2 A Primer on the Microstructure of Financial Markets

To understand the issues and problems faced in the design and implementation of trading strategies, we must consider the economics that drive these trading strategies. To do this we look to the market microstructure literature. Section 2.1 considers the basic market making model that focuses on inventory and inventory risk, as well as the trade-off between execution frequency and profit per trade. It also looks at the conceptual basis for the basic measures of liquidity. The last two sections look at trading when there are informational differences between traders. Section 2.2 from the point of view of the better informed trader, and Section 2.3 from that of the less informed market maker.

For the ecomomics of trading we look to market microstructure, as it is the subfield of finance which focuses on how trading takes place in very specific settings: it "is the study of the process and outcomes of exchanging assets under explicit trading rules" (O'Hara (1995)). Thus, it encompasses the subject of this book, algorithmic and high-frequency trading. It is within the microstructure literature that we find studies of the process of exchanging assets: trading strategies, and their outcomes: asset prices, volume, risk transfers, etc.

A key dimension of the trading and price setting process is that of information. Who has what information, how does that information affect trading strategies, and how do those trading strategies affect trading outcomes in general, and asset prices in particular. Forty years ago finance theory introduced the tools to explicitly incorporate and evaluate the notion of price efficiency, the idea that "market prices are an efficient way of transmitting the information required to arrive at a Pareto optimal allocation of resources" (Grossman & Stiglitz (1976)). This dimension naturally appears in microstructure studies which look into the details of how different trading rules and trading conditions incorporate or hinder price efficiency. What differentiates microstructure studies from more general asset pricing ones is that they focus on two aspects that are key to trading: liquidity and price discovery, and these are the two primary aspects that drive the questions and issues behind the design of effective algorithmic and high-frequency trading.[1]

Trading can take place in a number of possible ways: via personal deals settled over a handshake in a club, via decentralised chat rooms where traders engage

[1] Abergel, Bouchaud, Foucault, Lehalle & Rosenbaum (2012) provides a general overview of the determinants and effects of liquidity in security markets and related policy issues.

each other in bilateral personal transactions, via broker-intermediated over-the-counter (OTC) deals, via specialised broker-dealer networks, on open electronic markets, etc. Our focus is on trading and trading algorithms that take place in large electronic markets, whether they be open exchanges, such as the NASDAQ stock market, or in electronic private exchanges (run by a broker-dealer, a bank, or a consortium of buy-side investors).

2.1 Market Making

As we saw in Chapter 1, an important type of market participant is the 'passive' market maker (MM), who facilitates trade and profits from making the spread and from her execution skills, and must be quick to adapt to changing market conditions. Another type is the 'active' trader, who exploits her ability to anticipate price movements and must identify the optimal timing for her market intervention. We start with the first group, the 'passive' traders.

Because we are focusing on trading in active exchanges, it is natural to assume that there are many market makers (MMs) in competition. Naturally, trading in a market dominated by a few MMs would need to additionally incorporate how the MMs exercise their market power and how it affects the market as a whole.

MMs play a crucial role in markets where they are responsible for providing liquidity to market participants by quoting prices to buy and sell the assets being traded, whether they be equities, financial derivatives, commodities, currencies, or others. A key dimension of liquidity as provided by MMs is immediacy: the ability of investors to buy (or sell) an asset at a particular point in time without having to wait to find a counterparty with an offsetting position to sell (or buy). By quoting buy and sell prices (or posting limit orders (LOs) on both sides of the book), the MM is willing to provide liquidity to the market, but in order to make this a sustainable business the MM quotes a buy price lower than her quoted sell price. For example an MM is willing to purchase shares of company XYZ at \$99 and willing to sell at \$101 per share. Note that by posting LOs, the MM is providing liquidity to other traders who may be looking to execute a trade quickly, e.g. by entering a market order (MO). Hence, we have the usual dichotomy that separates MMs as liquidity providers from other traders, considered as liquidity takers.

If our MM is the one offering the best prices, so that the ask is \$101 and the bid \$99, then the quoted spread is \$2. There are a number of theories that explain what determines the spread in a competitive market. Before delving into some of these theories, we consider the issues faced by someone willing to provide liquidity.

2.1.1 Grossman–Miller Market Making Model

The first issue faced by an MM when providing liquidity is that by accepting one side of a trade (say buying from someone who wants to sell), the MM will hold an asset for an uncertain period of time, the time it takes for another person to come to the market with a matching demand for liquidity (wanting to buy the asset the MM bought in the previous trade). During that time, the MM is exposed to the risk that the price moves against her (in our example, as she bought the asset, she is exposed to a price decline and hence having to sell the asset at a loss in the next trade).

Recall that the MM has no intrinsic need or desire to hold any inventory, so she will only buy (sell) in anticipation of a subsequent sale (purchase). Grossman & Miller (1988) provide a model that captures this problem and describes how MMs obtain a liquidity premium from liquidity traders that exactly compensates MMs for the price risk of holding an inventory of the asset until they can unload it later to another liquidity trader.

Let us consider a simplified version of their model, with a finite number, n, of identical MMs for some given asset and three dates $t \in \{1, 2, 3\}$. To simplify the situation, there is no uncertainty about the arrival of matching orders: if at date $t = 1$ a liquidity trader, denoted by LT1, comes to the market to sell i units of the asset, there will be (for sure) another liquidity trader (LT2) who will arrive at the market to purchase i units (or more generally, to trade $-i$ units, so that LT1's trade (of i units) could be negative or positive (LT1 could be buying or selling). However, LT2 does not arrive to the market until $t = 2$. Let all agents start with an initial cash amount equal to W_0, MMs hold no assets, LT1 holds i units and LT2 $-i$ units.

There are no trading costs or direct costs for holding inventory. The focus is on price changes: the asset will have a cash value at $t = 3$ of $S_3 = \mu + \epsilon_2 + \epsilon_3$, where μ is constant, ϵ_2 and ϵ_3 are independent, normally distributed random variables with mean zero and variance σ^2. These will be publicly announced between dates $t-1$ and t, that is ϵ_3 is announced between $t = 2$ and $t = 3$, and ϵ_2 is announced between $t = 1$ and $t = 2$. Hence, the realised cash value of the asset can increase or decrease (ignore the fact that there are realisations of ϵ_2 and ϵ_3 that could make the asset value negative – the model serves to illustrate a point). Because the shocks to the value of the asset are on average zero a risk-neutral trader has no cost at all from holding the asset. The model becomes interesting if we assume that all traders, MMs and liquidity traders, are risk-averse. To be more specific, suppose they have the following expected utility for the future random cash value of the asset (X_3): $\mathbb{E}[U(X_3)]$ where $U(X) = -\exp(-\gamma X)$, and where $\gamma > 0$ is a parameter capturing the utility penalty for taking risks (the risk aversion parameter).

Solving the model backwards we obtain a description of trading behaviour and prices. At $t = 3$ the cash value of the asset is realised, $S_3 = \mu + \epsilon_2 + \epsilon_3$. At $t = 2$, the n MMs and LT1 come into the period with asset holdings q_1^{MM} and q_1^{LT1}

respectively. LT2 comes in with $-i$ and they all exit with asset holdings q_2^j, where $j \in \{MM, LT1, LT2\}$. Note that if, for example, $q_t^j = 2$ this denotes that agent j is holding 2 units when exiting date t, so that the agent will be long (that is, has an inventory of) two units. Given the problem as described so far, at $t = 2$ agent j chooses q_2^j to maximise his expected utility knowing the realisation of ϵ_2 that was made public before $t = 2$:

$$\max_{q_2^j} \mathbb{E}\left[U\left(X_3^j\right) \mid \epsilon_2\right]$$

subject to

$$X_3^j = X_2^j + q_2^j\, S_3, \qquad X_2^j + q_2^j\, S_2 = X_1^j + q_1^j\, S_2\,.$$

These two constraints capture:

(i) the fact that the cash value of the agent's assets at $t = 3$, X_3, is equal to the agent's cash holdings at $t = 3$, which is equal to X_2 plus the cash value of the agent's asset inventory q_2^j, and

(ii) the fact that the cash value of the agent's assets when exiting date $t = 2$ (X_2, and the inventory q_2^j) was equal to the cash value of the agent's assets when entering date $t = 2$ (X_1, and the inventory q_1^j).

Given the normality assumption and the properties of the expected utility function it is straightforward to show that

$$\mathbb{E}\left[U\left(X_3^j\right) \mid \epsilon_2\right] = -\exp\left\{-\gamma\left(X_2^j + q_2^j\,\mathbb{E}[S_3 \mid \epsilon_2]\right) + \frac{1}{2}\gamma^2\left(q_2^j\right)^2\sigma^2\right\}.$$

Thus, the problem is concave and the solution is characterised by

$$q_2^{j,*} = \frac{\mathbb{E}[S_3 \mid \epsilon_2] - S_2}{\gamma\sigma^2},$$

for all agents: the n MMs, LT1, and LT2.

As at date $t = 2$ demand and supply for the asset have to be equal to each other, we can solve for the equilibrium price S_2:

$$n\, q_1^{MM} + q_1^{LT1} + q_1^{LT2} = n\, q_2^{MM} + q_2^{LT1} + q_2^{LT2}\,, \tag{2.1}$$

where we use the convention that q_1^{LT2}, the assets LT2 came into period 2 with, is equal to his desired trade, $-i$. As we have established above, all q_2^j are equal, so that the right-hand side of the above equation is equal to

$$n\, q_2^{MM} + q_2^{LT1} + q_2^{LT2} = (n+2)\,\frac{\mathbb{E}[S_3 \mid \epsilon_2] - S_2}{\gamma\sigma^2}\,. \tag{2.2}$$

We also know that at date 1 the total quantity of the asset available was equal to the quantity of assets LT1 brought to the market, so that the LHS of (2.1) is

$$n\, q_1^{MM} + q_1^{LT1} + q_1^{LT2} = i + q_1^{LT2} = i - i = 0\,.$$

Hence, substituting into (2.2) and solving, we obtain that in equilibrium, at date

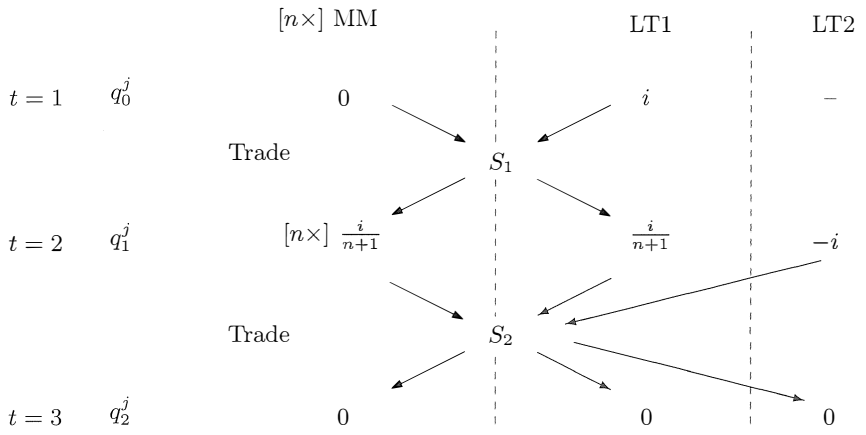

Figure 2.1 Trading and price setting in the Grossman–Miller model.

$t = 2$, $S_2 = \mathbb{E}[S_3] = \mu + \epsilon_2 + \mathbb{E}[\epsilon_3] = \mu + \epsilon_2$, and therefore, $q_2^j = 0$. This makes sense, as at $t = 2$ there are no asset imbalances, the price of the asset reflects its 'fundamental value' (efficient price) and no one will want to hold a non-zero amount of the risky asset. This analysis is captured in the bottom half of Figure 2.1, where we see the asset holdings of the three types of participants as they enter $t = 2$, q_1^j, $j \in \{MM, LT1, LT2\}$, and how after trading at a price equal to S_2 they end up with holdings, q_2^j, equal to zero.

Consider now what happens at date $t = 1$. Participating agents (the n MMs and LT1 – recall that LT2 will not appear until $t = 2$) anticipate that whatever they do, the future market price will be efficient and they will end up exiting date $t = 2$ with no inventories, so that $X_3 = X_2$. Thus, their portfolio decision is given by

$$\max_q \mathbb{E}\left[U\left(X_2^j\right)\right],$$

subject to

$$X_2^j = X_1^j + q_1^j\, S_1, \qquad X_1^j + q_1^j\, S_1 = X_0^j + q_0^j\, S_1\,.$$

Repeating the analysis of $t = 2$ at $t = 1$, we obtain that the optimal portfolio solution is

$$q_1^{j,*} = \frac{\mathbb{E}[S_2] - S_1}{\gamma\sigma^2},$$

for all agents, j, that are present: the n MMs, and LT1. Also, at date $t = 1$ demand and supply for the asset have to be equal to each other, so that

$$n\, q_0^{MM} + q_0^{LT1} = n\, q_1^{MM} + q_1^{LT1},$$

where $q_0^{LT1} = i$ (recall that if $i > 0$, LT1 is holding i shares he wants to sell),

and $q_0^{MM} = 0$. This gives us the following equation:

$$i = (n+1)\frac{\mu - S_1}{\gamma\sigma^2} \iff S_1 = \mu - \gamma\sigma^2\frac{i}{n+1}\,.$$

The top half of Figure 2.1 reflects how the MMs and LT1 enter the market with asset holdings q_0^j and after trading at S_1 they exit date one and enter date $t = 2$ with q_1^j.

With this expression we can interpret how the market reaches a solution for LT1's liquidity needs: LT1, a trader who wants to sell a total of $i > 0$ units at $t = 1$, finds that there is no one currently in the market with a balancing liquidity need. There are traders in the market, but they will not accept trading at the efficient price of μ because if they do, they will be taking on risky shares (they are exposed to the price risk from the realisation of ϵ_2) without compensation. But, if they receive adequate compensation (which we call a liquidity discount, as for $i > 0$, $S_1 < \mathbb{E}[S_2] = \mu$), the n MMs will accept the LT1's shares. However, LT1 is price-sensitive, so if he has to accept a discount on the shares, he will not sell all the i shares at once. In equilibrium, both the n MMs and LT1 end up holding $q_1^{j,*} = \frac{i}{n+1}$ units of the asset each, that is LT1 sells $\frac{n}{n+1}i$ units and holds on to $\frac{i}{n+1}$ units to be sold later. Trading occurs at a price below the efficient price, $S_1 = \mu - \gamma\sigma^2\frac{i}{n+1}$. The difference between the trading price and the efficient price, namely $|S_1 - \mu| = \left|\gamma\sigma^2\frac{i}{n+1}\right|$, represents the (liquidity) discount the MMs receive in order to hold LT1's shares. This size of the discount is influenced by the variables in the model: the size of the liquidity demand ($|i|$), the amount of competition amongst MMs (captured by n), the market's risk aversion (γ), and the risk/volatility of the underlying asset (σ^2). These variables all affect the discount in an intuitive way: the size of the liquidity shock, risk aversion, and volatility all increase the discount, while competition reduces it. This occurs when LT1 wants to sell, i.e. $i > 0$. If LT1 wanted to buy, $i < 0$, then the solution would be the same except that instead of a discount, the MMs would receive a premium equal to $|S_1 - \mu|$ per share when selling to LT1.

From this analysis we can also see that as competition (n) increases, the liquidity premium goes to zero, the price converges to the efficient level, $S_1 = \mu$, and LT1's optimal initial net trade, $q_1^{LT1,*} - q_0^{LT1}$, converges to his liquidity need (i).

2.1.2 Trading Costs

We have seen how the Grossman & Miller (1988) framework helps to understand how the cost of holding assets (in this case, via the uncertainty it generates to the risk-averse MMs) affects liquidity via the cost of trading ($|S_1 - \mu|$) and the demand for immediacy (as at $t = 1$ LT1 only executes $\frac{n}{n+1}i$ rather than her desired i). Also, competition between MMs is crucial in determining these trading costs. But what drives n? A natural answer is that n is driven by the trading costs borne by the MMs. In this case, we must distinguish between participation

costs, which are needed to be present in the market and do not depend on trading activity, and trading costs that do depend on trading activity, such as trading fees (which we ignored in the previous analysis).

Grossman & Miller (1988) link competition, n, with participation costs. They do this by introducing an earlier stage to the model in which potential MMs decide whether they want to actively participate in the market and provide liquidity or prefer to do something else. The decision is determined as a function of a participation cost parameter c which proxies for the time and investments needed to keep a constant, active and competitive presence in the market, as well as the opportunity cost the MM gives up by being in the market and not doing something else. The conclusion, which can be obtained without going into the details of the analysis, is that the level of competition decreases monotonically with supplier's participation costs. Thus, participation costs, proxied by the cost parameter c, increase the size of the liquidity premium (via its effect on competition, n).

The parameter c captures the fixed costs of participating in the market, but we could also consider introducing into the model a cost of trading that depends on the level of activity in the market. In particular, we introduce trading costs that depend on the quantity traded, like actual exchange trading fees. Exchange trading fees are usually proportional to dollar-volume but here, for simplicity, we use fees proportional to number of shares traded parameterised by η. Given that fees are known, these fees act like a participation cost for liquidity traders.

The first effect of having $\eta > 0$ is that liquidity traders with a desired trade ($|i|$) that is small relative to trading fees, will find trading too expensive and refrain from trading (we invite the interested reader to compute the minimum desired trade size $\hat{i}$ as a function of η). For sufficiently large desired trades (so that trading is preferred to not trading by all participants) the model gives us the following solution. Suppose every trader pays η per share regardless of whether they are buying or selling the asset. To simplify, assume that any remaining inventories after $t = 2$ are liquidated at $t = 3$. Also, assume LT1 wants to sell $|i|$ units ($i > 0$), and LT2 wants to buy the same amount (the reverse case looks the same but the trading fees enter the problem with the opposite sign).

At $t = 2$, since the MMs and LT1 enter the period with a positive inventory (and will be wanting to sell now or at $t = 3$) their optimal final period holdings are

$$q_2^j = \frac{\mathbb{E}[S_3 - \eta \,|\, \epsilon_2] - (S_2 - \eta)}{\gamma \sigma^2}, \quad j \in \{MM, LT1\},$$

while the demand for shares by LT2 is

$$q_2^{LT2} = \frac{\mathbb{E}[S_3 + \eta \,|\, \epsilon_2] - (S_2 + \eta)}{\gamma \sigma^2}.$$

As everyone anticipates that their trading positions need to be liquidated anyway, the trading fees do not affect the price at $t = 2$, and we obtain $S_2 = \mathbb{E}[S_3 \,|\, \epsilon_2] = \mu + \epsilon_2$ (as before when there were no fees, $\eta = 0$).

At $t = 1$, LT1 has a similar position to that at $t = 2$, as any quantities he doesn't sell now he will have to sell later, so that η disappears from the solution and his supply will be given by:

$$q_1^{LT1} = \frac{\mathbb{E}[S_2 - \eta] - (S_1 - \eta)}{\gamma\sigma^2}.$$

On the other hand, MMs anticipate that whatever they buy, they will have to sell later, which changes their asset demand functions to

$$q_1^{MM} = \frac{\mathbb{E}[S_2 - \eta] - (S_1 + \eta)}{\gamma\sigma^2}.$$

The resulting market equilibrium condition is now

$$i = n\, q_1^{MM} + q_1^{LT1} = \frac{\mu - S_1}{\gamma\sigma} + n\,\frac{\mu - S_1 - 2\eta}{\gamma\sigma}.$$

This gives us the following equation:

$$i = (n+1)\,\frac{\mu - S_1}{} - \frac{2\,n\,\eta}{} \iff S_1 = \mu - \gamma \quad \frac{i}{n+1} - 2\left(\frac{n}{n+1}\right)\eta,$$

and recall that for LT1, $i > 0$.

Thus, we conclude that the presence of trading fees introduces an extra liquidity discount to the initial price S_1. What the model tells us is that almost all the trading fees are paid by the liquidity trader initiating the transaction: he pays his own trading fee, η per share, plus a substantial fraction $(n/(n+1))$ of the two transaction fees paid by the MMs (2η) though indirectly, via a lower sale price, a lower S_1 It also affects the immediacy he obtains from the market, as his holdings at the end of $t = 1$ are no longer $q_1^{LT1,*} = -i/(n+1)$ but

$$q_1^{LT1,*} = \frac{i}{n+1} + 2\left(\frac{n}{n+1}\right)\frac{\eta}{}.$$

If we look at competition, we can see that participation costs and fees have very different effects. Participation costs enter directly through c while trading fees enter through expected future profits, which will be lower as MMs must bear a fraction of the trading fees. In particular, for each trade, the MM pays 2η, but recovers $2n/(n+1)\eta$ through the liquidity discount. Therefore, an increase in trading fees has a smaller effect on liquidity via competition but a greater direct effect on immediacy and the liquidity discount.

2.1.3 Measuring Liquidity

We have seen how in the Grossman & Miller (1988) model, trading costs, whether setup costs or trading fees, are mostly paid by liquidity traders, whether explicitly (as their own trading fees) or implicitly in the price (greater liquidity discount when selling and larger premium when buying). We now consider how these divergences from 'efficient' prices may be observed in electronic exchanges.

The Grossman & Miller (1988) model avoids looking into the details of the

trading mechanism by solving for equilibrium prices and demands in a 'Walrasian auctioneer'-type context where all trading takes place at once, and at a single price.[2] In electronic asset markets, decisions are not taken all at the same point in time, but the equilibrium analysis can be easily reinterpreted in the context of an electronic market. For example, suppose liquidity traders are very eager to trade and do so by sending MOs into the exchange. When the liquidity trader's orders hit the market, they meet the LOs that were posted by the patient MMs and are resting in the limit order book (LOB).

Then, the Grossman–Miller model would correspond to the following sequence of events: as LT1's MOs enter the market, they execute against LOs in the LOB which adjusts to the incoming MO. As the execution price changes, so does LT1's strategy and eventually, after selling $i\,\frac{n}{n+1}$ shares, the price has moved too far and LT1 stops trading. Overall, LT1's market order executes at the average price of S_1, either because it was sent as a large order that walked the LOB (or LOBs, if the order is routed to multiple markets), or because it was split up into several small orders that triggered a gradual move of the bid side in the LOB away from the initial starting point. Then, the discount received by LT1 is the difference between the average price received, S_1, and the initial midprice when the first MO hit the market (which is the usual proxy for the efficient price, $\mathbb{E}[S_2]$).

We can rewrite S_1 as a linear function of the quantity traded, q^{LT1}:

$$S_1 = \mu + \lambda\, q^{LT1},$$

so that in the Grossman–Miller model we would have

$$\lambda = -\frac{1}{n}\,\gamma\sigma^2, \qquad \text{and} \qquad q^{LT1} = i\,\frac{n}{n+1}\,.$$

The λ parameter captures the market's price reaction to LT1's total order, its price impact. The notion of price impact is very important both for trading and for theoretical work, and the use of a linear structure such as the one described by the parameter λ is very common. In particular, λ is used to describe the liquidity of the market for this asset – a more liquid market will have a lower λ, either because of greater competition (n), lower risk tolerance (γ), or lower volatility (σ^2), and this results in a lower liquidity discount/premium for liquidity traders.

There is a second popular way to measure liquidity based on price changes, and it is quite easy to see how this model works. The measure is based on the autocovariance of the asset's return, though for the Grossman–Miller model it is easier to describe it when looking at the autocovariance in asset price changes rather than returns. To see how this measure is constructed, let us introduce an additional date $t = 0$ prior to LT1's order submission ($t = 1$), and a random public news event, ϵ_1, that affects the asset's final liquidation price, $S_3 = \mu +$

[2] The notion of a Walrasian auctioneer comes from the work of Léon Walras who describes the prices that arise under perfect competition as the result of a simultaneous auction in which supply is equated to demand. The Walrasian auctioneer is the abstract manager of this auction.

$\epsilon_1 + \epsilon_2 + \epsilon_3$. The public news is announced between $t = 0$ and $t = 1$. Define the following constants:

$$\mu_0 = \mathbb{E}[S_3]\,, \qquad \mu_1 = \mathbb{E}[S_3 \,|\, \epsilon_1]\,, \qquad \mu_2 = \mathbb{E}[S_3 \,|\, \epsilon_1, \epsilon_2] \qquad \mu_3 = S_3\,,$$

and let ϵ_t, $t = 1, 2, 3$ be normal, i.i.d. random variables with mean zero and variance σ^2. The discrete process μ_t is a martingale, and we refer to it as the efficient market price.

According to the model above, at $t = 0$ there are no liquidity traders and no trade so that $S_0 = \mathbb{E}[S_3] = \mu_0$ will be the equilibrium price. The model shows that the subsequent equilibrium prices at $t = 1$ and $t = 2$ are:

$$S_1 = \mu_1 + \lambda\, q^{LT1}, \qquad \text{and} \qquad S_2 = \mu_2\,.$$

To construct the autocovariance of price changes, let $\Delta_1 = S_1 - S_0$ and $\Delta_2 = S_2 - S_1$, and the autocovariance of price changes be given by the following expression:

$$\begin{aligned}\mathrm{Cov}\,[\Delta_1, \Delta_2] &= \mathrm{Cov}\,\left[\mu_1 + \lambda\, q^{LT1} - \mu_0, \mu_2 - \mu_1 - \lambda\, q^{LT1}\right] \\ &= \mathrm{Cov}\,\left[\epsilon_1 + \lambda\, q^{LT1}, \epsilon_2 - \lambda\, q^{LT1}\right] = -\lambda^2\,\mathrm{Var}\left[q^{LT1}\right] < 0\,.\end{aligned}$$

In this simple (essentially static) model, where all the action takes place at $t = 1$, the autocovariance of price changes captures market liquidity just like price impact does. An interesting effect that we see here is that as liquidity increases and λ goes to zero, so the autocovariance of price changes, and the price process converges to the underlying ('efficient price') martingale process μ_t.

The two measures, price impact and the autocovariance of price changes (or returns), become distinct in richer dynamic settings, and capture different dimensions of the market's reaction to incoming MOs. For example, in the continuous-time models of later chapters, the average growth of the efficient price is affected by the rate at which MOs arrive to the market and this effect decays at an exponential rate. This permanent effect of the efficient price of the asset affects all market participants and is different from the temporary effect that each trader sees in their execution prices, which is captured by the parameter (λ) and does not affect the dynamics of the efficient price.

2.1.4 Market Making using Limit Orders

In the transition from the Walrasian auctioneer in the Grossman–Miller model to the measurement of price impact, we have proposed that MMs participate through the posting of LOs. We now consider why an MM would behave in this way and the simplest solution to how she does it.

The usual first reference for this is the model of Ho & Stoll (1981), but working with the original model requires familiarity with the techniques for solving stochastic dynamic programming problems which we see in Part II. Instead, we set up a static version of the model that captures some of the basic elements of the MM's problem. As in the Grossman & Miller (1988) model, the MM is

a professional trader who profits from intermediating between different liquidity traders. In this case, we consider a small risk-neutral trader with costless inventory management and infinite patience. She does not require compensation for her services, but makes a profit from optimally choosing how to provide liquidity in an uncertain environment populated by other MMs who do not react to our MM's decisions.

Uncertainty in this context comes from the timing and size of large incoming MOs, and there are no information problems: all information is public so that everyone agrees what the current value of the asset is, which we denote by S_t and refer to as the midprice. Our trader is one of many MMs. We take other MMs' behaviour as given, and this behaviour is represented by a fixed LOB, unaffected by our MM's decisions. Our MM makes money by adding her LOs to the book and clearing the resulting inventory at later dates. Because our MM has no inventory costs, incurs no trading costs, is risk-neutral and infinitely patient, we can assume that she liquidates her inventory at the midprice at no cost.

Then, the MM's problem is to choose where on the LOB to place her LOs so as to maximise her profit per trade, optimally balancing the increase in the price per trade received as she increases the distance of the LO from the midprice, with the frequency with which she will trade, which decreases with that distance from the midprice. Formally, the MM's problem is to choose the distance from the midprice, the depths $\delta^{\pm}$. Then, she will post her sell LO at $S_t + \delta^+$ and her buy LO at $S_t - \delta^-$. The uncertainty from MOs comes from the probability that an MO arrives ($p_{\pm}$) and the probability that once it arrives it walks the book up to where the MM's LOs are resting ($\delta^{\pm}$ away from the midprice), which is described by the cdf $P_{\pm}$. Thus, the probability that the buy LO will be filled is $p_- P_-(\delta^-)$. If we assume that the distribution of other LOs in the LOB is described by an exponential distribution with parameter κ^-, we have $p_- P_-(\delta^-) = p_- e^{-\kappa^- \delta^-}$. Similarly, the probability that the sell LO is filled is $p_+ e^{-\kappa^+ \delta^+}$. Clearly, as the MM posts her LOs deeper in the LOB, the probability that her order (once an MO arrives) decreases, though her profit per trade ($\delta^{\pm}$) increases.

The left panel of Figure 2.2 illustrates a hypothetical LOB around a midprice of S_t and two possible limit orders: a sell LO on the ask side at $S_t + \delta^+$, and a buy LO on the bid side at $S_t - \delta^-$. The right panel describes the corresponding probability distribution, P^+ (P^-), of execution of the order posted at a distance δ^+ (δ^-) from the midprice, conditional on the arrival of a buy (sell) MO.

Using Π to denote the MM's profit per trade, the MM's optimisation problem is given by the following expression:

$$\max_{\delta^+,\,\delta^-} \mathbb{E}\left[\,\Pi(\delta^+,\,\delta^-)\,\right] = \max_{\delta^+,\,\delta^-} \left\{ p^+ \, e^{-\kappa^+ \delta^+} \, \delta^+ + p^- \, e^{-\kappa^- \delta^-} \, \delta^- \right\}. \tag{2.3}$$

It is straightforward to see that the solution is to post LOs at the following depths:

$$\delta^{\pm,*} = \frac{1}{\kappa^{\pm}}.$$

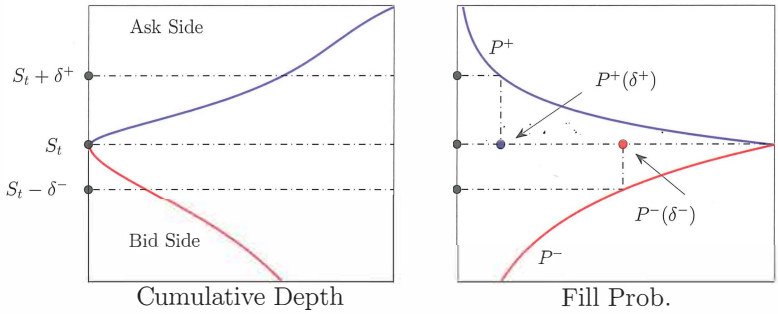

Figure 2.2 The LOB and the probability of execution.

Given our parametric choice of $P_{\pm}$, the optimal depth is equal to the mean depth in the LOB.

This model captures in a simple way the trade-off between the probability of execution and margin per trade. But, it is very unrealistic in several dimensions: the functional form of all the stochastic components of the model ($P_{\pm}$, and $p^{\pm}$) is very special, constant and exogenous, the MM's decision and that of other traders (as captured by $P(\delta)$) are independent, the MM's objective function is static and very simple. However to address these other issues we need more sophisticated methods and models, so after developing those methods in the following chapters we will revisit some of them. For instance, in Chapter 10 we see how MMs decide how to post limit orders in a fully-fledged dynamic inventory model and how she adjusts her posts if trading with better informed counterparties – a topic that we discuss next.

2.2 Trading on an Informational Advantage

So far we have side-stepped one of the main issues in trading: informational differences. Many trades originate not because someone needs cash and sells an asset, or has extra cash and wants to invest, but because one party has (or believes she has) better information about what the price is going to do than is reflected in current prices. So, having seen the basic market making models in the context of public information we turn to the next fundamental issue: how to exploit an informational advantage while taking into account one's price impact. The primary reference in this case is Kyle (1985).

Kyle (1985) looks at the decision problem of a trader who has a strong informational advantage (the case of several competing informed traders is studied in Kyle (1989)) in a context where the price is 'efficient'. The model in Kyle (1985) tells us how the informed trader optimally adjusts his trading strategy to take into account the market reaction, and in particular, the price impact that his trades generate in equilibrium.

To get into the details of the model we first need to define what we mean by 'a strong informational advantage' and price efficiency in this context. To keep things simple we only consider the investor's static decision problem. The same basic idea extends to a dynamic setting. The formal static model is as follows: there is a market for an asset that opens at one point in time. The asset is traded at price S, and after trading, the asset has a cash value equal to v. The future cash value of the asset, v, is uncertain. In particular, v is assumed to be normally distributed with mean μ and variance σ^2. In the market, there are three types of traders: an informed trader, an anonymous mass of price-insensitive liquidity traders (traders who need to execute trades whatever the cost), and a large number of MMs that observe and compete for the order flow – that is, the MMs observe and compete for the flow of incoming buy and sell orders from the informed and the liquidity traders.

In contrast to the Grossman & Miller (1988) setting, MMs are risk-neutral, so they do not need a liquidity premium to compensate for the price risk from holding inventory. Therefore, any liquidity premium that arises will come from the need to compensate MMs for their informational disadvantage – and which will be borne by the price-insensitive liquidity traders. These liquidity traders will have, in aggregate, a net demand represented by the random quantity u, such that if $u > 0$, on aggregate liquidity traders want to buy u units, while if $u < 0$, these traders want to sell $|u|$ units of the asset. Assume that u is normally distributed with mean zero, variance σ_u^2, and is independent of v. In principle, as liquidity traders are not sensitive to the price (u does not depend on S) MMs could charge very large liquidity premia, but competition for order flow between MMs drives the liquidity premium to zero, so that (when there are only MMs and liquidity traders) $S = \mathbb{E}[v]$.

Now consider the possibility that a new trader enters the market, and that this trader (the "insider") knows the exact value of v. The insider is the only one who knows v and chooses how much to trade. Let $x(v)$ denote the number of shares traded by the insider. MMs, on the other hand, know that there is an informed trader in the market, but do not know who this trader is.

To make the analysis formal, the model is structured as follows: (i) the insider observes v, (ii) on observing v the insider chooses $x(v)$, (iii) u is realised, (iv) the MMs observe the net order flow, $x(v) + u$, (v) based on the net order flow MMs compete to set the asset price, S.

To solve the model we use the solution concept of (Bayesian) Nash equilibrium; without going into all the details, this means that all agents optimise given the decisions of all other players, according to their beliefs (which are updated according to Bayes' rule whenever possible). Thus, we require that in equilibrium the insider chooses $x(v)$ to maximise his expected profit, taking into account the dynamics of the game (i.e. that his order will be mixed in with those of the liquidity traders), and anticipating that MMs will set their prices on the basis of what they learn from observing the order flow and what they know about the informed trader's decision problem. Also, we require that MMs choose

their prices taking into account the strategy of the insider (in particular, they anticipate the functional form of $x(v)$) and the properties of the uninformed order flow that comes from liquidity traders. In particular, MMs set the market price as a function of net order flow, $S(x+u)$. This is important, as the model naturally tells us that prices are affected by the order flow, so that trading automatically generates a price impact – the average price per unit traded, S, moves with the net order flow, $x+u$. We need to look at the equilibrium of the model to see what that price impact function looks like. Nevertheless, in equilibrium, the insider will anticipate the functional form of $S(x+u)$, that is, she will incorporate price impact when choosing $x(v)$.[3] The equilibrium is a fixed point in the optimisation of x given the functional form of S, and of S given the functional form of x.

Consider what the insider should do. The most natural response is: sell if $v < \mathbb{E}[v] = \mu$ and buy if $v > \mu$, and whether selling or buying, do so as much as possible to leverage his informational advantage. This seems natural, but we must take into account that MMs will adjust their prices to the order flow they observe. Hence, even if $v < \mu$, the insider cannot expect $S = \mu$. In the extreme case where there are no liquidity traders everyone knows that any trade comes from the insider and so the MMs, anticipating the demand as a function of the realisation of v, behave optimally and set prices that incorporate all information on v in $x(v)$. Fortunately for the insider, there are liquidity traders that add noise into order flow and allow the insider to camouflage his trade to gain positive expected profits.

So, how do MMs set their prices? The first thing to note is that as MMs compete for order flow, any profits they could extract are competed away. Thus, whatever the price strategy, it will lead to zero expected profits for our (risk-neutral) MMs – though never negative profits as they can always choose not to trade. The zero (expected) profit condition forces prices to have a very specific property: $S = \mathbb{E}[v \,|\, \mathcal{F}]$, where $\mathcal{F}$ represents all information available to MMs. This property is known as **semi-strong efficiency**: prices reflect all publicly available information (which in our case is order flow which is all the information MMs have).[4] This is why we can readily identify a fundamental property of the MMs' equilibrium strategy:

$$S(x+u) = \mathbb{E}[v \,|\, x+u]\,.$$

To solve the model we need to find an $x(v)$ that is optimal, i.e. it maximises the insider's expected trading profits, conditional on this pricing rule. Because of the normality of v and u, we hypothesise that $S(x+u)$ is linear in net order flow. In particular, let

$$S(x+u) = \mu + \lambda\,(x+u),$$

[3] Formally, liquidity traders are substituted by a "nature" player that executes the random demand u.

[4] The notion of price efficiency was introduced by the recent Nobel Laureate, Eugene Fama, see Fama (1970).

where λ is an unknown parameter representing the linear sensitivity of the market price to order flow.

Taking this particular functional form as given, consider the insider's problem:

$$\max_x \mathbb{E}\left[x\left(v - S(x+u)\right)\right].$$

Substituting for $S(x+u) = \mu + \lambda(x+u)$ and taking expectations with respect to u, we obtain that the objective function is concave and the first-order condition yields

$$x^*(v) = \beta(v - \mu),$$

where $\beta = (2\lambda)^{-1}$.

Because we have hypothesised the functional form of the price function, we must now confirm that the functional form is consistent with the optimal $x(v)$ and at the same time we can characterise λ. We know that $S = \mathbb{E}[v \mid x+u]$. From the optimal x, we know that

$$x + u = \beta(v - \mu) + u = \beta\mu + \beta v - u.$$

As v and u are independent and normal, $x+u$ is normal with mean $\mu(1+\beta)$ and variance $\beta^2\sigma^2 + \sigma_u^2$. We can now compute the joint distribution of v and $x+u$, and from it we can derive $S = \mathbb{E}[v \mid x+u]$, which (using the projection theorem for normal random variables and simplifying) is given by

$$S = \mu + 2(x+u)\frac{\sigma_u}{\sigma},$$

so that the linear sensitivity parameter is $\lambda = 2\sigma_u/\sigma$. This confirms that the hypothesised equilibrium is indeed an equilibrium (for a formal proof, see Kyle (1985)).

Even within the simple, static version of the Kyle model we can clearly see the issues that arise when facing informed trading (also referred to as "toxic order flow"). While in the previous models MMs just needed a liquidity premium (discount) to cover the expected cost from future price uncertainty, the presence of informed traders implies that MMs will be adversely selected, buying when informed traders know it would be better to sell and selling when it would be better to buy. This adverse selection requires a higher premium borne by other (more impatient liquidity) traders. In this model, the additional premium takes the form of price adjustment to order flow (price impact) as described by Kyle's lambda (the λ parameter we have just derived). This premium accounts not for the risk that future price movements will be random, as described in Section 2.1.1, but for the adverse selection faced by MMs, as prices will on average move *against* the MMs' position because they trade with better informed traders in the market. The sign of λ will be the same as in Grossman & Miller (1988): prices move with the order flow, increasing as buy MOs hit the market and falling as traders sell aggressively.

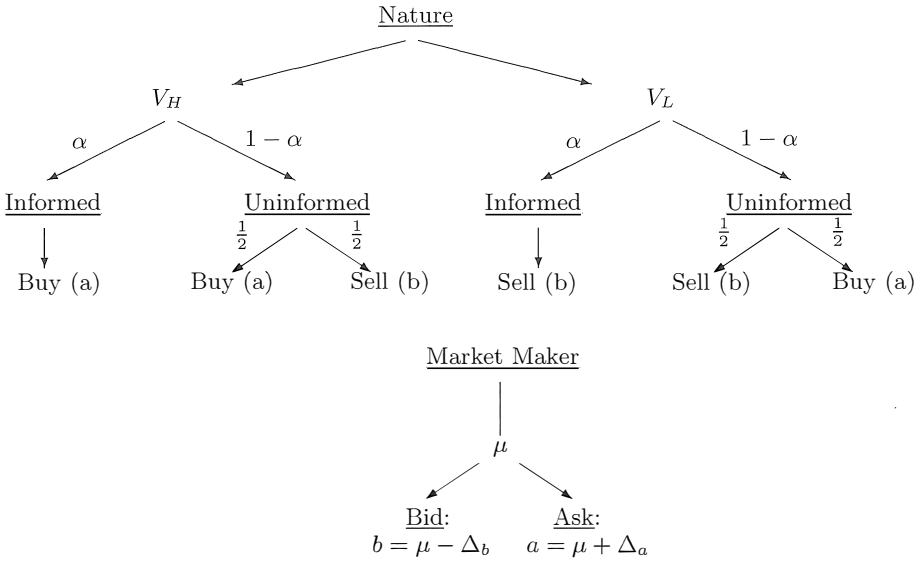

Figure 2.3 The Glosten-Milgrom model.

2.3 Market Making with an Informational Disadvantage

The Kyle model focuses on the informed trader's problem, while using competition to characterise the MM's decisions. As we are very interested in the MM's problem, we now turn to Glosten & Milgrom (1985) for a model that puts the MM at the centre of the problem of trading with counterparties who have superior information.

Again, we look at a simplified and (essentially) static version of the model that allows us to capture the nature of the MM's decision problem. The situation is as before: there are liquidity traders, informed traders, and a competitive group of MMs. The MM is risk-neutral and has no explicit costs from carrying inventory.

Our simple model (described in Figure 2.3) has a future cash value of the asset equal to v which we limit to two possible values: $V_H > V_L$, that is a High, and a Low value. The unconditional probability of $v = V_H$ is p. All orders are of one unit, MMs post an LO to sell one unit at price a, and a buy LO for one unit at price b. We start by assuming that liquidity traders are price insensitive and want to buy with probability 1/2 and want to sell with probability 1/2. There are many informed traders, all of whom know v but are limited to trade a single unit, which simplifies their decision: when $v = V_H$ they buy one unit if $a < V_H$, and do nothing otherwise, while when $v = V_L$ they sell one unit if $b > V_L$ and do nothing otherwise. The total population of liquidity and informed traders is normalised to one, and of these, a proportion α are informed and a proportion $(1 - \alpha)$ are uninformed liquidity traders.

Figure 2.3 captures the probabilistic structure of the model: Nature randomly

determines whether the underlying state is V_H or V_L. Independently of the state, a trader is picked at random from the population, so that with probability α she is informed, and with probability $1-\alpha$ she is uninformed. An informed trader will always buy at the ask price (a) when the asset's value is V_H and sell at the bid (b) when the asset's value is V_L, while an uninformed trader will buy or sell with equal probability, independent of the true (unknown) value of the asset.

The MM's problem is to choose a and b in this setting. Because liquidity traders are price-insensitive, the optimal solution is trivial: set $a = V_H$ and $b = V_L$, but since MMs compete for business, prices will be set to their (semi-strong) efficient levels – again, this happens because MMs use only public information, *which includes order flow*. Were the MMs to have private information in addition to the order flow, in this setting competition for order flow would incorporate some of that information into prices.

Competition between MMs drives their expected profits to zero. Hence, a and b are determined by the no-profit condition. Rather than solve for a and b directly, define the ask- and bid-halfspreads, Δ_a and Δ_b respectively. The sum of the two, $\Delta_a + \Delta_b$, represents the (quoted) spread. Let the expected value of the asset $\mu = \mathbb{E}[v \mid \mathcal{F}]$ where $\mathcal{F}$ represents all public information prior to trading. Then, as described at the bottom of Figure 2.3, MMs will choose $a = \mu + \Delta_a$ and $b = \mu - \Delta_b$ optimally. To determine the effect of choosing a and b on the expected profit and loss, consider what happens when a buy order comes in:

- if it comes from an uninformed liquidity trader she makes an expected profit of $a - \mu = \Delta_a$,
- if it comes from an informed trader she makes an expected loss of $a - V_H = \Delta_a - (V_H - \mu)$.

From the point of view of the MM, the probability that a liquidity trader wants to buy is $1/2$, while the probability that an informed trader wants to buy is p (as all informed traders will buy if the state is $v = V_H$ which occurs with probability p). As there are $1-\alpha$ liquidity traders and α informed ones, the expected profit from posting a price $a = \mu + \Delta_a$ is

$$\frac{(1-\alpha)/2}{\alpha p + (1-\alpha)/2}\Delta_a + \frac{p\alpha}{\alpha p + (1-\alpha)/2}\left(\Delta_a - (V_H - \mu)\right) .$$

Setting this expected profit to zero we obtain

$$\Delta_a = \frac{\alpha p}{\alpha p + (1-\alpha)/2}\left(V_H - \mu\right) = \frac{1}{1 + \frac{1-\alpha}{\alpha}\frac{1/2}{p}}\left(V_H - \mu\right) ,$$

and following similar reasoning,

$$\Delta_b = \frac{1}{1 + \frac{1-\alpha}{\alpha}\frac{1/2}{1-p}}\left(\mu - V_L\right) .$$

To interpret these equations let us label the variables. If we think of asymmetric information as 'toxicity' then we can think of α as the prevalence of toxicity, $1-p$

and $V_H - \mu$ as the magnitude of buy-toxicity and $1-p$ and $\mu - V_L$ that of sell-toxicity. Then, the equations above describe how MMs adjust the ask-half-spread and the bid-half-spread, and increase it with the prevalence and magnitude of buy- and sell-toxicity.

In later chapters we show how trading algorithms are built to either take advantage of informational advantages or to adjust the depth at which LOs are posted so as to recover losses from trading agents to more informed traders. For example, in Section 7.3 we develop trading algorithms that use the information provided in the order flow to adjust acquisition or liquidation rates when the agent seeks to enter or exit a large position. We also show how the strategy of the MM depends on whether she knows detailed high-frequency information about short-term deviations in the drift of the asset she is trading, see for example Section 10.4.2.

2.3.1 Price Dynamics

This simple model can be extended in two different and complementary ways: by incorporating a time dimension and by making liquidity traders price-sensitive. The former is straightforward. In order to avoid having to keep track of the interest rate, set it equal to zero. Then index all variables by time t and set the time of the determination of the cash value of the asset to T. Moreover, ensure that probabilities and expectations are adjusted to incorporate the accumulation of public information from trade, as captured by the filtration $\mathcal{F}_t$. As MMs observe different sequences of buy and sell orders they adjust (using Bayes' rule) the estimation of the distribution of v, and in particular they set $p_t = \mathbb{P}(v = V_H \mid \mathcal{F}_t)$, and $\mu_t = \mathbb{E}[v \mid \mathcal{F}_t]$. Then, bid and ask prices will adjust in response to the history of trading, so that

$$a_t = \mu_t + \Delta_{a,t} = \mu_t + \frac{1}{1 + \frac{1-\alpha}{\alpha}\frac{1/2}{p_t}} (V_H - \mu_t) ,$$

and

$$b_t = \mu_t - \Delta_{b,t} = \mu_t - \frac{1}{1 + \frac{1-\alpha}{\alpha}\frac{1/2}{1-p_t}} (\mu_t - V_L) .$$

The resulting bid-ask prices display dynamic changes that reflect the public information embedded in the order flow. Note also that at every execution, the execution price (a_t if it is the execution of a market buy order, and b_t for a sell) is equal to the expectation of the underlying asset conditional on the history of order flow, $\mathcal{F}_t$, and also on the information in the execution (that is a buy or a sell). Hence, it can be seen that the realised price process (the price process at execution times) is a martingale (with respect to the objective measure).

2.3.2 Price Sensitive Liquidity Traders

An interesting extension of the static model (which can be further extended to include the dynamics we have just seen) is to allow liquidity traders to avoid trading if the half-spread, Δ, is too high. A direct way to do this is to assume that liquidity traders get an additional (exogenous) value from executing their desired trade, so that trader i gets a cash equivalent utility gain of c_i if he manages to execute his desired trade. Thus, if the transaction cost imposed by the half-spread is too high, higher than c_i, trader i will prefer not to execute his trade. Assume that the distribution of the parameter c_i in the population of liquidity traders is described by the cumulative distribution function F, such that $F(c)$ is the proportion of liquidity traders that have $c_i \leq c$. We refer to c_i as the agent's urgency parameter.

Then, we can recompute the expected profit of the MM from setting an ask price $a = \mu + \Delta_a$ as above, which will now be given by

$$\frac{(1 - F(\Delta_a))(1-\alpha)/2}{\alpha p + (1 - F(\Delta_a))(1-\alpha)/2}\Delta_a + \frac{p\alpha}{\alpha p + (1 - F(\Delta_a))(1-\alpha)/2}\left(\Delta_a - (V_H - \mu)\right).$$

In this expression we have incorporated the fact that whenever a liquidity trader wants to buy $(1-\alpha)/2$, only $1-F(\Delta_a)$ will have sufficient urgency to execute the trade with a buy-half-spread equal to Δ_a. Introducing this parameter increases the half-spreads, which are now implicitly defined by the following expressions:

$$\Delta_a = \frac{1}{1 + \frac{1-\alpha}{\alpha}\frac{(1-F(\Delta_a))/2}{p}}\left(V_H - \mu\right),$$

and following similar reasoning,

$$\Delta_b = \frac{1}{1 + \frac{1-\alpha}{\alpha}\frac{(1-F(\Delta_a))/2}{1-p}}\left(\mu - V_L\right).$$

A key issue now is that as the MM increases the halfspread, she faces a smaller population of liquidity traders. If the urgency parameters in the population are relatively small, the MM may find that the above expressions have only the extreme solutions $\Delta_a = V_H - \mu$ and $\Delta_b = \mu - V_L$.[5] These extreme solutions correspond to the solutions without liquidity traders and represent market collapse. With those spreads no one gains anything from trade, and any trade that may occur will come from the informed agents who are indifferent to either trading or not trading – though any trade will immediately reveal the underlying value of the asset and the price will be strong-efficient.

2.4 Bibliography and Selected Readings

Grossman (1976), Grossman (1977), Grossman (1978), Ho & Stoll (1981), Gross-

[5] By small urgency parameters we mean that no one has an urgency parameter higher than the expected value of the asset, that is, there exists $\epsilon > 0$, such that $F(\mu - \epsilon) = 1$.

man & Miller (1988), Glosten & Milgrom (1985), Kyle (1985), Kyle (1989), de Jong & Rindi (2009), O'Hara (1995), Abergel et al. (2012), Easley, López de Prado & O'Hara (2012), Vayanos & Wang (2009), SEC (2013*b*), SEC (2013*a*), O'Hara, Yao & Ye (2014), Foucault, Kadan & Kandel (Winter 2005), Rosu (2009), Easley, Engle, O'Hara & Wu (2008), Easley & O'Hara (1992), Biais, Glosten & Spatt (2005), Cartea & Penalva (2012), Boehmer, Fong & Wu (2014), Pascual & Veredas (2009), Martinez & Rosu (2013), Martinez & Rosu (2014), Hoffmann (2014), Cvitanic & Kirilenko (2010), Vives (1996), Colliard & Foucault (2012), Foucault & Menkveld (2008), Gerig (2008), Farmer, Gerig, Lillo & Waelbroeck (2013), Gerig & Michayluk (2010), Cohen & Szpruch (2012), Jarrow & Li (2013), Moallemi & Saglam (2013).

3 Empirical and Statistical Evidence: Prices and Returns

3.1 Introduction

The next two chapters contain empirical analysis of different aspects of trading: prices, returns, spreads, volume, etc., using primarily millisecond stamped data, though we start with daily data that will give us a general overview of the main issues. Chapter 3 focuses on prices and returns, while Chapter 4 is dedicated to volume and market quality measures such as spreads, volatility, or depth.

This chapter, first looks at millisecond data. We then turn to look at the properties of returns both at the daily and at much shorter (one second) time intervals, as well as looking at the interarrival times of price changes. Section 3.4 looks at how market conditions may change when facing latency, as well as the issue of tick size. This is followed by a discussion on price dynamics. Section 3.6 provides a glimpse of the issue of market fragmentation in the US, while the last section provides a first look at the empirics of pairs trading.

In addition to the empirical analysis, we also include plausible interpretations and speculation as to what could be behind some of the results of that analysis. These speculations are included to make the chapter more engaging and to encourage the reader to think about the results. However, they should not be interpreted as anything other than speculative theorising, and should be kept separate from the descriptive analysis of the empirical facts that is limited in scope to the data sample we are using.

3.1.1 The Data

We use data from several sources. For daily and monthly data we use publicly available aggregated data from Yahoo! Finance, and data from the Center for Research in Security Prices (CRSP). We also use millisecond timestamped **ITCH** data (publicly available industry standard data, similar to the direct data feed, recently timestamps go to nanosecond resolution). Our data have been converted into table format for easier processing and is in binary for speed and storage reasons. For illustration purposes we convert these to more human-readable form. The data are made up of the following fields (we drop two fields that are irrelevant here):

- Timestamp: number of milliseconds after midnight