here X_t^n is the inventory of each market maker and $-S_t$ is the inventory of the liquidity traders, which evolve according to

$$dX_t^n = q_t^n dt \tag{4.3}$$

$$dS_t = \mathcal{N}_t dt. \tag{4.4}$$

D_t is the exogenous valuation of the asset shared by all agents, and it evolves according to[1]

$$dD_t = \sigma_D dB_t^D. \tag{4.5}$$

$\tilde{S}_t$ is an exogenous component of the liquidity traders' inventory, and it evolves according to

$$d\tilde{S}_t = -\psi\tilde{S}_t dt + \sigma_{\tilde{S}} dB_t^{\tilde{S}}. \tag{4.6}$$

[1] In this chapter we set the drift of the valuation, μ, equal to 0 for simplicity. Including the drift does not have a material effect on the results.

$\{B_t^D\}$ and $\{B_t^{\tilde{S}}\}$ are independent Brownian motions.

The justification for the liquidity traders' objective function is as follows. Firstly, the liquidity traders' inventory is $-S_t$ and it's derivative, the trading rate of the liquidity traders, is $-\mathcal{N}_t$. Thus $\mathcal{N}_t$ is the *selling* rate of the liquidity traders, as it's the negative derivative of their inventory. This is why the first term in the liquidity traders' objective function, $\mathcal{N}_t(p_t - D_t)$, corresponding to costs from trading, does not have a minus sign like the corresponding term in (1). In contrast, q_t^n is the *buying* rate of the market makers, as it's the positive derivative of their inventories.

Secondly, we assume that in addition to the inventory built from trading on the exchange, the liquidity traders also have another inventory of $\tilde{S}_t$, perhaps from trading with clients. Thus the aggregate inventory of the liquidity traders is $-S_t + \tilde{S}_t$, and so their quadratic inventory holding costs are $(-S_t + \tilde{S}_t)^2 = (S_t - \tilde{S}_t)^2$. The risk aversion parameter of the liquidity traders is θ, giving rise to the second term in their objective function.

This sets up the model as a stochastic differential game between the N market makers *and* the liquidity traders. The initial conditions for the game are the initial conditions for equations (3), (5) and (6). Since the asset is in zero net supply, the initial conditions for (3) also provide an initial condition for (4): $S_0 = X_0^1 + \cdots X_0^N$. We are interested in studying the Nash equilibria of this game.

As before, it's necessary to specify admissibility conditions for the agents' controls, and also to place restrictions on the class of equilibria that will be studied. This is all fairly analogous to chapter two, but the definitions are included below for completeness. After this, the next section characterizes equilibria in terms of the zeros of a quartic polynomial. The following section analyzes properties of equilibrium.

Definition 4.0.1. Given initial conditions $(\vec{x}, \tilde{s}, d) \in \mathbb{R}^N \times \mathbb{R} \times \mathbb{R}$ for $(X_0^1, \cdots, X_0^N)$, $\tilde{S}_0$, and D_0, we say that the profile of progressively measurable demand schedules $\{(\alpha_t^1, \beta_t^1)\}, \cdots, \{(\alpha_t^N, \beta_t^N)\}$ and liquidity trader selling rates $\{\mathcal{N}_t\}$ is **admissible starting from** $(\vec{x}, \tilde{s}, d)$ if:

1. $\beta_t^n > 0\ \forall t \geq 0,\ \forall n = 1, \cdots, N$ almost surely

2. $\int_0^T |q_t^n| dt < \infty\ \forall T \geq 0,\ \forall n = 1, \cdots, N$ almost surely

3. The double integral (1) converges absolutely

4. $\alpha_t^n, \beta_t^n \in \sigma\Big(\{D_s\}_{0 \leq s \leq t}, \{p_s\}_{0 \leq s < t}, \{X_s^n\}_{0 \leq s \leq t}, S_0\Big)\ \forall t \geq 0,\ \forall n = 1, \cdots N$

5. $\int_0^T |\mathcal{N}_t| dt < \infty\ \forall T \geq 0$ almost surely

6. The double integral (2) converges absolutely

7. $\mathcal{N}_t \in \sigma\Big(\{D_s\}_{0 \leq s \leq t}, \{p_s\}_{0 \leq s < t}, \{S_s\}_{0 \leq s \leq t}, \{\tilde{S}_s\}_{0 \leq s \leq t}\Big)\ \forall t \geq 0.$

Given a profile $\{(\alpha_t^1, \beta_t^1)\}, \cdots, \{(\alpha_t^N, \beta_t^N)\}, \{\mathcal{N}_t\}$ that's admissible starting from $(\vec{x}, \tilde{s}, d)$, denote by $J^n\big(\vec{x}, \tilde{s}, d, \{(\alpha_t^1, \beta_t^1)\}, \cdots, \{(\alpha_t^N, \beta_t^N)\}, \{\mathcal{N}_t\}\big)$ the value of the double integral (1) and by $J^{LT}\big(\vec{x}, \tilde{s}, d, \{(\alpha_t^1, \beta_t^1)\}, \cdots, \{(\alpha_t^N, \beta_t^N)\}, \{\mathcal{N}_t\}\big)$ the value of the double integral (2). These are the payoffs the market makers and liquidity traders receive when everyone's strategies are $\{(\alpha_t^1, \beta_t^1)\}, \cdots, \{(\alpha_t^N, \beta_t^N)\}$, $\{\mathcal{N}_t\}$ and the initial conditions are $(X_0^1, \cdots, X_0^N) = \vec{x},\ \tilde{S}_0 = \tilde{s}$, and $D_0 = d$.

Definition 4.0.2. Given initial conditions $(\vec{x}, \tilde{s}, d)$, we say that a profile $\{(\alpha_t^1, \beta_t^1)\}, \cdots,$ $\{(\alpha_t^N, \beta_t^N)\}, \{\mathcal{N}_t\}$ is a **Nash Equilibrium starting from** $(\vec{x}, \tilde{s}, d)$ if:

1. The profile is admissible starting from $(\vec{x}, \tilde{s}, d)$

2. For any $n = 1, \cdots, N$, and for any demand schedule process $\{(\alpha_t, \beta_t)\}$ such that
$\{(\alpha_t^1, \beta_t^1)\}, \cdots, \{(\alpha_t^{n-1}, \beta_t^{n-1})\}, \{(\alpha_t, \beta_t)\}, \{(\alpha_t^{n+1}, \beta_t^{n+1})\}, \cdots, \{(\alpha_t^N, \beta_t^N)\}\big), \{\mathcal{N}_t\}$
is admissible starting from $(\vec{x}, \tilde{s}, d)$, we have that

$$J^n\big(\vec{x}, \eta, d, \{(\alpha_t^1, \beta_t^1)\}, \cdots, \{(\alpha_t^N, \beta_t^N)\}, \{\mathcal{N}_t\}\big) \geq$$
$$J^n\big(\vec{x}, \eta, d, \{(\alpha_t^1, \beta_t^1)\}, \cdots, \{(\alpha_t^{n-1}, \beta_t^{n-1})\}, \{(\alpha_t, \beta_t)\}, \{(\alpha_t^{n+1}, \beta_t^{n+1})\}, \cdots, \{(\alpha_t^N, \beta_t^N)\}, \{\mathcal{N}_t\}\big).$$

3. For any liquidity trader selling process $\{\mathcal{N}_t'\}$ such that
$\{(\alpha_t^1, \beta_t^1)\}, \cdots, \{(\alpha_t^N, \beta_t^N)\}, \{\mathcal{N}_t'\}$ is admissible starting from $(\vec{x}, \tilde{s}, d)$, we have

that

$$J^{LT}\big(\vec{x}, \tilde{s}, d, \{(\alpha_t^1, \beta_t^1)\}, \cdots, \{(\alpha_t^N, \beta_t^N)\}, \{\mathcal{N}_t\}\big) \geq$$
$$J^{LT}\big(\vec{x}, \tilde{s}, d, \{(\alpha_t^1, \beta_t^1)\}, \cdots, \{(\alpha_t^N, \beta_t^N)\}, \{\mathcal{N}_t'\}\big).$$

Definition 4.0.3. A profile $\{(\alpha_t^1, \beta_t^1)\}, \cdots, \{(\alpha_t^N, \beta_t^N)\}, \{\mathcal{N}_t\}$ is said to be **linear symmetric** if $\exists\ a, b, c, \lambda, \phi, M \in \mathbb{R}$ s.t.

$$\alpha_t^n = aX_t^n + bD_t + cS_t \tag{4.7}$$

$$\beta_t^n = \lambda \tag{4.8}$$

$$\mathcal{N}_t = -\phi S_t + M\tilde{S}_t \tag{4.9}$$

$\forall t \geq 0,\ \forall n = 1, \cdots, N.$

Definition 4.0.4. We say that $a, b, c, \lambda, \phi, M \in \mathbb{R}$ are a **linear symmetric Nash equilibrium** if the profile defined by (7), (8), and (9) is a Nash equilibrium starting from any set of initial conditions.

4.1 Equilibrium Characterization

Theorem 4.1.1. *Fix exogenous parameters $N \geq 3$ and $\rho, \gamma, \theta, \sigma_D, \psi, \sigma_{\tilde{S}} > 0$.*
There is a one to one correspondence between linear symmetric Nash equilibria and $\kappa > -\frac{\rho}{2} + \frac{1}{2}\sqrt{\rho^2 + \rho(\rho+\psi)(N-2)}$ that satisfy the quartic equation (29).
In equilibrium the trading price is

$$p_t = D_t - \Gamma S_t - \Lambda \mathcal{N}_t, \tag{4.10}$$

the LTs selling rate is

$$\mathcal{N}_t = -\phi(S_t - \pi\tilde{S}_t) \tag{4.11}$$

and the MMs trading rates are

$$q_t^n = -\kappa(X_t^n - \frac{S_t}{N}) + \frac{\mathcal{N}_t}{N}. \tag{4.12}$$

The endogenous parameters Γ, Λ, ϕ *and* π *are given in terms of* κ *by*

$$\phi = \frac{2\kappa^2}{(N-2)(\rho+\psi)} + \frac{2\rho\kappa}{(N-2)(\rho+\psi)} - \rho \tag{4.13}$$

$$\pi = \frac{\theta\sigma_D^2}{2\Lambda(\rho+\psi+\phi)\phi} \tag{4.14}$$

$$\Lambda = \frac{\gamma\sigma_D^2}{N}\frac{N-1}{2\kappa^2+\rho\kappa} \tag{4.15}$$

$$\Gamma = \left(\rho+\psi+\phi-\frac{\kappa}{N-2}\right)\frac{N-2}{N-1}\Lambda. \tag{4.16}$$

Proof. For the first half of the proof we need to show that any linear symmetric Nash equilibrium satisfies all the properties in the statement of the theorem. To this end, fix parameters $a, b, c, \lambda, \phi, M \in \mathbb{R}$ and suppose that they provide a linear symmetric Nash equilibrium. Thus for any initial conditions the strategy profile given by (7)-(9) satisfies the admissibility and optimality conditions in Definition 1 and Definition 2.

The optimality condition for the LTs strategy implies that for any initial conditions

$$-\phi S_t + M\tilde{S}_t \in \arg\max_{\{\mathcal{N}_t\}} \mathbb{E}\left[\int_0^\infty e^{-\rho t}\left(\mathcal{N}_t(p_t - D_t) - \frac{\theta\sigma_D^2}{2}(S_t-\tilde{S}_t)^2\right)dt\right], \tag{4.17}$$

where

$$p_t = bD_t - \Gamma S_t - \Lambda\mathcal{N}_t \tag{4.18}$$

with $\Gamma := -(\frac{a}{N}+c)$ and $\Lambda := \frac{\lambda}{N}$. The relevant dynamics are given by (4), (5), and (6) and the optimization is constrained to those $\{\mathcal{N}_t\}$ for which the profile with (7) and (8) satisfies the conditions in Definition 1.

Denote by $V(d,s,\tilde{s})$ the value function for the optimization (17), i.e. the value of the supremum when the initial conditions are $D_0 = d, S_0 = s, \tilde{S}_0 = \tilde{s}$. The Hamiltonian for the optimization is

$$H(d,s,\tilde{s},\eta,y,z) := \eta y_s - \psi\tilde{s}y_{\tilde{s}} + \frac{\sigma_{\tilde{S}}^2}{2}z_{\tilde{s}\tilde{s}} + \frac{\sigma_D^2}{2}z_{dd} + \big((b-1)d - \Gamma s - \Lambda\eta\big)\eta - \frac{\theta\sigma_D^2}{2}(s-\tilde{s})^2. \tag{4.19}$$

As in Chapter 2 we have that the value function and Hamiltonian are smooth, and so the following hold:

$$-\phi s + M\tilde{s} \in \arg\max_{\eta} H(d, s, \tilde{s}, \nabla V, \nabla^2 V) \tag{4.20}$$

$$\rho V = \sup_{\eta} H(d, s, \tilde{s}, \nabla V, \nabla^2 V). \tag{4.21}$$

Antidifferentiating the first order condition for (20) implies that ∃ a smooth function $w : \mathbb{R}^2 \to \mathbb{R}$ such that

$$V(d, s, \tilde{s}) = \left(\frac{\Gamma}{2} - \Lambda\phi\right)s^2 + 2\Lambda M s\tilde{s} - (b-1)sd + w(d, \tilde{s}). \tag{4.22}$$

As in Chapter 2, combining this with equations (20) and (21), and equating like terms, we obtain the following equilibrium consistency conditions for the parameters:

$$b = 1 \tag{4.23}$$

$$\rho(\Gamma - 2\Lambda\phi) = -\theta\sigma_D^2 + 2\Lambda\phi^2 \tag{4.24}$$

$$(\rho + \psi)M = \frac{\theta\sigma_D^2}{2\Lambda} - \phi M. \tag{4.25}$$

An immediate consequence of equation (25) is that $M \neq 0$ because $\theta\sigma_D^2 > 0$. Next we show, by contradiction, that (24) implies $\phi \neq 0$. If $\phi = 0$, then by (24) we have $\Gamma = -\frac{\theta\sigma_D^2}{\rho} < 0$. On the other hand, if $\phi = 0$ then $\mathcal{N}_t = M\tilde{S}_t$. Since $M \neq 0$, it follows that $\{\mathcal{N}_t\}$ is an Ornstein-Uhlenbeck process with mean reversion parameter $\psi > 0$, exactly as in Chapter 2. Thus repeating the argument from the Theorem 1 in Chapter 2, the optimality conditions for the MMs imply that $\Gamma = -(\frac{a}{N} + c) = \frac{1}{\rho}\frac{\gamma\sigma_D^2}{N} > 0$, which is a contradiction. Hence becasue of (24) it must be that $\phi \neq 0$.

We can now define $\pi := \frac{M}{\phi} \neq 0$. Then the LTs selling rate is as in (11) and equation (25) implies that (14) holds. Also, admissibility of the MMs strategies impies $\lambda > 0$, so we can define $\kappa := -\frac{a}{\lambda}$. Then the MMs trading rates are as in (12). Next we will derive the quartic equation for κ and show that the relations (13), (15) and (16) hold. This will be done by looking at the optimality conditions satisfied by the MMs strategies.

As we saw in Corollary 1.2 of Chapter 3, the presence of the non-zero parameter π in the LTs selling rate (11) has no effect on the market makers' optimal strategies. Hence, the argument from Theorem 1.1 in Chapter 3 carries through verbatim[2], and we obtain the following equilibrium consistency conditions for the parameters:

$$-\frac{\rho}{2}\frac{\kappa}{N-1} = \frac{\kappa^2}{N-1} - \frac{\gamma\sigma_D^2}{2N}\frac{1}{\Lambda} \tag{4.26}$$

$$-(\rho+\psi+\phi)\frac{N-2}{N-1}\Lambda = -\Gamma - \frac{\kappa\Lambda}{N-1} \tag{4.27}$$

$$\rho\Big(\frac{N-2}{N-1}\phi\Lambda - \Gamma + \frac{\kappa\Lambda}{N-1}\Big) = -\frac{2\kappa^2\Lambda}{N-1} - \frac{N-2}{N-1}\phi^2\Lambda + \phi\Big(\frac{\kappa\Lambda}{N-1} + \Gamma\Big). \tag{4.28}$$

These are simply the equations (8) - (10) derived in the proof of Theorem 1.1 in Chapter 3, but they are written in terms of κ, Γ and Λ instead of a, c and λ.

Rearranging (26) and (27) gives (15) and (16). Using (27) to substitute for Γ in (28) and rearranging gives (13). Finally, inserting (13), (15), and (16) into the remaining equilibrium relation (24) gives the following quartic equation for κ:

$$\begin{aligned} &\frac{8}{(N-2)^2(\rho+\psi)^2}\kappa^4 + \frac{16\rho}{(N-2)^2(\rho+\psi)^2}\kappa^3 \\ &+\Big(\frac{8\rho^2}{(N-2)^2(\rho+\psi)^2} - \frac{2\rho(3N-4)}{(N-1)(N-2)(\rho+\psi)} - \frac{N\theta}{\gamma}\frac{2}{N-1}\Big)\kappa^2 \\ &+\Big(\frac{\rho}{N-1} - \frac{2\rho^2(3N-4)}{(N-1)(N-2)(\rho+\psi)} - \frac{N\theta}{\gamma}\frac{\rho}{N-1}\Big)\kappa - \frac{N-2}{N-1}\rho\psi = 0. \end{aligned} \tag{4.29}$$

To finish the first half of the proof, it remains to be shown that κ satisfies the inequality constraint in the statement of the theorem. It will be shown momentarily below that the admissibility conditions satisfied by the LTs and MMs strategies imply $\phi, \kappa > -\frac{\rho}{2}$. The inequality for ϕ combined with the relation (13) implies that either $\kappa > -\frac{\rho}{2} + \frac{1}{2}\sqrt{\rho^2 + \rho(\rho+\psi)(N-2)}$ or $\kappa < -\frac{\rho}{2} - \frac{1}{2}\sqrt{\rho^2 + \rho(\rho+\psi)(N-2)}$. Since $\kappa > -\frac{\rho}{2}$, the first inequality must hold, as in the statement of theorem.

Next we show that the admissibility conditions imply $\phi, \kappa > -\frac{\rho}{2}$. First note that the bound on ϕ implies the bound on κ. Indeed, the bound on ϕ implies that $e^{-\rho t}S_t^2$ is

[2] In Chapter 3 we assumed $\phi > 0$, which we have not yet shown. However, this assumption was not used when deriving equilibrium consistency relations between the parameters from the MMs optimality conditions, and that is the part of the argument used here.

integrable for any initial conditions, and admissibility of the MMs strategies implies that the double integral (1) converges absolutely. We can now apply the argument[3] from Lemma 1.5 in Chapter 2 to conclude that $e^{-\rho t}(X_t^n)^2$ is integrable for any initial conditions, and thus $\kappa > -\frac{\rho}{2}$. Hence we need only prove the bound on ϕ.

Now, the admissibility conditions for the LTs strategy imply that the double integral in (2) converges absolutely. The integrand is of the form $e^{-\rho t}Q(S_t, \tilde{S}_t)$ for a quadratic polynomial Q. The coefficient of S_t^2 in Q is $\Gamma\phi - \Lambda\phi^2 - \frac{\theta\sigma_D^2}{2}$. We will need to deal separately with the cases where this coefficient is and isn't zero. If the coefficient is non-zero, then we can apply the argument from Lemma 1.5 in Chapter 2 to conclude that $e^{-\rho t}S_t^2$ is integrable for any initial conditions, and thus $\phi > -\frac{\rho}{2}$.

In the case where the coefficient is 0 we have that

$$\phi(\Gamma - \Lambda\phi) = \frac{\theta\sigma_D^2}{2} > 0. \tag{4.30}$$

We consider two sub-cases, $\phi \geq 0$ and $\phi < 0$. Obviously in the first sub-case we have $\phi > -\frac{\rho}{2}$, so we only need to deal with the second sub-case. In the second sub-case we must have by (30) that $\Gamma - \Lambda\phi < 0$. The MMs admissibility conditions imply that $\Lambda > 0$, and the relation (16) implies that $\Gamma - \Lambda\phi = \frac{1}{N-1}\Lambda\big((N-2)(\rho+\psi) - \phi - \kappa\big)$. Thus it follows that $\kappa > (N-2)(\rho+\psi) - \phi > 0$, and so $e^{-\rho t}(X_t^n)^2$ is integrable for any initial conditions.

Now, consider the double integral (1), which converges absolutely by the MMs admissibility conditions. The integrand is of the form $e^{-\rho t}Q(S_t, \tilde{S}_t, X_t^n)$ for a quadratic polynomial Q with an S_t^2 coefficient of $\frac{1}{N}(\kappa - \phi)(\Gamma - \Lambda\phi)$. This coefficient must be non-zero because $\kappa > 0$, $\phi < 0$ and $\Gamma - \Lambda\phi > 0$. Since we just showed that $e^{-\rho t}(X_t^n)^2$ is integrable for any initial conditions, it follows by the argument from Lemma 1.5 in Chapter 2 that $e^{-\rho t}S_t^2$ is integrable for any initial conditions, and so $\phi > -\frac{\rho}{2}$.

We now turn to the second half of the proof, which requires us to show that if $\kappa > -\frac{\rho}{2} + \frac{1}{2}\sqrt{\rho^2 + \rho(\rho+\psi)(N-2)}$ satisfies (29) then there is a linear symmetric

[3] See the last paragraph of the proof of Lemma 1.5 in Chapter 2. The same argument is used here and twice more below.

Nash equilibrium of the form given in the theorem. To this end, take κ as such and define Γ, Λ, ϕ and π by (13) - (16), Also define $b = 1, \lambda = N\Lambda, a = -N\kappa\Lambda$, $c = -\Gamma + \kappa\Lambda$ and $M = \pi\phi$. We need to show that $a, b, c, \lambda, \phi, M \in \mathbb{R}$ are a linear symmetric Nash equilibrium.

The bound on κ implies that $\lambda > 0$, $\kappa > 0$ and $\phi > -\frac{\rho}{2}$, so the profile given by (7) - (9) is admissible. The MMs optimality condition follows by repeating the proof of Theorem 1.1 in Chapter 3. The LTs optimality condition also follow similarly. The argument is outlined here for completeness.

We can find a second order polynomial $w(d, \tilde{s})$ satisfying the equation

$$\rho w = -\psi \tilde{s} w_{\tilde{s}} + \frac{\sigma_{\tilde{S}}^2}{2} w_{\tilde{s}\tilde{s}} + \frac{\sigma_D^2}{2} w_{dd} + (\Lambda M^2 - \frac{\theta\sigma_D^2}{2})\tilde{s}^2 \tag{4.31}$$

on all of $\mathbb{R}^2$. Now define the function V using formula (22), so that by construction (20) and (21) hold. The first order condition for (20) is sufficent because $\lambda > 0$. The transversality condition for V is satisfied because $\psi > 0$, $\kappa > 0$ and $\phi > -\frac{\rho}{2}$. Hence (17) follows by the verification theorem. □

4.2 Equilibrium Analysis

Finding the zeros of (29) is straightforward using any standard numerical software. As benchmark exogenous parameters we take $N = 10$, $\rho = \gamma = \theta = \sigma_D^2 = \psi = 1$. Extensive numerical simulations, both near and far from these benchmarks, always find a unique zero satisfying the constraint. This suggests that there always exists a unique equilibrium.

There are five endogenous parameters that describe the equilibrium: $\pi, \phi, \kappa, \Gamma$ and Λ. The parameters κ, Γ, and Λ are familiar from Chapters 2 and 3; they describe equilibrium between the market makers taking as given the behavior of the liquidity traders. κ is the rate of convergence to efficiency between the market makers, Γ captures a risk premium, and Λ is the price impact the liquidity traders face when trading. These parameters behave almost identically to the ones in Chapter 2, so for brevity we don't repeat their analysis here.

The endogenous parameters π and ϕ are new in this model, and they describe the trading behavior of the liquidity traders taking as given the pricing function (10) generated by the market makers. These parameters do show up in the model in Chapter 3, but they are exogenous there. Below we study the economic properties of equilibrium between the liquidity traders and market makers by looking at comparative statics of π and ϕ.

Instead of varying each market makers risk aversion γ directly, we will instead vary their aggregate risk aversion $\gamma_0 := \frac{\gamma}{N}$. The benchmark value for γ_0 is $\gamma_0 = 0.1$. Varying N while leaving γ_0 fixed allows us to focus on the effects of competition among the market makers without changing their aggregate risk bearing capacity. Conversely for varying γ_0 while leaving N fixed.

In the plots that follow, unless a parameter is being explicitly varied it's set equal to its benchmark value.

π: The steady state allocation between the LTs and the MMs

Recall that the LTs have an inventory of $\tilde{S}_t - S_t$. The first term comes from trading with clients, and the second term come from selling on the exchange to the market makers. The clients exogenously want to hold $-\tilde{S}_t$ shares of the zero net supply asset, so they trade with the LTs, giving the LTs an inventory of $\tilde{S}_t$. The LTs then turn to the exchange to unload some of this inventory on to the market makers, specifically S_t. Thus the LTs hold $\tilde{S}_t - S_t$ shares and the market makers in aggregate hold S_t shares.

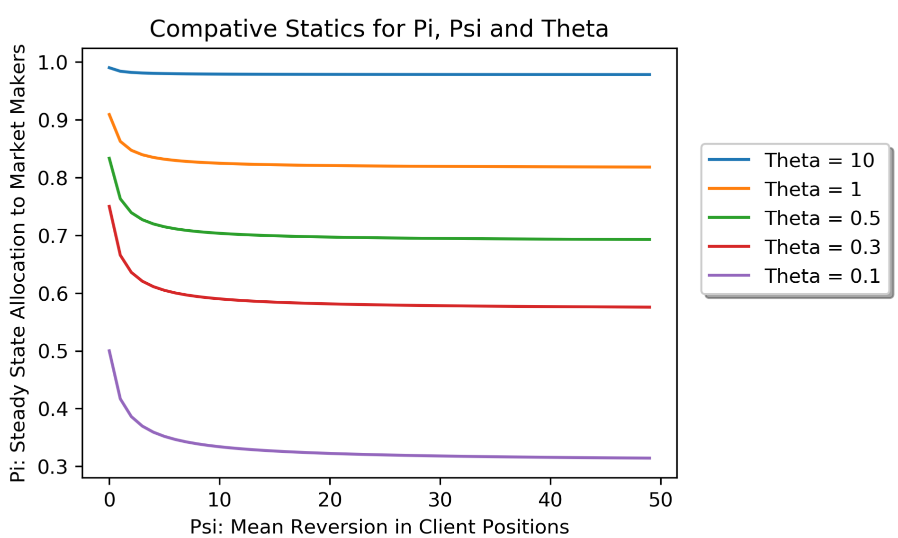
Together the LTs and MMs always have a net shareholding of $\tilde{S}_t$, which is the offsetting position to the exogenous clients. Trading between the LTs and the MMs occurs in order to determine how they share the exposure of $\tilde{S}_t$ between themselves. We should expect there to be a steady state, at which the LTs and MMs have reached an allocation from which they find it optimal not to trade with each other. This is indeed the case, and it's the role of the endogenous parameter π.

Consider the formula for the LTs selling rate in Theorem 1. We see that the LTs

stop trading on the exchange ($\mathcal{N}_t = 0$) when $S_t = \pi \tilde{S}_t$. S_t is the number of shares held by the MMs in aggregate, so they stop trading with the LTs when they hold a fraction π of the clients' offsetting position. We refer to the allocation of $\pi \tilde{S}_t$ shares to the MMs and $(1 - \pi)\tilde{S}_t$ shares to the LTs as the steady state allocation.

Note that no strategy profile can perfectly implement the steady state allocation at all times. This is because the steady state allocation to the market makers $\pi \tilde{S}_t$ evolves as a diffusion in time, whereas the true allocation to the market makers S_t evolves smoothly in time. Instead, in equilibrium allocations are continually hit with shocks, coming from the clients' orders, and trading always moves allocations closer to the steady state. This is discussed in more detail in the next subsection.

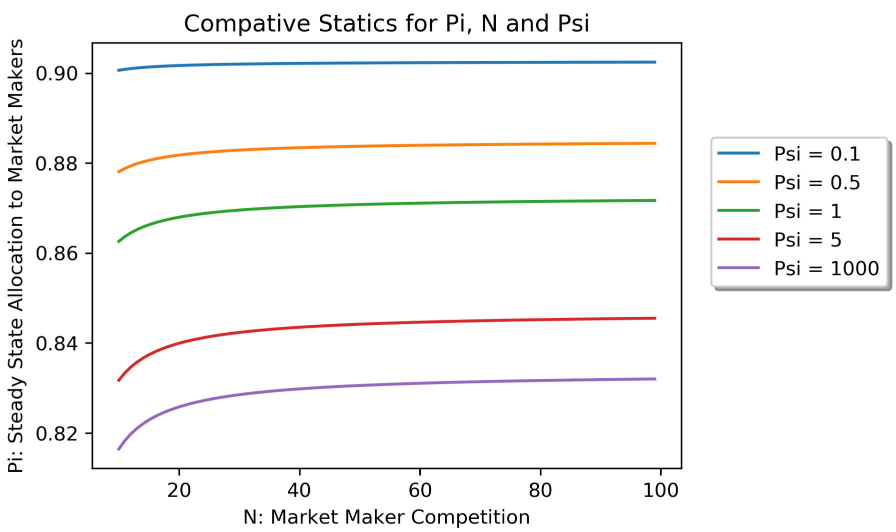
What might one expect π to be? At the very least we should expect $\pi \in [0, 1]$, so that the LTs and MMs always have positions in the same direction, opposite to the clients. The numerics certainly always give this. More specifically, we might expect π to be $\frac{\theta}{\theta+\gamma_0}$. To arrive at this, note that the aggregate risk tolerance of the LTs and the MMs together is $\frac{1}{\gamma_0} + \frac{1}{\theta}$, and the proportion of this coming from the MMs is $\frac{\frac{1}{\gamma_0}}{\frac{1}{\gamma_0}+\frac{1}{\theta}} = \frac{\theta}{\theta+\gamma_0}$. Thus a steady state allocation of $\frac{\theta}{\theta+\gamma_0}\tilde{S}_t$ would be one that is proportional to risk bearing capacity, which tends to be typical in equilibrium models. This is not quite the case though, as the next plot shows.



This plot shows that when $\psi = 0$ it's true that $\pi = \frac{\theta}{\theta+\gamma_0}$, but in general we have

$\pi < \frac{\theta}{\theta+\gamma_0}$, with π decreasing as ψ increases. This means that in general the LTs hold more in the steady state than is proportional to their risk bearing capacity. The reason is that ψ govern the mean reversion in the clients position. Suppose $\psi > 0$ and $\tilde{S}_t > 0$. Then going forward, in the absence of subsequent shocks, one should expect $\tilde{S}_t$ to decrease towards 0. Thus the LTs are happy holding slightly more than their risk bearing capacity, because they expect to shortly be able to reduce their exposure by trading with clients, and this is better than trading on the exchange now and suffering price impact.

Comparing the colored lines in the plot above also shows that π increases as θ increases. This is because the larger θ is, the more risk averse the LTs are, so they hold less in the steady state, and thus the MMs hold more, meaning π is larger. The next plot below shows that as N increases π increases, so the LTs steady state allocation is going down. This is because as MM competition increases, price impact decreases, and price impact is the reason that LTs choose to hold a larger than risk bearing capacity allocation. The effect is more pronounced when ψ is large. This is because the alternative to suffering price impact is waiting for client positions to revert to 0, so changes in price impact are more relevant when the mean reversion is stronger, i.e. changes in N matter more when ψ is large.



ϕ: The rate of reversion to the steady state

The parameter ϕ governs the rate at which equilibrium allocations revert to steady state allocation. The numerics always yield $\phi > 0$, so from (10) it follows that the liquidity traders sell on the exchange, $\mathcal{N}_t > 0$, when they hold more than the steady state allocation, $S_t < \pi\tilde{S}_t$, and vice versa. Thus the steady state is attractive, as one would expect, and ϕ governs the the strength of this attraction. The larger ϕ is, the more aggressively the liquidity traders sell on the exchange when they are above the steady state allocation.

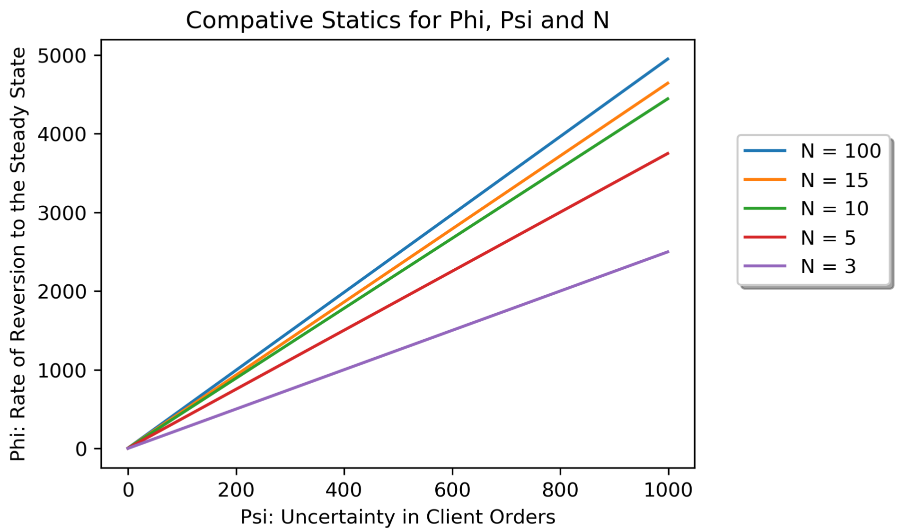
As mentioned before, the steady state cannot be implemented by any strategy profile, which is why we refer to ϕ as a reversion rate as opposed to a convergence rate. One way to think about this is in terms of the equilibrium dynamics of the distance of the system from the steady state. Define $\mathcal{E}_t := \pi\tilde{S}_t - S_t$, which is the inventory in excess to the steady state held by the liquidity traders. When $\mathcal{E}_t > 0$, the liquidity traders would need to sell $\mathcal{E}_t$ to the market makers in order to reach the steady state, and when $\mathcal{E}_t < 0$ they would need to buy $\mathcal{E}_t$.

Using equations (4) and (11) we can solve for the equilibrium dynamics of $\{\mathcal{E}_t\}$. Assuming the system starts at the steady state, i.e. $\mathcal{E}_0 = 0$, we have that

$$\mathcal{E}_t = \int_0^t e^{-\phi(t-u)} d\tilde{S}_u. \tag{4.32}$$

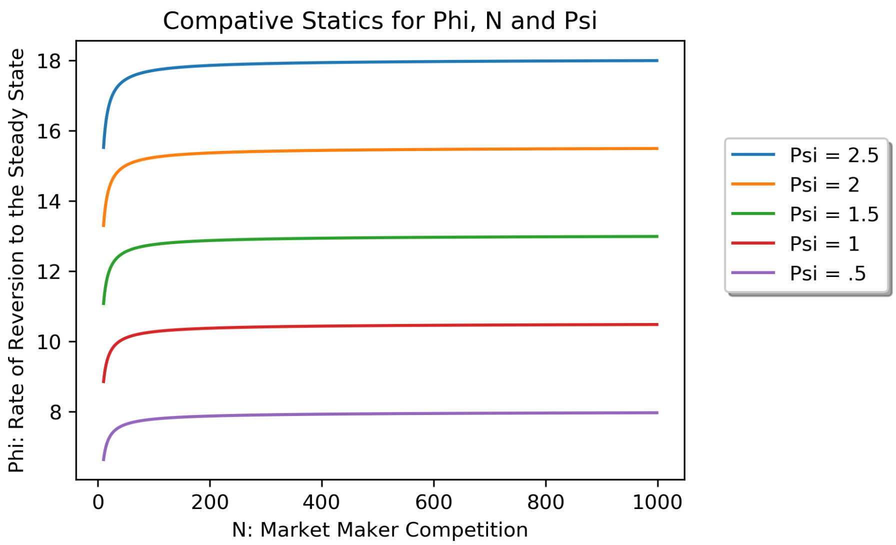
$d\tilde{S}_u$ should be thought of as the random quantity sold by the clients to the liquidity traders at time u. Hence (32) gives $\mathcal{E}_t$ as a weighted average of the history of clients' random orders. The random arrival of client orders is what keeps moving the system away from the steady state, and trading between the liquidity traders and market makers attempts to move it back. One can intuitively think of $\{\mathcal{E}_t\}$ as an Ornstein-Uhlenbeck process with mean reversion parameter ϕ. This is not quite true because $\{\tilde{S}_t\}$ is not a Brownian motion but is itself an Ornstein-Uhlenbeck process. Nonetheless, ϕ governs the amount of mean reversion in $\{\mathcal{E}_t\}$.

So why don't we have $\phi = \infty$? The logic is analogous to the one given in Chapter 2 for why market maker allocations don't converge to efficiency arbitrarily fast. The larger ϕ is, the faster the liquidity traders' inventory reverts to the steady state, but the larger the price impact the liquidity traders suffer. On the other hand, a fortuitous realization of subsequent client orders could automatically move the liquidity traders closer to the steady state. Thus the liquidity traders weight the chances of fortuitous client orders against the costs of price impact in order to pin down a finite value of ϕ. This is illustrated in the plot below.



This plot shows that as ψ increases, ϕ increases, with $\phi \to \infty$ as $\psi \to \infty$. The reason is that ψ governs the randomness in client orders. This can be seen from (6), where the larger ψ is the the stronger the deterministic first term is, and the weaker the stochastic second term is. Said another way, the stationary distribution of $\{\tilde{S}_t\}$ has variance $\frac{\sigma_{\tilde{S}}}{2\psi}$, which is decreasing in ψ. Thus the larger ψ is, the less likely are fortuitous client orders, and so the liquidity traders trade faster, i.e. ϕ is larger. In the limit when $\psi \to \infty$, there is no randomness in client orders, and so there is no reason not to trade arbitrarily fast, i.e. $\phi \to \infty$.

The plot also shows that ϕ increases as N increases. This is because price impact is decreasing in N, and so the downside to trading fast is reduced as N is increased. Looking at the next plot below, we note that $\phi \not\to \infty$ as $N \to \infty$. This is because in the competitive limit price impact does not vanish, and there is still randomness in the clients' orders. In fact, the reason price impact does not vanish is precisely because there is still randomness in client orders. Price impact is the mechanism through which trading prices reflect realizations of clients' order shocks. Thus even in the competitive limit the liquidity traders make a decision between guaranteed trades for price impact with the market makers, or random trades for free with clients. This trade-off pins down a finite value for ϕ.



Chapter 5

CONSUMPTION BASED MODEL

This chapter recasts the model from Chapter 2 into a consumption based setting with a dividend paying asset. This is the standard setting for most asset pricing models in the literature. The point of this chapter is to show that working in a consumption based setting does not significantly alter the results presented earlier in the thesis. As we will see below, the consumption based model is not as analytically tractable, which is why it wasn't used in previous chapters.

The first section lays out the model along the lines of Chapter 2, but with appropriate modifications to accommodate for dividends and consumption. The second section characterizes equilibrium in terms of the zeros of a polynomial system. Some details of the proof are in the appendix to this chapter. The last section analyzes the models equilibrium, showing that the economic properties of the equilibrium in Chapter 2 continue to hold here.

5.1 The Model

Fix a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ equipped with two independent Brownian motions, $\{B_t^D\}$ and $\{B_t^N\}$, and satisfying the usual conditions. We consider a market on an infinite horizon where shares of a zero net supply risky asset are traded for cash (the risk-free asset). Cash is in infinite elastic supply, earns an interest rate[1] of $r > 0$ and is the numeraire. The risky asset is a claim to the dividend

[1]Cash earns no interest in the model in Chapter 2, though it is straightforward to incorporate an interest rate there. In the consumption based setting here it is necessary to include a positive interest otherwise the model is ill-posed [CK93]. We will see below that r in this model plays the role of ρ in Chapter 2's model.

stream[2] $\{D_t\}$ which evolves according to

$$dD_t = \mu dt + \sigma_D dB_t^D, \tag{5.1}$$

where $\mu \in \mathbb{R}$ and $\sigma_D > 0$.

There are two types of traders in the model, $N \in \mathbb{N}$ market makers and a collection of liquidity traders. Trading occurs through a demand schedule auction. At each instant the market makers submit demand schedules (α_t^n, β_t^n) and then their trading rates q_t^n and the trading price p_t are determined by

$$\alpha_t^n - \beta_t^n q_t^n = p_t \qquad \forall n = 1, \cdots, N \tag{5.2}$$

$$q_t^1 + \cdots + q_t^N = \mathcal{N}_t. \tag{5.3}$$

Here $-\mathcal{N}_t$ is the current trading rate of the liquidity traders, interpreted as a market order (vertical line demand schedule) to sell $\mathcal{N}_t dt$ shares over the time interval $[t, t + dt]$.

As in Chapter 2, the liquidity traders' market orders to sell are exogenously given by an Ornstein-Uhlenbeck process

$$d\mathcal{N}_t = -\psi \mathcal{N}_t dt + \sigma_{\mathcal{N}} dB_t^{\mathcal{N}}, \tag{5.4}$$

where $\psi, \sigma_{\mathcal{N}} > 0$. The liquidity traders' time t inventory is denoted $-S_t$, which evolves as a result of trading according to

$$dS_t = \mathcal{N}_t dt. \tag{5.5}$$

The markets makers' time t inventories are denoted by X_t^n, and they evolve as a result of trading according to

$$dX_t^n = q_t^n dt. \tag{5.6}$$

The initial conditions for (6) are the market makers' inventories at the start of the model, $X_0^1, \cdots, X_0^N$. Since the asset is in zero net supply, this also determines

[2] The zero net supply asset should be though of as a futures contract that is perpetually rolled over. The dividends should be thought of as payments to/from a margin account. It's no issue to work with a finite supply asset, thought of as a stock, which is the standard setting for consumption based models. However, the motivation for the thesis is trading of standardized derivatives contracts on exchanges, so we work in the zero net supply setting.

the initial inventory of the liquidity traders, which is the initial condition for (5): $S_0 = X_0^1 + \cdots X_0^N$. Of course we also have $S_t = X_t^1 + \cdots X_t^N \ \forall t > 0$, which follows from (3), (5) and (6). Thus S_t is the supply faced by the market makers at time t.

In addition to demand schedules, market makers also choose at each instant a cash consumption rate $\mathfrak{C}_t^n$, which is their ultimate source of utility. Each market maker's maximization objective is

$$\mathbb{E}\left[-\int_0^\infty e^{-\rho t} e^{-\gamma \mathfrak{C}_t^n} dt\right], \tag{5.7}$$

where $\rho, \gamma > 0$. The amount of cash held by each market maker at time t is denoted by M_t^n, and it evolves according to

$$dM_t^n = rM_t^n dt + X_t^n D_t - p_t q_t^n dt - \mathfrak{C}_t^n dt. \tag{5.8}$$

Here the first term comes from the interest rate on cash, the second term comes from dividends on the risky asset, the third term comes from trading and the fourth term comes from consumption.

This sets up the model as a stochastic differential game between the N market makers. Below we formulate the equilibrium concept analogously to Chapter 2. At first glance one might be surprised that the market makers' objectives (7) depend only on their choice of consumption streams, and not on any other state variables. However, when formulating the equilibrium concept below we include a No-Ponzi condition, as is typical in the literature [Bac17] [KOW18]. This condition ties the other state variables into the choice of an optimal consumption stream.

Equilibrium Concept

Definition 5.1.1. Given initial conditions $(\vec{x}, \eta, d) \in \mathbb{R}^N \times \mathbb{R} \times \mathbb{R}$ for $(X_0^1, \cdots, X_0^N)$, $\mathcal{N}_0$, and D_0, we say that the profile of progressively measurable demand schedules $\{(\alpha_t^1, \beta_t^1)\}, \cdots, \{(\alpha_t^N, \beta_t^N)\}$ and consumption streams $\{\mathfrak{C}_t^1\}, \cdots, \{\mathfrak{C}_t^N\}$ is **admissible starting from** $(\vec{x}, \eta, d)$ if:

1. $\beta_t^n > 0 \ \forall t \geq 0, \ \forall n = 1, \cdots, N$ almost surely

2. $\int_0^T |q_t^n| dt < \infty \; \forall T \geq 0, \; \forall n = 1, \cdots, N$ almost surely

3. $\mathbb{E}\left[-\int_0^\infty e^{-\rho t} e^{-\gamma \mathfrak{C}_t^n} dt \right] > -\infty \; \forall n = 1, \cdots, N$

4. $\alpha_t^n, \beta_t^n, \mathfrak{C}_t^n \in \sigma\Big(\{D_s\}_{0 \leq s \leq t}, \{p_s\}_{0 \leq s < t}, \{X_s^n\}_{0 \leq s \leq t}, \{M_s^n\}_{0 \leq s \leq t}, S_0 \Big) \; \forall t \geq 0, \; \forall n = 1, \cdots N$

5. $\liminf_{T \to \infty} \mathbb{E}_t \left[e^{-r(T-t)} M_T^n + \int_T^\infty e^{-r(u-t)} X_u^n D_u du \right] \geq 0 \; \forall t \geq 0, \; \forall n = 1, \cdots, N.$

Nash equilibrium is defined as usual. The admissibility requirement is given by the definition above. Optimality is required for both the choice of consumption streams and the choice of demand schedules. As before we will only by concerned with equilibria where the market makers submit demand schedules that have a linear and symmetric structure. We don't place any explicit restrictions on the nature of the consumption streams in equilibrium.

Definition 5.1.2. Demand schedules $\{(\alpha_t^1, \beta_t^1)\}, \cdots, \{(\alpha_t^N, \beta_t^N)\}$ are said to be **linear symmetric** if $\exists \; a, \lambda, b, c, \xi \in \mathbb{R}$ s.t.

$$\alpha_t^n = a X_t^n + b D_t + c S_t + \xi \tag{5.9}$$

$$\beta_t^n = \lambda \tag{5.10}$$

$\forall t \geq 0, \; \forall n = 1, \cdots, N.$

Definition 5.1.3. We say that $a, \lambda, b, c, \xi \in \mathbb{R}$ are a **linear symmetric Nash equilibrium** if for any initial conditions there exist consumption streams that when used with the linear symmetric demand schedules (9)-(10) give a Nash equilibrium.

In previous chapters we saw that focusing on linear symmetric equilibria allows us to formulate the equilibrium conditions in terms of a single optimization problem over a trading rate. An analogue of this continues to hold here, as stated in the next proposition. The proof is omitted because it's identical to the one in Chapter 2. The presence of the consumption streams does not change anything; only the linear symmetric structure of the demand schedules is of relevance.

Proposition 5.1.4. *$a, \lambda, b, c, \xi \in \mathbb{R}$ are a linear symmetric Nash equilibrium if and only if $\lambda > 0$ and for any initial conditions $\exists\{\hat{\mathfrak{C}}_t\}$ such that*

$$\frac{a}{\lambda}\Big(X_t - \frac{S_t}{N}\Big) + \frac{\mathcal{N}_t}{N}, \{\hat{\mathfrak{C}}_t\} \in \underset{\{q_t\},\{\mathfrak{C}_t\}}{\arg\max}\, \mathbb{E}\bigg[-\int_0^\infty e^{-\rho t} e^{-\gamma \mathfrak{C}_t} dt\bigg],$$

where

$$p_t = \Big(\frac{a}{N-1} + c\Big)S_t + bD_t + \xi - \frac{\lambda}{N-1}\mathcal{N}_t - \frac{a}{N-1}X_t + \frac{\lambda}{N-1}q_t$$

the relevant dynamics are

$$\begin{aligned} dX_t &= q_t dt \\ dD_t &= \mu dt + \sigma_D dB_t^D \\ d\mathcal{N}_t &= -\psi \mathcal{N}_t dt + \sigma_{\mathcal{N}} dB_t^{\mathcal{N}} \\ dS_t &= \mathcal{N}_t dt \\ dM_t &= rM_t dt + X_t D_t dt - p_t q_t dt - \mathfrak{C}_t dt \end{aligned}$$

and the optimization is constrained to those processes $\{q_t\}, \{\mathfrak{C}_t\}$ such that

1. $\int_0^T |q_t| dt < \infty\ \forall T \geq 0$ *almost surely*
2. $\mathbb{E}\Big[-\int_0^\infty e^{-\rho t} e^{-\gamma \mathfrak{C}_t} dt\Big] > -\infty$
3. $q_t, \mathfrak{C}_t \in \sigma\Big(\{D_u\}_{0\leq u\leq t}, \{X_u\}_{0\leq u\leq t}, \{M_u\}_{0\leq u\leq t}, \{S_u\}_{0\leq u\leq t}, \{\mathcal{N}_u\}_{0\leq u\leq t}\Big)\ \forall t \geq 0$
4. $\liminf_{T\to\infty} \mathbb{E}_t\left[e^{-r(T-t)}M_T + \int_T^\infty e^{-r(u-t)} X_u D_u du\right] \geq 0\ \forall t \geq 0.$

5.2 Equilibrium Characterization

Theorem 5.2.1. *Fix exogenous parameters $N \geq 3$, $\rho, \gamma, r, \sigma_D, \psi, \sigma_{\mathcal{N}} > 0$ and $\mu \in \mathbb{R}$. If $\kappa, \Lambda > 0$ and $F_1 \in \mathbb{R}$ satisfy equations (33) - (35) in the appendix, then there is a linear symmetric equilibrium with trading price*

$$p_t = \frac{1}{r}D_t + \frac{\mu}{r^2} - \Gamma S_t - \Lambda \mathcal{N}_t \tag{5.11}$$

and trading rates

$$q_t^n = -\kappa\big(X_t^n - \frac{S_t}{N}\big) + \frac{\mathcal{N}_t}{N}. \tag{5.12}$$

A formula for the parameter Γ *in terms of* κ, Λ *and* F_1 *is given in equation (32) in the appendix.*

Proof. Consider the optimization problem in Proposition 1.4 associated to a linear symmetric profile with parameters a, b, c, ξ and $\lambda > 0$. The is a standard Markovian stochastic control problem on an infinite horizon with state space $(m, d, s, \eta, x) \in \mathbb{R}^5$ and control space $(q, c) \in \mathbb{R}^2$. The Hamiltonian for the problem is

$$H(m, d, s, \eta, x, y, z, q, c) = -e^{-\gamma c} + rmy_m - cy_m + xdy_m - qP(d, s, \eta, x, q)y_m + qy_x + \mu y_d + \frac{\sigma_D^2}{2} z_{dd} + \eta y_s - \psi \eta y_\eta + \frac{\sigma_N^2}{2} z_{\eta\eta}.$$

Here $P : \mathbb{R}^5 \longrightarrow \mathbb{R}$ is the the function specifying prices in terms of state and control in Proposition 1.4, i.e.

$$P(d, s, \eta, x, q) = \Big(\frac{a}{N-1} + c\Big)s + bd + \xi - \frac{\lambda}{N-1}\eta - \frac{a}{N-1}x + \frac{\lambda}{N-1}q.$$

Also denote by $Q : \mathbb{R}^3 \longrightarrow \mathbb{R}$ the function specifying the optimal trading rate in feedback form in Proposition 1.4, i.e.

$$Q(s, \eta, x) = \frac{a}{\lambda}\big(x - \frac{s}{N}\big) + \frac{\eta}{N}.$$

By the verification theorem, to show that the optimality condition in Proposition 1.4 is satisfied, it suffices to find a smooth function $V(m, d, s, \eta, x)$ and a measurable function $\hat{\mathfrak{C}}(m, d, s, \eta, x)$ such that

$$\rho V(m, d, s, \eta, x) = \sup_{q,c\in\mathbb{R}} H(m, d, s, \eta, x, \nabla V, \nabla^2 V, q, c) \tag{5.13}$$

$$Q(s, \eta, x), \hat{\mathfrak{C}}(m, d, s, \eta, x) \in \arg\max_{q,c\in\mathbb{R}} H(m, d, s, \eta, x, \nabla V, \nabla^2 V, q, c) \tag{5.14}$$

$\forall (m, d, s, \eta, x) \in \mathbb{R}^5$. Furthermore, the transversality condition for V and the admissibility conditions for the Markov controls given by Q and $\hat{\mathfrak{C}}$ must hold.

Now, to demonstrate that (13) and (14) hold, it suffices to find a function smooth function $w(d, s, \eta, x)$ such that

$$rw = \mu w_d + \eta w_s - \psi \eta w_\eta + \frac{\sigma_D^2}{2}\big(w_{dd} - r\gamma (w_d)^2\big) + \frac{\sigma_N}{2}\big(w_{\eta\eta} - r\gamma (w_\eta)^2\big) + xd + \frac{\rho}{r\gamma} + \frac{\log r}{\gamma} - \frac{1}{\gamma} + \frac{\lambda}{N-1}\Big(\frac{a}{\lambda}(x - \frac{s}{N}) + \frac{\eta}{N}\Big)^2 \tag{5.15}$$

$$w_x(d,s,\eta,x) = bd + \xi + \left(\frac{N-2}{N(N-1)}a + c\right)s - \frac{N-2}{N(N-1)}\lambda\eta + \frac{a}{N-1}x \qquad (5.16)$$

$\forall(d,s,\eta,x) \in \mathbb{R}^4$. Indeed then we can take

$$V(m,d,s,\eta,x) = -e^{-r\gamma\left(m+w(d,s,\eta,x)\right)}$$
$$\hat{\mathfrak{C}}(m,d,s,\eta,x) = -\frac{\log r}{\gamma} + rm + rw(d,s,\eta,x)$$

to satisfy (13) and (14). For (14) we use the first order conditions, which are sufficient because $\gamma > 0$ and $\lambda > 0$.

Furthermore, if w is a second order polynomial and $\frac{a}{\lambda} < 0$, then the admissibility and transversality conditions hold. Also, antidifferentiating shows that to satisfy (16) w must be of the form

$$w(d,s,\eta,x) = bxd + \xi x + \left(\frac{N-2}{N(N-1)}a + c\right)xs - \left(\frac{N-2}{N(N-1)}\lambda\right)x\eta + \frac{1}{2}\frac{a}{N-1}x^2 + h(d,s,\eta) \qquad (5.17)$$

for some function $h : \mathbb{R}^3 \to \mathbb{R}$.

Combining the arguments in the last two paragraphs we have proven the following statement. If $a < 0$, $\lambda > 0$, and $b, c, \xi \in \mathbb{R}$ are such that there exists a second order polynomial h that satisfies (15) and (17), then a, b, c, ξ, λ are a linear symmetric equilibrium.

In the appendix we show that if $\kappa, \Lambda > 0$ and $F_1 \in \mathbb{R}$ satisfy equations (33)-(35), Γ is defined as in (32), and

$$a = -N\kappa\Lambda \qquad (5.18)$$
$$b = \frac{1}{r} \qquad (5.19)$$
$$c = \kappa\Lambda - \Gamma \qquad (5.20)$$
$$\xi = \frac{\mu}{r^2} \qquad (5.21)$$
$$\lambda = N\Lambda \qquad (5.22)$$

then such a second order polynomial h can be found. Thus the parameters above give a linear symmetric equilibrium, and the formulas in the theorem follow by computing equilibrium prices and trading rates using (2), (3), (9) and (10).

□

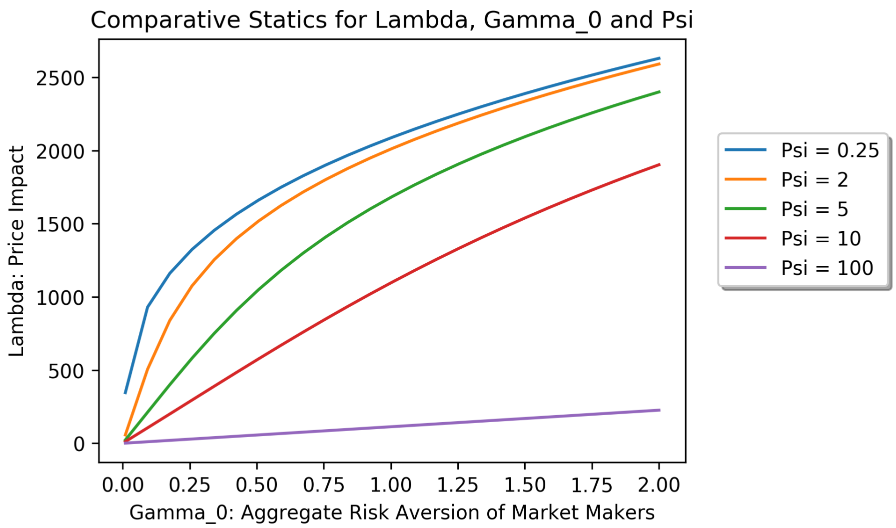
5.3 Equilibrium Analysis

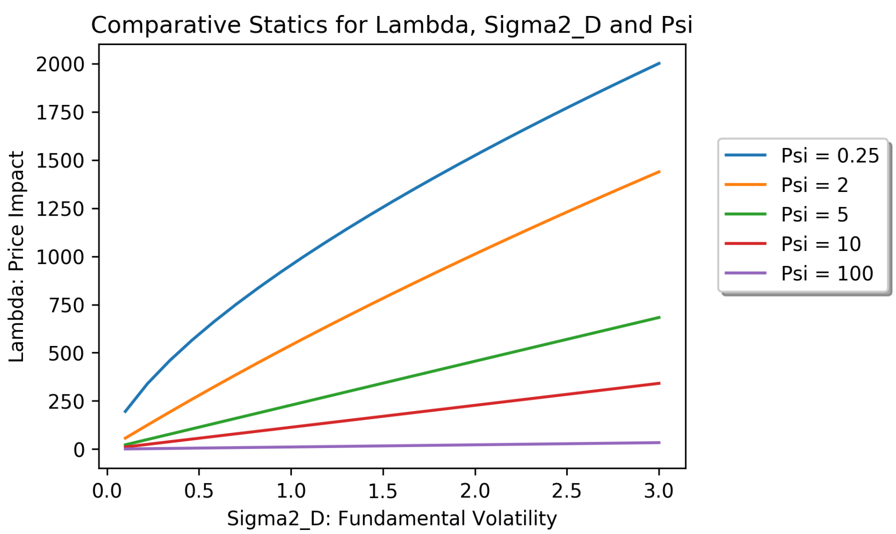
The equilibrium equations are in the form of a system of multivariate polynomial equations. There does not appear to be a way to reduce these equations down to a single one-dimensional polynomial equation. In general, solving such systems is not particularly tractable, but there are a variety of numerical methods that tend to work. Here we use the Homotopy Continuation Method, implemented with the open source software PHCpack [Ver14].

This method works by starting with a proxy polynomial system that is easy to solve, forming a homotopy between the coefficients of the proxy system and the actual system of interest, and then tracking the solutions over the course of the homotopy. The software automatically chooses the proxy system, but one can also choose one manually. A natural choice for a proxy polynomial system would be the one derived in Chapter 2, which we were able to solve explicitly. Using this as the initial proxy gives the same results as allowing the software to choose its own initial proxy. However, the software works slightly faster when allowed to choose its own proxy, so that was the methodology used.

As benchmark parameters we take $r = .01$, $N = 10$ and $\gamma = \psi = \sigma_D^2 = \sigma_N^2 = 1$. Extensive numerical simulations, both near and far from these benchmarks, always find a unique solution satisfying the constraint. This suggests that there always exists an equilibrium of the form stated in the theorem. Close to the benchmarks the software finds the solution very fast, and the results are very stable over repeated runs with same parameters. When very far from these benchmarks the software can be slow, taking sometimes a couple of minutes to solve a thousand systems. There is also slight numerical instability when very far from the benchmarks. Roughy 1 in every 100 runs of the software on the same parameters fails to find any solutions. The remaining runs always give the same solution.

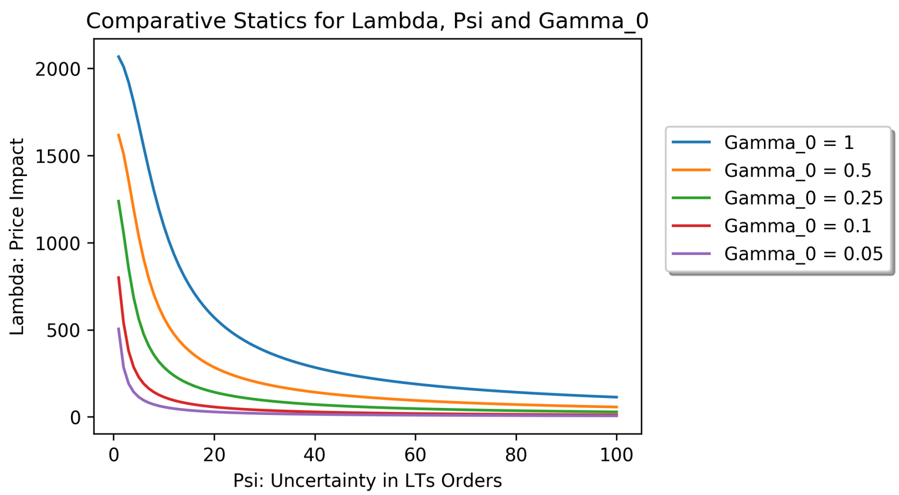
In the plots that follow, unless a parameter is being explicitly varied it's set equal to its benchmark value. Instead of varying each market maker's risk aversion γ directly, we will instead vary their aggregate risk aversion $\gamma_0 := \frac{\gamma}{N}$. The benchmark value for γ_0 is $\gamma_0 = 0.1$. Varying N while leaving γ_0 fixed allows us to focus on the effects of competition among the market makers without changing their aggregate risk bearing capacity. Conversely for varying γ_0 while leaving N fixed.

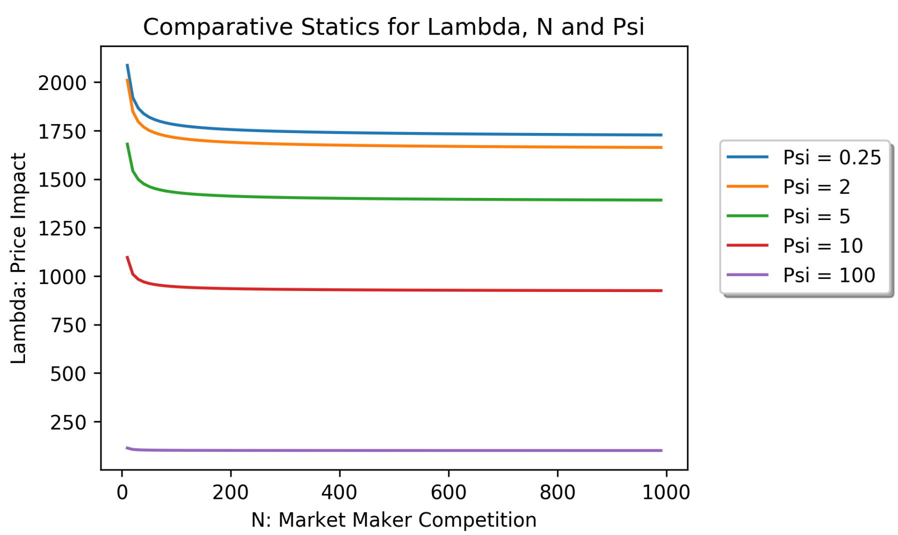




The first result is that price impact is increasing in the market makers' risk aversion, as can be seen in the first plot above. The second result is that price impact is increasing in fundamental volatility, as can be seen in the second plot above.

The next two results are more interesting and are shown in the two plots below. The first plot shows that price impact is decreasing in ψ. Here ψ is thought of as governing the uncertainty in the liquidity traders' orders. When ψ is large there is very little uncertainty, and when ψ is small there is a lot of uncertainty. The second plot shows that price impact is decreasing in market maker competition.[3]



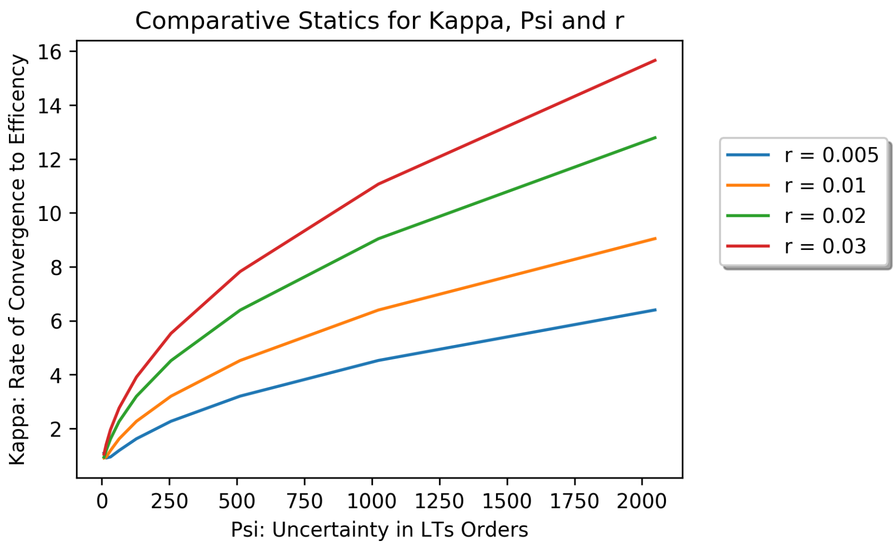


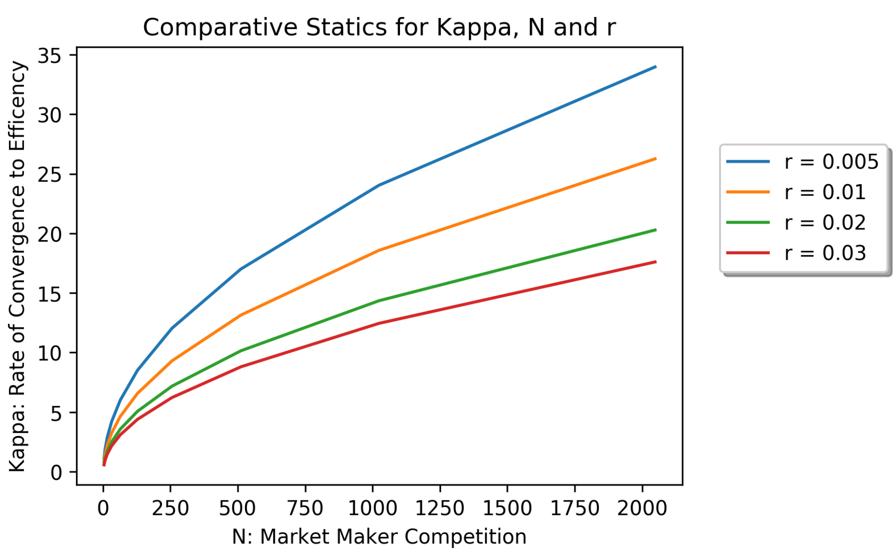
The two plots above also show that the limiting behavior of price impact discussed in Chapter 2 continues to hold. From the second plot we see that price impact does not vanish in the competitive limit. The reason is that $\psi \neq \infty$, and thus there is

[3] In this plot we don't use the benchmark value of γ_0 but instead take $\gamma_0 = 1$. This is purely for visualization purposes. When $\gamma_0 = .1$ and we vary ψ the lines all look flat due to the scale of the y-axis. The pattern of decreasing to a non-zero limit holds regardless of what parameters are chosen.

uncertainty in the liquidity traders' orders. In the first plot we see that price impact vanishes in the limit as $\psi \to \infty$.

We next turn to the rate of convergence to efficiency between the market makers. Comparative statics and limiting behavior can be seen in the two plots below.





Three key properties that were seen in Chapter 2 continue to hold here. Firstly, the rate of convergence goes to infinity as N or ψ go to infinity. Secondly, the rate of convergence is concave as a function of N and as a function of ψ. In Chapter 2 the concavity was specifically of the square root form. Finally, the rate of convergence depends, for the most part, only on the parameters N, ψ and r. Indeed the plots show the dependence on these parameters. We verified that there is essentially no change in these plots when varying other parameters; the graphs all sit right on top of each other and are indiscernible visually. The actual numerical values do depend ever so slightly on other parameters. This is analogous to the formula for the rate of convergence in Chapter 2, which only features N, ψ and r.

5.4 Appendix: Equilibrium Equations

We will show that a second order polynomial of the form $h(d,s,\eta) = F_1\eta^2 + F_2\eta s + F_3 s^2 + F_4$ can satisfy equations (15) and (17) for certain values of a, b, c, ξ and λ. Inserting this functional form for h into equations (15) and (17) yields an equality between two second order polynomials in (d,s,η,x). Equating coefficients of the polynomials gives the following system of nine equations.

$$rb = 1 \tag{5.23}$$

$$r\xi = \mu b \tag{5.24}$$

$$r\Big(\frac{N-2}{N(N-1)}a + c\Big) = r\gamma\sigma_N^2\frac{N-2}{N(N-1)}\lambda F_2 - \frac{2}{N(N-1)}\frac{a^2}{\lambda} \tag{5.25}$$

$$-r\frac{N-2}{N(N-1)}\lambda = \frac{1}{N-1}a + c + \psi\frac{N-2}{N(N-1)}\lambda + 2r\gamma\sigma_N^2\frac{N-2}{N(N-1)}\lambda F_1 \tag{5.26}$$

$$\frac{r}{2}\frac{a}{N-1} = -\frac{\sigma_D^2}{2}r\gamma b^2 - \frac{\sigma_N^2}{2}r\gamma\Big(\frac{N-2}{N(N-1)}\Big)^2\lambda^2 + \frac{1}{N-1}\frac{a^2}{\lambda} \tag{5.27}$$

$$rF_1 = F_2 - 2\psi F_1 - 2r\gamma\sigma_N^2 F_1^2 + \frac{1}{N^2(N-1)}\lambda \tag{5.28}$$

$$rF_2 = 2F_3 - \psi F_2 - 2r\gamma\sigma_N^2 F_1 F_2 - \frac{2}{N^2(N-1)}a \tag{5.29}$$

$$rF_3 = -\frac{r\gamma\sigma_N^2}{2}F_2^2 + \frac{1}{N^2(N-1)}\frac{a^2}{\lambda} \tag{5.30}$$

$$rF_4 = \frac{\rho}{r\gamma} + \frac{\log r}{\gamma} - \frac{1}{\gamma} + \sigma_N^2 F_1 \tag{5.31}$$

Equations (23) and (24) are equivalent to (19) and (21). Equations (28), (30) and (31) specify F_2, F_3 and F_4 in terms of F_1, a and λ. Next we change variables from a, λ and c to κ, Λ and Γ using (18), (20) and (22). Then equation (25) gives Γ in terms of κ, Λ and F_1 as

$$\Gamma = \frac{2}{r(N-1)}\kappa^2\Lambda + \frac{1}{N-1}\kappa\Lambda - \gamma\sigma_N^2\frac{N-2}{N-1}\Lambda\Big((r+2\psi)F_1 + 2r\gamma\sigma_N^2 F_1^2 - \frac{1}{N(N-1)}\Lambda\Big). \tag{5.32}$$

Finally, we are left with the following system for κ, Λ and F_1.

$$-\frac{r}{2}\frac{N}{N-1}\kappa\Lambda + \frac{\sigma_D^2}{2}\frac{\gamma}{r} + \frac{\sigma_N^2}{2}r\gamma\Big(\frac{N-2}{N-1}\Big)^2\Lambda^2 - \frac{N}{N-1}\kappa^2\Lambda = 0 \tag{5.33}$$

$$-\frac{2}{r(N-1)}\kappa^2\Lambda - \frac{2}{N-1}\kappa\Lambda + \sigma_N^2\gamma\frac{N-2}{N-1}F_2\Lambda + 2r\sigma_N^2\gamma\frac{N-2}{N-1}F_1\Lambda + (r+\psi)\gamma\frac{N-2}{N-1}\Lambda = 0 \tag{5.34}$$

$$-\gamma\sigma_N^2 F_2^2 + \frac{2}{rN(N-1)}\kappa^2\Lambda - (\psi + r)F_2 - 2r\gamma\sigma_N^2 F_1 F_2 + \frac{2}{N(N-1)}\Lambda = 0. \tag{5.35}$$

In this system one needs to plug in F_2 as

$$F_2 = (r + 2\psi)F_1 + 2r\gamma\sigma_N^2 F_1^2 - \frac{1}{N(N-1)}\Lambda.$$

DIRECTIONS FOR FUTURE RESEARCH

This thesis leaves open numerous directions for future research. Firstly, the model in Chapter 4 features clients whose flow the liquidity traders clear at the theoretical value. This is inspired by the practice of traders buying flow from clients, though typically this flow is not cleared at a theoretical value, but at the prevailing mid-price on the exchange. It should be straightforward to consider such an extension of the model. Secondly, it would be interesting to explicitly model the buying of flow from brokers/clients. It's probably difficult to obtain an equilibrium model where traders pay a positive price for broker flow, which is what occurs in reality.

More generally, there has been an interesting evolution in market microstructure that could potentially be modeled using the techniques in this thesis. It used to be the case that market makers traded on the exchange, and certain broker-traders executed orders on the exchange on behalf of clients. This is roughly the setting modeled in this thesis, with the liquidity traders playing the role of the broker-traders. However, nowadays these broker-traders are just brokers, and they sell their clients' flow directly to market makers. The bulk of trading on the exchange now takes place between market makers, after they have netted the various flows they have purchased from brokers. There are also some sophisticated clients, typically hedge funds, who no longer use brokers and instead trade directly on the exchange with market makers.

This is arguably the result of specialization. The brokers have specialized in client services, and the market makers have specialized in trading technology. This has probably raised broker fees to the point where hedge funds prefer to also invest in trading technology. It would be interesting to model this evolution by extending the models in this thesis, perhaps by allowing agents to invest in trading and client service technologies. One place to start looking for ideas to build a model such as this could be [BFM15].

BIBLIOGRAPHY

[AC01] Robert Almgren and Neil Chriss. "Optimal Execution of Portfolio Transactions". *The Journal of Risk* 3.2 (Jan. 2001), pp. 5–39.

[Bac17] Kerry Back. *Asset Pricing and Portfolio Choice Theory*. Second edition. Financial Management Association Survey and Synthesis Series. Oxford ; New York: Oxford University Press, 2017. 722 pp. ISBN: 978-0-19-024114-8.

[Bar+15] Nicholas Barberis et al. "X-CAPM: An Extrapolative Capital Asset Pricing Model". *Journal of Financial Economics* 115.1 (Jan. 1, 2015), pp. 1–24.

[BFM15] Bruno Biais, Thierry Foucault, and Sophie Moinas. "Equilibrium Fast Trading". *Journal of Financial Economics* 116.2 (May 1, 2015), pp. 292–313.

[Bou+18] Bruno Bouchard et al. "Equilibrium Returns with Transaction Costs". *Finance and Stochastics* 22.3 (July 1, 2018), pp. 569–601.

[Car16] R. Carmona. *Lectures on BSDEs, Stochastic Control, and Stochastic Differential Games with Financial Applications*. Philadelphia: Society for Industrial and Applied Mathematics, 2016. 265 pp. ISBN: 978-1-61197-423-2.

[CJP15] Álvaro Cartea, Sebastian Jaimungal, and José Penalva. *Algorithmic and High-Frequency Trading*. Cambridge, United Kingdom: Cambridge University Press, 2015. 343 pp. ISBN: 978-1-107-09114-6.

[CK93] John Y. Campbell and Albert S. Kyle. "Smart Money, Noise Trading and Stock Price Behaviour". *The Review of Economic Studies* 60.1 (Jan. 1, 1993), pp. 1–34.

[Coc05] John H. Cochrane. *Asset Pricing*. Rev. ed. Princeton, N.J: Princeton University Press, 2005. 533 pp. ISBN: 978-0-691-12137-6.

[FPR13] Thierry Foucault, Marco Pagano, and Ailsa Röell. *Market Liquidity: Theory, Evidence, and Policy*. Oxford ; New York: Oxford University Press, 2013. 424 pp. ISBN: 978-0-19-993624-3.

[GM88] Sanford J. Grossman and Merton H. Miller. "Liquidity and Market Structure". *The Journal of Finance* 43.3 (1988), pp. 617–633.

[GP16] Nicolae Gârleanu and Lasse Heje Pedersen. "Dynamic Portfolio Choice with Frictions". *Journal of Economic Theory* 165 (Sept. 1, 2016), pp. 487–516.

[Gué16] Olivier Guéant. *The Financial Mathematics of Market Liquidity: From Optimal Execution to Market Making*. Chapman & Hall/CRC Financial Mathematics Series. Boca Raton: CRC Press, Taylor & Francis Group, 2016. 278 pp. ISBN: 978-1-4987-2547-7.

[KOW18] Albert S. Kyle, Anna A. Obizhaeva, and Yajun Wang. "Smooth Trading with Overconfidence and Market Power". *The Review of Economic Studies* 85.1 (Jan. 1, 2018), pp. 611–662.

[Kyl89] Albert S. Kyle. "Informed Speculation with Imperfect Competition". *The Review of Economic Studies* 56.3 (1989), pp. 317–355.

[OHa11] Maureen O'Hara. *Market Microstructure Theory*. repr. Malden, Mass.: Blackwell, 2011. 290 pp. ISBN: 978-0-631-20761-0.

[Pha09] Huyên Pham. *Continuous-Time Stochastic Control and Optimization with Financial Applications*. Stochastic Modelling and Applied Probability 61. Berlin: Springer, 2009. 232 pp. ISBN: 978-3-540-89499-5.

[SS16] Yuliy Sannikov and Andrzej Skrzypacz. "Dynamic Trading: Price Inertia and Front-Running". *Working Paper* (2016).

[Tou13] Nizar Touzi. *Optimal Stochastic Control, Stochastic Target Problems, and Backward SDE*. Fields Institute Monographs v. 29. New York: Springer, 2013. 214 pp. ISBN: 978-1-4614-4286-8.

[Ver14] Jan Verschelde. "Modernizing PHCpack through Phcpy" (Apr. 29, 2014). arXiv: `1310.0056 [cs, math]`.

[Wil79] Robert Wilson. "Auctions of Shares". *The Quarterly Journal of Economics* 93.4 (1979), pp. 675–689.