Liquidity and Information in Order Driven Markets

Ioanid Roşu[*]

April 21, 2008

Abstract

This paper analyzes the interaction between liquidity traders and informed traders in a dynamic model of an order-driven market. Agents freely choose between limit and market orders by trading off execution price and waiting costs. In equilibrium, informed patient traders generally submit limit orders, except when their privately observed fundamental value of the asset is far away from the current market-inferred value, in which case they become impatient and submit a market order. As a result, a market buy order is interpreted as a strong positive signal; by contrast, a limit buy order is a much weaker signal, and in some cases even negative. The model generates a rich set of relationships between prices, spreads, trading activity, and volatility. In particular, the order flow is autocorrelated if and only if there are informed traders, with a higher autocorrelation when there is a larger percentage of informed traders, or a larger volatility. Moreover, higher volatility and smaller trading activity generate larger spreads, while a higher percentage of informed traders surprisingly generates smaller spreads.

KEYWORDS: Bid-ask spread, price impact, volatility, trading volume, limit order book, waiting costs.

---

1 Introduction

This article studies the role of information in order-driven markets, where trading is done via limit orders and market orders in a limit order book.[1] Today more than half of the world's stock exchanges are order-driven, with no designated market makers (e.g. Euronext, Helsinki, Hong Kong, Tokyo, Toronto), while in many hybrid markets designated market makers have to compete with a limit order book (NYSE, Nasdaq, London).

Given the importance of order-driven markets, there have been relatively few models which describe price formation in these markets. This is partly due to the difficulty of the problem. Since there is no centralized decision maker, prices arise from the interaction of a large number of traders, each of which can be fully strategic. The presence of traders who are informed about the asset's fundamental value complicates the problem even more.

This paper proposes a dynamic model of an order-driven market where agents can strategically choose between limit and market orders. The model modifies the framework of Roşu(2008) in several important ways. First, the model introduces informed traders, who get a glimpse of the fundamental value $v(t)$ of an asset (assumed to follow a diffusion process). The market then forms an expectation $\hat{v}(t)$ (the efficient price), based on all the publicly available information.

Second, the liquidity traders use the efficient price $\hat{v}$ to make an initial choice whether or not to trade. This is done by comparing their expected profit with a private one-time cost, uniformly and independently distributed on an interval $[-C, C]$. This assumption generates a downward-sloping demand function from liquidity buyers, and an upward-sloping supply function from liquidity sellers. For example, the probability that a liquidity seller enters the market increases with his expected utility (expected price minus waiting costs). This has two main implications. One, the limit order book naturally becomes resilient, i.e. the bid-ask spread tends to revert to small values, and the limit order book tends to be centered around the mid-point (zero in this case). The other consequence is that the bounds of the limit order

---

[1]A limit order is a price-contingent order to buy (sell) if the price falls below (rises above) a prespecified price. A sell limit order is also called an *offer*, while a buy limit order is also called a *bid*. The limit order book (or simply the book) is the collection of all outstanding limit orders. The lowest offer in the book is called the *ask price*, or simply *ask*, and the highest bid is called the *bid price*, or simply *bid*.

book become endogenously determined.

Most other results from Roşu(2008) hold true in the current model: e.g. the existence of discrete spreads in this model (due to waiting costs); and the comovement effect between bid and ask prices, which was obtained in the absence of information.[2]

How does private information affect the limit order book? First, the limit order book is always centered around the efficient price $\hat{v}$, but otherwise depends only on the number of sellers m and number of buyers n in the book. The strategy of a patient informed trader depends on how far the fundamental price v is from $\hat{v}$. If $v - \hat{v}$ is above or below two cutoffs (that depend on the state of the book), the patient informed trader optimally behaves in an impatient way and submits a market order. By contrast, an impatient informed trader only submits market orders: a buy market order, if v is larger than the ask; a sell market order, if v is smaller than the bid; and does nothing if v is between the bid and the ask.

As a result, the efficient price $\hat{v}$ tends to always converge towards the fundamental value v. The speed of convergence depends on how the market reacts to market and limit orders. Since a fundamental value v away from $\hat{v}$ makes market orders more likely, market orders are correctly interpreted by the market to contain a lot of information about v. For example, a buy market order is a clear positive signal: with positive probability it comes from either a patient informed trader (and so the fundamental value v is higher than the efficient price $\hat{v}$ plus a cutoff), or from an impatient trader (so v is larger than the ask price). By contrast, it can be shown that a limit buy order is a weaker positive signal, and in some cases it can even be a negative signal. Notice that a higher percentage of informed traders means a higher adjustment of the efficient price $\hat{v}$ to a market order, which means that prices converge faster when there there are more informed traders.

In general, the model generates a rich set of relationships between prices, spreads, trading activity, volatility, and information asymmetry (measured by the percentage of informed

---

[2]The comovement effect is the fact that e.g. a market sell order not only decreases the bid price—this can in part be due to the mechanical execution of limit orders on the buy side—but also decreases the ask price. Moreover, the decrease in the bid price is larger than the decrease in the ask price, which leads to a wider bid-ask spread. The intuition is the following: the decrease in the bid price lowers the reservation value for the sellers. So competition between them also drives the ask price down, albeit by a lesser amount: this is because the reservation value only becomes fully relevant in the future state when the bid-ask spread is at a minimum.

traders). In particular, consistent with previous literature, one can show that smaller trading activity and higher volatility generate larger spreads. Smaller trading activity implies a smaller flow of impatient traders, which increases the waiting costs of patient traders, and therefore makes them separate each other with higher spreads. Higher volatility makes extreme fundamental values more likely, and increases the percentage of market orders (from patient informed traders), which makes the spreads wider.

A surprising new prediction is that, controlling for volatility and trading activity, a higher percentage of informed traders should generate *smaller* spreads. This is because a higher percentage of informed traders generates a quicker adjustment of prices to fundamentals, and therefore extreme values of v become less likely. This generates fewer market orders from patient informed sellers, and hence smaller spreads. This prediction raises some interesting questions about why spreads are larger around earnings announcements. According to this paper, it is not the amount of asymmetric information, since this by itself would generate smaller spreads. Then it must be due to either lower trading activity, or to larger volatility, or perhaps a combination of both. While it is not clear why there would necessarily be lower trading activity in times of high uncertainty, one may argue that the high volatility surrounding these events causes the higher spreads just by itself.

Also, this model contributes towards an explanation of the "diagonal effect" of Biais, Hillion and Spatt (1995), namely that the order flow is positively autocorrelated (e.g. a market buy order makes a future market buy order more likely). In this model one can show that the order flow is autocorrelated if and only if there exist informed traders, with a higher autocorrelation when there is a larger percentage of informed traders, or a larger volatility. To see this, consider the case with only liquidity traders. In that case, we saw that the limit order book is resilient and tends to revert the the mid-point $\hat{v}$. This means that the ex-ante entry probability is approximately equal for all types of traders. So since the market arrivals are independent and the entry decision is made with approximately the same probability, the order flow is approximately uncorrelated (i.e. a buy market order does not make future buy market orders more likely). By contrast, in the presence of informed traders, a market order reflects a fundamental price v far away from the efficient price $\hat{v}$, and makes the (ex-post) probability of another market order relatively higher. In conclusion, the present model

produces a diagonal effect as in Biais, Hillion and Spatt (1995), and makes the extra prediction that the diagonal effect should be stronger when volatility or asymmetric information are higher.

Can one use this paper to estimate the extent of information asymmetry in the market (e.g. by using the level of autocorrelation of order flow)? The answer is, it depends on what one means by information asymmetry. What if for example some agents do not observe the fundamental value v, but instead have information about the future order flow (they are informed in the sense of Evans and Lyons, 2002)? Then these agents would would behave in the same way as the informed agents in the current model, and become impatient if the information about the future order flow is extreme enough. Could we then distinguish these agents from truly informed ones? Perhaps, if one looked at price reversals. If the order-flow-informed traders do not really have information about v, then the efficient price $\hat{v}$ would then revert to the true value v after the truly informed traders have brought prices in line with fundamentals.[3]

Notice that this model in the end generates two types of resilience. One is the resilience relative to the limit order book (we can call it "micro resilience"), and is due to the action of discriminatory liquidity traders. The other makes the efficient price $\hat{v}$ eventually converge to the fundamental value (we can call it "macro resilience"), and is due to the action of informed traders.

The choice between market orders and limit orders has been analyzed in various contexts, see e.g. Chakravarty and Holden (1995), Cohen, Maier, Schwartz and Whitcomb (1981), Handa and Schwartz (1996), Kumar and Seppi (1993). Dynamic models of order-driven markets include Foucault (1999), Foucault, Kadan and Kandel (2003), Parlour (1998), and Roşu(2008). The price behavior in limit order books has been analyzed theoretically by Biais,

---

[3]Information about order flow is not the only alternative explanation of the diagonal effect. Another explanation comes from the possibility of large orders (if one relaxes the assumption that agents can only trade one unit of the asset). Then an agent who wants to trade a large quantity and is patient enough to work the order (divide it into smaller orders) and thus take advantage of the resilience of the limit order book would also generate a positively autocorrelated order flow. Can one test if this is the right explanation? If one had access to identity of the order flow, then one could check if there was just one trader working the order. If so, then the model could still incorporate working orders, by treating each order as coming from a separate trader. In that case, one finds again the problem of distinguishing information about fundamentals from information about order flow.

Martimort and Rochet (2000), Glosten (1994), OHara and Oldfield (1986), Rock (1990), and Seppi (1997). Models that analyze liquidity traders, the dynamics of prices and trades and the convergence of prices to the fundamental value include Glosten and Milgrom (1985), Kyle (1985), Admati and Pfleiderer (1988), Easley and O'Hara (1987).

Empirical papers include Biais, Hillion, and Spatt (1995), who document the diagonal effect (positive autocorrelation of order flow) and the comovement effect (e.g. a downward move in the bid due to a large sell market order is followed by a smaller downward move in the ask – which increases the bid-ask spread); Sandas (2001), who uses data from the Stockholm exchange to reject the static conditions implied by the information model of Glosten (1994), and also finds that liquidity providers earn superior information; Harris and Hasbrouck (1996) who obtain a similar result for the NYSE SuperDOT system; Hollifield, Miller and Sandas (2004) who test monotonicity conditions resulting from a dynamic model of the limit order book and provides some support for it; Hollifield, Miller, Sandas and Slive (2006) who use data from the Vancouver exchange to find that agents supply liquidity (by limit orders) when it is expensive and demand liquidity (by market orders) when it is cheap.

2 The Model

This section describes the assumptions of the model. Consider a market for an asset which pays no dividends, and whose fundamental value (or *full-information* price) $v(t)$ moves according to a diffusion process with constant volatility σ: $dv(t) = \sigma \, dW(t)$, where $W(t)$ is a standard Brownian motion. Based on all available public information until t, the market forms an estimate (the *efficient* price): $\hat{v}(t) = \mathsf{E}\{v(t) \mid \text{Public Information at } t\}$. For simplicity, it is assumed that $v(t)$ is normally distributed: $v(t) \sim N\big(\hat{v}(t), \hat{\sigma}(t)\big)$, with mean $\hat{v}(t)$ and standard deviation $\hat{\sigma}(t)$.[4]

The buy and sell prices for this asset are determined as the bid and ask prices resulting from trading based on the rules given below. Prices can take any value, i.e., the tick size is zero.

---

[4]The connection between this assumption and the fact that $v(t)$ is a diffusion process is discussed in Footnote 12.

Trading The time horizon is infinite, and trading in the asset takes place in continuous time. The only types of trades allowed are market orders and limit orders, which are executed with no delays. There is no cost of cancellation for limit orders.[5]

Trading is based on a publicly observable limit order book, which is the collection of all the limit orders that have not yet been executed. The limit orders are subject to the usual price priority rule; and, when prices are equal, the time priority rule is applied. If several market orders are submitted at the same time, only one of them is executed, at random, while the other orders are canceled.

Agents The market is composed of two types of agents: liquidity traders and informed traders. Both types of traders can be patient and impatient, in a sense to be precisely described below. They trade at most one unit, after which they exit the model forever. The traders' types are fixed from the beginning and cannot change.

All agents in this model are risk-neutral, so their instantaneous utility function (felicity) is linear in price. By convention, felicity is equal to price for sellers, and minus the price for buyers. Traders discount the future in a way proportional to the expected waiting time. If τ is the random execution time and P_τ is the price obtained at τ, the expected utility of a seller is $f_t = \mathsf{E}_t\{P_\tau - r(\tau - t)\}$. (The expectation operator takes as given the strategies of all the players.) Similarly, the expected utility of a buyer is $-g_t = \mathsf{E}_t\{-P_\tau - r(\tau - t)\}$, where by notation $g_t = \mathsf{E}_t\{P_\tau + r(\tau - t)\}$. One calls f_t the value function, or utility, of the seller at t; and similarly g_t is the value function, or utility, of the buyer, although in fact g_t equals minus the expected utility of a buyer.

The discount coefficient r is constant.[6] It can take only two values: if it is low, the corresponding traders are called *patient*, otherwise they are *impatient*. For simplicity, it is assumed that the impatient agents always submit market orders. From now on, denote by r only the time discount coefficient of the patient agents.

---

[5]In most financial markets cancellation of a limit order is free, although one may argue that there are still *monitoring* costs. The present model ignores such costs, but one can take the opposite view that there are infinite cancellation / monitoring costs. See e.g. Foucault, Kadan and Kandel (2005).

[6]The nature of waiting costs is intentionally vague in this paper. One can interpret it as an opportunity cost of trading. Another interpretation is that waiting costs reflect traders' uncertainty aversion: if uncertainty increases with the time horizon, an uncertainty averse trader loses utility by waiting.

Liquidity Traders These are of four types: patient buyers, patient sellers, impatient buyers, and impatient sellers. All types of liquidity traders arrive at the market according to independent Poisson processes with the same arrival intensity rate λ^U.[7] They are liquidity traders, in the sense that they want to trade the asset for reasons exogenous to the model. But they do have discretion whether to enter the market, and once they enter, whether to use market or limit orders.

The decision to enter the market is based on a private one-time cost of trading c, uniformly distributed on the interval $[-C, C]$. This means that each trader who arrives at the market makes the entry decision based on the private cost c. E.g. a seller who expects utility f from trading must satisfy $f - \hat{v} \geq c$.[8] This assumption generates an upward-sloping supply of sell orders: the probability that a liquidity seller enters the market increases with expected utility f. Similarly, this also generates a downward-sloping supply of buy orders: the probability that a liquidity buyer enters the market decreases with their expected utility.

For simplicity, it is assumed that after the submission of an order the market adjusts the efficient price $\hat{v}$ to the appropriate extent, but then it immediately forgets the history of orders. This does not change the qualitative behavior of the model, but it does change the quantitative behavior.

Informed Traders These are of two types: patient and impatient. Both types of informed traders arrive at the market according to independent Poisson processes with the same arrival intensity rate λ^I.

When informed traders arrive at time t, they observe the fundamental value $v(t)$, and decide whether that want to become buyers or sellers, and whether they want to use market or limit orders. For example, an informed patient trader becomes a seller and submits a limit order (which gives expected utility f) if the difference $f - v(t)$ is the maximum possible. After the initial entry decision, it is further assumed that the informed agents do not use

---

[7] By definition, a Poisson arrival with intensity λ implies that the number of arrivals in any interval of length T has a Poisson distribution with parameter λT. The inter-arrival times of a Poisson process are distributed as an exponential variable with the same parameter λ. The mean time until the next arrival is then $1/\lambda$.

[8] Alternatively, one can think of $\hat{v} + c$ as the seller's private opinion about the fundamental value; the seller trades only if the expected utility f is above the opinion $\hat{v} + c$.

their information anymore. This is admittedly an important assumption, but in Section 4 it will be argued that the solution described in this paper is robust to this assumption.

Strategies Since this is a model of continuous trading, it is desirable to set the game in continuous time. There are also technical reasons why that would be useful: in continuous time, with Poisson arrivals the probability that two agents arrive at the same time is zero, and this simplifies the analysis of the game. But setting the game in continuous time requires extra care, see Roşu(2008) for details.

3   Benchmark: No Informed Traders

Without informed traders, the efficient price $\hat{v}$ does not evolve in time, and so one can assume that $\hat{v} = 0$. The intuition for the solution follows that of the model in Roşu(2008). In equilibrium, the limit order book is formed only of limit orders submitted by patient liquidity traders. If limit orders are pictured on the vertical axis, the limit order book is formed by two queues: the sell (or ask, or offer) side, with sell limit orders descending to a minimum price, called the *ask price*; and the buy (or bid) side, with buy limit orders ascending to a maximum price, called the *bid price*. The two sides are disjoint, i.e. the ask price is above the bid price.

3.1   Equilibrium

In the equilibrium limit order book, the patient limit sellers compete for the incoming market orders from impatient buyers, and the patient buyers compete for the order flow from impatient sellers. The sellers for example have their limit orders placed at different prices, but they get the same expected utility: otherwise, they would undercut by a penny those with higher utility. Thus, the sellers with a higher limit order obtain in expectation a higher price, but also have to wait longer. Similarly, all the buyers have the same expected utility. This makes the equilibrium Markov, and the numbers of buyers and sellers in the book becomes a state variable.

Denote by m the number of sellers, and by n the number of buyers in the limit order book at a given time. Denote by $a_{m,n}$ the ask price, $b_{m,n}$ the bid price, $f_{m,n}$ the expected utility of

the sellers, and $g_{m,n}$ (minus) the expected utility of the buyers, as defined in Section 2. One can prove the following formulas:

$$a_{m,n} = f_{m-1,n} \quad \text{and} \quad b_{m,n} = g_{m,n-1}, \tag{1}$$

which follows from the fact that e.g. all sellers have the same utility in state (m,n): if an impatient buyer comes, then the bottom seller gets the ask price $a_{m,n}$, while all the other sellers get $f_{m-1,n}$. As in Roşu(2008), one defines the *state region* Ω as the collection of all pairs (m,n) where in equilibrium the m sellers and the n buyers wait in expectation for some positive time. Also one defines the *boundary* γ of Ω as the set of (m,n) where at least some agent has a mixed strategy. This is the set of states where the limit order becomes full, and any extra incoming patient seller or buyer would immediately place a market order.

Recall that the each agent's decision to enter the market is based on a private one-time cost of trading c, uniformly distributed on the interval $[-C, C]$. Take for example an impatient seller with cost c who arrives at the market in state (m,n). If the seller placed a sell market order, he would get the bid price $b_{m,n}$. According to the assumption, the seller then decides to enter the market and place a market order if and only if $b_{m,n} - \hat{v} = b_{m,n} \geq c$. The ex-ante probability of this event is $P(c \leq b_{m,n}) = \frac{1+b_{m,n}/C}{2}$. Similarly, if a patient seller decided to place a limit order, the book would go to the state $(m+1, n)$ with one extra limit seller, which yields expected utility of $f_{m+1,n}$. This even happens with ex-ante probability of $P(c \leq f_{m+1,n}) = \frac{1+f_{m+1,n}/C}{2}$. Also, an impatient buyer would place a market order at the ask price $a_{m,n}$ if the gain $\hat{v} - a_{m,n} = -a_{m,n} \geq c$. This event has the ex-ante probability $P(c \leq -a_{m,n}) = \frac{1-a_{m,n}/C}{2}$.

From now on normalize $C = 1$. From a typical state (m,n) the system can go to the following neighboring states:

- $(m-1, n)$, if an impatient buyer arrives and places a market order at the ask;

- $(m+1, n)$, if a patient seller arrives and submits a limit order;

- $(m, n-1)$, if an impatient seller arrives and places a market order at the bid;

- $(m, n+1)$, if a patient buyer arrives and submits a limit order.

10

If a trader arrives at the market, but decides to place no order, the book remains in state (m, n).

Any of these four events arrives at a random time which is exponentially distributed with parameter λ^U, therefore the first of them arrives at a random time which is exponentially distributed with parameter $4\lambda^U$. This means that a patient seller loses an expected utility of $r \cdot \frac{1}{4\lambda^U}$ from waiting in state (m, n). Now, conditional on the first event, the arrival e.g. of an impatient buyer happens with probability $1/4$. Once the impatient buyer arrives, he places a market with probability $\frac{1-a_{m,n}}{2}$; otherwise, with probability $\frac{1+a_{m,n}}{2}$ he decides not to do anything, in which case the market remains in the same state. Doing this analysis for all states, we get

$$
\begin{aligned}
f_{m,n} &= \frac{1}{4}\left(\frac{1-a_{m,n}}{2}f_{m-1,n} + \frac{1+a_{m,n}}{2}f_{m,n}\right) + \frac{1}{4}\left(\frac{1+f_{m+1,n}}{2}f_{m+1,n} + \frac{1-f_{m+1,n}}{2}f_{m,n}\right) \\
&\quad + \frac{1}{4}\left(\frac{1+b_{m,n}}{2}f_{m,n-1} + \frac{1-b_{m,n}}{2}f_{m,n}\right) + \frac{1}{4}\left(\frac{1-f_{m,n+1}}{2}f_{m,n+1} + \frac{1+f_{m,n+1}}{2}f_{m,n}\right) \\
&\quad - r \cdot \frac{1}{4\lambda^U}.
\end{aligned}
\tag{2}
$$

A similar recursive system of difference equations holds for g, the expected utility of a buyer in state (m, n). It is not difficult to see that a solution to a system of recursive equations leads to a Markov equilibrium of this market. (See Theorem 3 in Roşu(2008).) The description of the solution is very difficult, because of the need to simultaneously determine the state space Ω and its boundary γ (where the book becomes full), and find a solution to the recursive system. To get a better understanding what the solution looks like, one can simplify the problem by looking only at the sell side of the book. This is a model where the liquidity traders are only patient sellers and impatient buyers.

3.2 One-Sided Limit Order Book

Assume now that the liquidity traders arriving at the market are only patient sellers and impatient buyers. Moreover, assume that there is a large supply of limit buy orders at zero, so that prices never go below zero. This is an artificial assumption (similar to the assumption of exogenous bounds for the limit order book, as in Foucault, Kadan and Kandel (2005) and

Roşu(2008)) which ensures that the book always has a bid price of zero.

The model for the one-sided book is easier to solve (although it still has to be solved numerically), and it provides important intuition for the two-sided case. In equilibrium, each new patient seller arrives to the market and places a limit order inside the bid-ask spread, thus lowering the ask price.[9] There is only one exception: there is a state where all the patient sellers have zero utility, in which case a new incoming patient seller has no incentive to wait and instead places a market order at the bid price and exits. In this case the limit order book is called "full," and there is a maximum number M of sellers, while the bid-ask spread is at a minimum.

As in the two-sided case, if the book has m patient sellers, they all have the same expected utility f_m. Then m becomes a state variable and f_m satisfies a recursive system (a_m is the ask price in state m): $f_m = \frac{1}{2}\left(\frac{1-a_m}{2}f_{m-1} + \frac{1+a_m}{2}f_m\right) + \frac{1}{2}\left(\frac{1+f_{m+1}}{2}f_{m+1} + \frac{1-f_{m+1}}{2}f_m\right) - r \cdot \frac{1}{2\lambda^U}$. Equation (1) in the one-sided case implies that the ask price a_m when there are m sellers in the book equals to the expected utility f_{m-1} in the state $m-1$ with one less seller.[10]

$$a_m = f_{m-1} \quad \text{for all } m = 2, \ldots, M. \tag{3}$$

This is true for $m < M$, when the limit order book is not full. In the case when $m = M$ any new seller submits a market order at zero and exits, and so the state $M+1$ never exists. Therefore the recursive equation becomes $f_M = 0 = \frac{1}{2}\left(\frac{1-f_{M-1}}{2}f_{M-1}\right) - \frac{r}{2\lambda^U}$. Now define the *granularity* parameter

$$\varepsilon = \frac{r}{\lambda^U}. \tag{4}$$

One gets the recursive equation

$$f_m = \frac{1}{2}\left(\frac{1-f_{m-1}}{2}f_{m-1} + \frac{1+f_{m-1}}{2}f_m\right) + \frac{1}{2}\left(\frac{1+f_{m+1}}{2}f_{m+1} + \frac{1-f_{m+1}}{2}f_m\right) - \frac{\varepsilon}{2}. \tag{5}$$

with $f_M = 0$ and $f_{M-1} = \frac{\varepsilon}{1+\sqrt{1-2\varepsilon}}$. This suggests that one could find f_m numerically by

---

[9]Biais, Hillion and Spatt (1995) empirically show in their study of the Paris Bourse (now Euronext) that the majority of limit orders are spread improving.

[10]Just like in Theorem 1 of Roşu(2008), the equation is true when $m < M$. When $m = M$ the equation is true only if one assumes that the bottom agent does not have a mixed strategy. Since we are interested here only in what happens to the average bid-ask spread, the single state $m = M$ does not affect such calculations.

rewriting the recursive equation in terms of f_{m-1} as a function of f_m and f_{m+1}:

$$f_{m-1} = \frac{1}{2}\left(1 + f_m - \sqrt{(1+f_m)^2 - 4\left[f_m(2+f_{m+1}) - f_{m+1}(f_{m+1}+1) + 2\varepsilon\right]}\right). \quad (6)$$

What are the bounds of the limit order book? In Roşu(2008), as in Foucault, Kadan and Kandel (2005), the boundaries are assumed to be exogenous. The present model gives a way to endogenize them. In the one-sided case, the lower bound is exogenous (zero), so one needs only to show how the upper bound is determined.

Proposition 1. *Let f_m be the utility of sellers in state $m = 1, 2, \ldots, M$, given by the solution of the recursive equation (5). Then the level a_1 of the sole limit order in state $m = 1$ is set at the monopoly price*

$$a_1 = \frac{1 + f_1}{2}.$$

Proof. See the Appendix. □

3.3 Resilience

The next numerical result shows that the average bid-ask spread and price impact are of the order of the granularity parameter $\varepsilon = \frac{r}{\lambda^U}$ to some power less than one. This shows that a higher trading activity λ^U and higher patience (lower r) indeed generate smaller spreads. This is because when agents do not have to wait much (high λ) or do not mind waiting (low r), they tolerate staying closer to each other, which generates smaller spreads.

Numerical Proposition 1. *Let granularity ε run over $10^{-1}, 10^{-2}, \ldots, 10^{-16}$. For each ε compute the solution f_m to the recursive system that describes the utility of the m sellers in the book. Denote by x_m the Markov stationary probability that the limit order book is in state m (has m sellers). Since f_m is the bid-ask spread in state m, denote by $\bar{s} = \sum_{m=0}^{M} \pi_m f_m$ the average bid-ask spread. Then regressing $\log(\bar{s})$ on $\log(\varepsilon)$ gives the approximate formula $\log(\bar{s}) \approx -0.42 + 0.34 \log(\varepsilon)$, with $R^2 = 0.9996$. This means that with a very good approximation the average spread*

$$\bar{s} \approx 0.66\, \varepsilon^{0.34}. \quad (7)$$

Moreover, denote by $I_m = a_{m-1} - a_m = f_{m-2} - f_{m-1}$ the price impact of a one-unit market order in state m. Denote by $\bar{I} = \sum_{m=0}^{M} \pi_m I_m$ the average price impact I_m. Then regressing $\log(\bar{I})$ on $\log(\varepsilon)$ gives the approximate formula $\log(\bar{I}) \approx 1.15 + 0.68 \log(\varepsilon)$, with $R^2 = 0.9998$. This means that with a very good approximation the average one-unit price impact

$$\bar{I} \approx 3.15 \, \varepsilon^{0.68}. \tag{8}$$

Proof. See the Appendix to see how the stationary probabilities π_m are computed. $\qquad\square$

The Numerical Proposition 1 shows that the average spread gets very close to zero when the granularity $\varepsilon = r/\lambda^U$ is small. Moreover, one can check that the system is more likely to remain in states with more sellers, i.e. π_m is increasing in m. This indicates that the one-sided limit order book is resilient: the bid-ask spreads tend to revert to the minimum value $a_M = f_{M-1} = \frac{\varepsilon}{1+\sqrt{1-2\varepsilon}}$.

Now, coming back to the two-sided case, recall that the mid-point of the book is zero (the mid-point of the exogenous private cost interval $[-1, 1]$ for the liquidity traders). Now, suppose the limit order book has m sellers and n buyers, i.e. is in state (m, n). Then the probability of entry for an impatient buyer (conditional on arrival to the market) is $(1 - a_{m,n})/2$, where $a_{m,n}$ is the ask price; and the probability of entry for a patient seller is $(1 + f_{m+1,n})/2$ which is the utility of patient sellers in state $(m + 1, n)$ with one more seller. When the granularity ε is small, $a_{m,n} = f_{m-1,n}$ and $f_{m+1,n}$ are very close to each other, with a difference of the order of $\varepsilon^{0.68}$.[11] If these two numbers are not close to zero, for example if $a_{m,n}$ is much significantly larger than zero, then $(1 - a_{m,n})/2$ is significantly smaller than $(1 + f_{m+1,n})/2$, which means that patient sellers arrive significantly faster than impatient buyers. This drives the ask price down to the point $a_{m,n}$ is not significantly larger than zero. A similar argument works for patient buyers and impatient sellers, which also brings the bid price $b_{m,n}$ close to zero. In conclusion, the two-sided limit order book is also resilient, and moreover tends to be centered around zero. According to the Numerical Proposition 1, the average bid-ask spread is of the order of $\varepsilon^{0.34}$, and the distance between the mid-point of the bid-ask spread and zero is also

---

[11]According to the Numerical Proposition 1, $f_{m-1,n} - f_{m+1,n} = (f_{m-1,n} - f_{m,n}) + (f_{m,n} - f_{m+1,n})$, and each term is of the order of $\varepsilon^{0.68}$.

of the order of $\varepsilon^{0.34}$. This result is to be compared with that of Farmer, Patelli and Zovko (2003), who in their cross-sectional empirical analysis of the London Stock Exchange find that the average bid-ask spread varies proportionally to $\varepsilon^{0.75}$.

In the theoretical model of Roşu(2008) there are two cases, depending on the competition parameter c, which is the ratio of arrival rate of patient traders to the impatient traders. When $c = 1$, the average bid-ask spread does not depend on the granularity parameter ε, and is in fact very large: $\bar{s} = (A - B)/2$, where A and B are the bounds of the limit order book. When $c > 1$, i.e. patient traders arrive faster than impatient traders, the book becomes resilient, and the average bid-ask spread becomes of the order of $\varepsilon \ln(1/\varepsilon)$. Notice that the present model is a mixture of the cases $c = 1$ and $c > 1$: when the bid and ask prices are close to zero, c is close to one, while when e.g. the ask price is significantly larger than zero, c is significantly larger than one.

4   General Case: Informed Traders

Now we tackle the general case, when besides the discretionary uninformed liquidity traders, patient and impatient buyers and sellers, there are also patient and impatient informed traders. A newly arrived informed trader observes the fundamental value (or *full-information price*) v, which is assumed to follow a diffusion process with constant volatility σ. The market forms an estimate $\hat{v}$ (called the *efficient price*) of v based on all publicly available information. It is further assumed that in each state of the book the fundamental value v is normally distributed with mean $\hat{v}$ and volatility $\hat{\sigma}$.[12]

How does an informed trader behave in such an environment? As it will be seen in Section 4.2, the strategy of a patient informed trader depends on how far the fundamental value v is from $\hat{v}$, i.e. how it behaves with respect to three cutoffs (that depend on the state of the book). If $v - \hat{v}$ is below the lower cutoff, the patient informed trader optimally behaves in an impatient way and submits a market sell order. If $v - \hat{v}$ is above that lower cutoff but

---

[12]The market volatility $\hat{\sigma}$ need not be constant, but by a stationarity argument similar to the volatility story in Section 4.3 one can to determine the relationship between $\hat{\sigma}$ and the fundamental volatility σ. Indeed, after each order the informativeness of the signal reduces the conditional volatility of $\hat{\sigma}$. But after some inter-transaction time elapses, because of the diffusion process that v satisfies, the uncertainty in σ builds back.

still below the middle cutoff, the patient seller submits a sell limit order and waits. Above the middle cutoff but below the upper cutoff, the trader places a limit buy order and waits, and finally above the upper cutoff the patient traders submits a market buy order. By contrast, an impatient informed trader only submits market orders: a buy market order, if v is larger than the ask; a sell market order, if v is smaller than the bid; and does nothing if v is between the bid and the ask.

4.1   Uninformed (Liquidity) Traders

How do the uninformed traders adjust to the presence of informed traders? The answer is relatively simple: the limit order book shifts up and down along with the efficient price $\hat{v}$, but its translation by $-\hat{v}$ moves the limit order book to a canonical limit order book centered at zero.

Proposition 2. *There is a canonical equilibrium in which the shape of the limit order book does not depend on the efficient price $\hat{v}$, and up to a translation only depends on the number of sellers m and number of buyers n in the book. The limit order book is centered at $\hat{v}$.*

Proof. The shape of the book depends on the arrival rates and strategies of the various market participants. The assumptions of Section 2 imply that the behavior of both the uninformed and informed traders only depends on the state of the book when they arrive: For the uninformed liquidity traders, entry is based on a one-time private cost that is compared to the expected profit from entering the book. For the informed traders, it was assumed that after the initial entry decision they do not use their information anymore. So as long as the strategy of the informed traders involves only the difference between v and $\hat{v}$, the existence of the canonical equilibrium follows in a straightforward way. In Section 4.2 it will be shown that indeed the strategy of the informed trader only depends on the difference $v - \hat{v}$ and the state of the book (m, n).   □

The previous result shows that one can consider the canonical equilibrium in the interval $[-1, 1]$, just like the equilibrium in Section 3.1. The difference is now that there are also informed traders, both patient and impatient. Their strategies depend on the difference $v - \hat{v}$

between their privately observed fundamental value v and the efficient price $\hat{v}$. The strategy is described in detail in Section 4.2 as a function that maps the informed trader of type $i \in \{PI, II\}$ (patient informed, impatient informed) and a pair $\big(v - \hat{v}, (m, n)\big)$ into a decision $j \in \{SMO, SLO, BLO, BMO, NO\}$ (sell market order, sell limit order, buy limit order, buy market order, or no order). For now it suffices to summarize the order-submission strategy by the market-estimated probabilities P_j^i that an informed trader i submits an order j.

To describe the canonical equilibrium, one gets a system of recursive equations similar to (2). This is written explicitly in Equation (12) in the Appendix. Such a recursive system is probably impossible to solve explicitly, but in Section 4.4 some qualitative results are given concerning the average spreads and price impact in a limit order book with informed traders.

4.2   Informed Traders

The strategy of an informed trader depends on how far the fundamental price v is from $\hat{v}$, and on the current state (m, n) of the book (with m sellers and n buyers). To understand better what this strategy is based on, consider an informed trader of type $(i \in \{PI, II\})$ who contemplates the choice of an order $j \in \{SMO, SLO, BLO, BMO, NO\}$.

Proposition 3. *Consider the decision of an informed trader of type $i \in \{PI, II\}$ (patient informed, impatient informed) to enter a limit order book in state (m, n) (with m sellers and n buyers). Denote by u_j^i the expected utility from choosing an order $j \in \{BMO, SLO, SMO, BLO, NO\}$ (sell market order, sell limit order, buy limit order, buy market order, or no order). Define*

$$k = \frac{3\lambda^U + 4\lambda^I}{4\lambda^U + 4\lambda^I} \in (3/4, 1)$$

which is increasing in the ratio of informed to uninformed λ^I / λ^U. Then u_j^i is given by

- $u_{SMO}^{PI} = u_{SMO}^{II} = \hat{v} + b_{m,n} = \hat{v} + g_{m,n-1}$,
- $u_{SLO}^{PI} \approx \hat{v} + f_{m+1,n} + k(v - \hat{v})$, *when v is sufficiently low.*
- $u_{BLO}^{PI} \approx \hat{v} + g_{m,n+1} + k(v - \hat{v})$, *when v is sufficiently high.*
- $u_{BMO}^{PI} = u_{BMO}^{II} = \hat{v} + a_{m,n} = \hat{v} + f_{m-1,n}$,
- $u_{NO}^{PI} = u_{NO}^{II} = v$.

17

Proof. See the Appendix. □

One can use the previous Proposition to derive the strategies of the informed traders.

Corollary 1. *For a typical limit order book with $f_{m,n} > 0$, $g_{m,n} < 0$, denote by $s_{m,n} = a_{m,n} - b_{m,n}$ the bid-ask spread in the state (m,n), with m sellers and n buyers, and by $p_{m,n} = \frac{a_{m,n}+b_{m,n}}{2}$ the mid-point of the bid-ask spread. Then there are three (approximate) cutoffs: $-\frac{s_{m,n}}{k} < \frac{p_{m,n}}{1-k} < \frac{s_{m,n}}{k}$, where $k = \frac{3\lambda^U + 4\lambda^I}{4\lambda^U + 4\lambda^I} \in \left(\frac{3}{4}, 1\right)$, such that the optimal strategy of the patient informed trader in state (m,n) is to submit: a sell market order if $v - \hat{v} < -\frac{s_{m,n}}{k}$; a sell limit order if $v - \hat{v} \in \left(-\frac{s_{m,n}}{k}, \frac{p_{m,n}}{1-k}\right)$; a buy limit order if $v - \hat{v} \in \left(\frac{p_{m,n}}{1-k}, \frac{s_{m,n}}{k}\right)$; and a buy market order if $v - \hat{v} > -\frac{s_{m,n}}{k}$.*

The strategy of an impatient informed trader is simpler, in that he only submits market orders: a sell market order if $v - \hat{v} < b_{m,n}$ (v is smaller than the bid), a buy market order if $v - \hat{v} > a_{m,n}$ (v is larger than the ask), and do nothing if $v - \hat{v} \in (a_{m,n}, b_{m,n})$ (v is between the bid and the ask).

Proof. See the Appendix. □

4.3 Behavior of the Efficient Price

As a result of the strategies of informed traders, the efficient price $\hat{v}$ always converge towards the fundamental value v. To see how this is done, consider the market reaction to a sell market order. By assumption, the market knows that v is distributed normally with mean $\hat{v}$ and mean $\hat{\sigma}$. Since the informed traders choose a sell market order only when the fundamental value v is below a certain cutoff (see Proposition 1), this leads to a downward adjustment to the efficient price following a sell market order (SMO). The exact formula is given by the following result.

Proposition 4. *Suppose the limit order book is in state (m,n) with m sellers and n buyers, and the efficient price equals $\hat{v}$. Denote by $\alpha_{m,n} = \frac{s_{m,n}}{k}$, where $s = s_{m,n} = a_{m,n} - b_{m,n}$ is the bid-ask spread and $k = \frac{3\lambda^U + 4\lambda^I}{4\lambda^U + 4\lambda^I}$. Then, following a sell market order, the efficient price $\hat{v}$*

changes to

$$E(v|SMO) \;=\; \hat{v} - \hat{\sigma} \; \frac{\mathrm{e}^{-\frac{\alpha_{m,n}^2}{2\hat{\sigma}^2}} + \mathrm{e}^{-\frac{b_{m,n}^2}{2\hat{\sigma}^2}}}{\Phi\left(-\frac{\alpha_{m,n}}{\hat{\sigma}}\right) + \Phi\left(\frac{b_{m,n}}{\hat{\sigma}}\right) + \frac{\lambda^U}{\lambda^I}}, \tag{9}$$

where $\Phi(x)$ is the cumulative density function for the standard normal distribution.

This shows that the size of the adjustment increases with the ratio $\frac{\lambda^I}{\lambda^U}$, and therefore with the percentage of informed traders of the total population. So, not surprisingly, prices converge faster to the fundamental value when there are more informed traders. Notice that the adjustment has two contributions, one from the informed patient traders, who only submit a market order when the mispricing $v - \hat{v}$ is below the cutoff $-\alpha_{m,n}$. The other is from the informed impatient traders, who "fine tune" the adjustment by submitting a market order every time the fundamental value v is below the bid price $\hat{v} + b_{m,n}$.

One of the goals of this section is to understand observed prices and volatility. The efficient price $\hat{v}$ is in principle not observed by the econometrician (although it is observed by the traders in the limit order book). But because the limit order book is resilient (see Section 3.3), the bid-ask spread mid-point $p_{m,n} = \frac{a_{m,n} + b_{m,n}}{2}$ is close to $\hat{v}$ (the difference is on average of the order of $\varepsilon^{0.34}$, where $\varepsilon = \frac{r}{2\lambda^U + 2\lambda^I}$ is the granularity parameter), so it can be taken as a good proxy for the efficient price. Moreover, because of the presence of impatient informed traders, every time the fundamental value is not in the bid-ask spread, they would place a market order. This brings the efficient price $\hat{v}$ eventually in the confines of the bid-ask spread.

The volatility story is more difficult. The econometrician does not observe $\hat{\sigma}$ (although the market is assumed to observe it). What the econometrician observes is the variance of the efficient price $\mathsf{Var}(\hat{v}) = \mathsf{Var}(E(v|OF))$, where OF represents the order flow. In particular, we can take the order flow to be represented only by sell market orders, since the market does not adjust as strongly after a sell limit order, and the buy market order side is symmetric.

In the rest of this section, one ignores the impatient informed traders, in order to get simpler formulas. By conditional probability theory, the observed volatility then becomes $\mathsf{Var}(\hat{v}) = \mathsf{Var}(E(v|SMO)) = \hat{\sigma}^2 - E(\mathsf{Var}(v|SMO))$. The conditional variance $\mathsf{Var}(v|SMO)$

can be computed as in the previous Proposition:

$$\mathsf{Var}(v|SMO) \;=\; \hat\sigma^2 - \hat\sigma^2 \, \frac{e^{-\frac{\alpha_{m,n}^2}{2\hat\sigma^2}}}{\Phi\left(-\frac{\alpha_{m,n}}{\hat\sigma}\right) + \frac{\lambda^U}{\lambda^I}} \left(\frac{e^{-\frac{\alpha_{m,n}^2}{2\hat\sigma^2}}}{\Phi\left(-\frac{\alpha_{m,n}}{\hat\sigma}\right) + \frac{\lambda^U}{\lambda^I}} - \frac{\alpha_{m,n}}{\sqrt{2\pi}} \right).$$

Finally, one gets the formula

$$\mathsf{Var}(\hat v) \;=\; \hat\sigma^2 \, \frac{e^{-\frac{\alpha_{m,n}^2}{2\hat\sigma^2}}}{\Phi\left(-\frac{\alpha_{m,n}}{\hat\sigma}\right) + \frac{\lambda^U}{\lambda^I}} \left(\frac{e^{-\frac{\alpha_{m,n}^2}{2\hat\sigma^2}}}{\Phi\left(-\frac{\alpha_{m,n}}{\hat\sigma}\right) + \frac{\lambda^U}{\lambda^I}} - \frac{\alpha_{m,n}}{\sqrt{2\pi}} \right). \tag{10}$$

So controlling for the observed variance of $\hat v$ (holding it fixed), a high asymmetric information ratio $\frac{\lambda^I}{\lambda^U}$ leads to a smaller volatility $\hat v$.

4.4 Empirical Predictions

This is a structural model of an order-driven market, so in principle one can test the various relationships among prices, spreads, trading activity, volatility, and information asymmetry by imposing this structure on the data.

But one can get directly some qualitative predictions from some comparative statics. For example, consistent with previous literature, it can be shown that smaller trading activity generate larger spreads: in section Section 3.3 it was shown that the average spread was of the order of $\varepsilon^{0.34}$, and by definition the granularity parameter $\varepsilon = r/(2\lambda^U + 2\lambda^I)$ is inversely related to trading activity $2\lambda^U + 2\lambda^I$. spreads.

In the model higher volatility also causes larger spreads. This is because the the market-inferred probability of an informed sell order is $\Phi(\frac{-g_{m,n}}{\hat\sigma})$, which increases with $\hat\sigma$. By the results in Roşu(2008), a smaller percentage of limit orders per market order makes the limit order book less resilient, and hence causes larger bid-ask spreads. The intuition is that higher volatility makes extreme fundamental values more likely, and increases the percentage of market orders (from patient informed traders), which makes the spreads wider.

A surprising new prediction is that, controlling for volatility and trading activity, a higher percentage of informed traders should generate *smaller* spreads. This is because a higher percentage of informed traders generates a quicker adjustment of prices to fundamentals, and

therefore extreme values of v become less likely. This generates fewer market orders from patient informed sellers, and hence smaller spreads.

Also, this model contributes towards an explanation of the "diagonal effect" of Biais, Hillion and Spatt (1995), namely that the order flow is positively autocorrelated (e.g. a market buy order makes a future market buy order more likely). In this model one can show that the order flow is autocorrelated if and only if there exist informed traders, with a higher autocorrelation when there is a larger percentage of informed traders, or a larger volatility. To see this, consider the case with only liquidity traders. In that case, we saw that the limit order book is resilient and tends to revert the the mid-point $\hat{v}$. This means that the ex-ante entry probability is approximately equal for all types of traders. So since the market arrivals are independent and the entry decision is made with approximately the same probability, the order flow is approximately uncorrelated (i.e. a buy market order does not make future buy market orders more likely). By contrast, in the presence of informed traders, a market order reflects a fundamental price v far away from the efficient price $\hat{v}$, and makes the (ex-post) probability of another market order relatively higher. In conclusion, the present model produces a diagonal effect as in Biais, Hillion and Spatt (1995), and makes the extra prediction that the diagonal effect should be stronger when volatility or asymmetric information are higher.

Can one use this paper to estimate the extent of information asymmetry in the market (e.g. by using the level of autocorrelation of order flow)? The answer is, it depends on what one means by information asymmetry. What if for example some agents do not observe the fundamental value v, but instead have information about the future order flow (they are informed in the sense of Evans and Lyons, 2002)? Then these agents would would behave in the same way as the informed agents in the current model, and become impatient if the information about the future order flow is extreme enough. Could we then distinguish these agents from truly informed ones? Perhaps, if one looked at price reversals. If the order-flow-informed traders do not really have information about v, then the efficient price $\hat{v}$ would then revert to the true value v after the truly informed traders have brought prices in line with fundamentals.

Information about order flow is not the only alternative explanation of the diagonal effect.

Another explanation comes from the possibility of large orders (if one relaxes the assumption that agents can only trade one unit of the asset). Then an agent who wants to trade a large quantity and is patient enough to work the order (divide it into smaller orders) and thus take advantage of the resilience of the limit order book would also generate a positively autocorrelated order flow. Can one test if this is the right explanation? If one had access to identity of the order flow, then one could check if there was just one trader working the order. If so, then the model could still incorporate working orders, by treating each order as coming from a separate trader. In that case, one finds again the problem of distinguishing information about fundamentals from information about order flow.

Notice that this model in the end generates two types of resilience. One is the resilience relative to the limit order book (we can call it "micro resilience"), and is due to the action of discriminatory liquidity traders. The other makes the efficient price $\hat{v}$ eventually converge to the fundamental value (we can call it "macro resilience"), and is due to the action of informed traders.

5 Conclusions

This paper has proposed a tractable models of an order-driven market with both liquidity traders and informed traders. The resulting equilibrium is quite intuitive. Informed patient traders generally submit limit orders, except when their privately observed fundamental value of the asset deviates a lot from the current market-inferred value, the efficient price. In that case, they become impatient and submit a market order. As a result, e.g. market buy orders are interpreted as strong positive signals; by contrast, a limit buy order are weaker signals, and in some cases can be showed to even provide a negative signal.

The model derives various testable implications, relating prices, spreads, trading activity, and volatility. In particular, the order flow is autocorrelated only in the presence of informed traders, and the order flow autocorrelation is higher whenever there is larger percentage of informed traders, or a larger volatility. Higher volatility and smaller trading activity are showed to generate larger spreads, while in contradiction with some of the asymmetric information literature, a higher percentage of informed traders actually generates smaller

spreads.

Some of the limitations of this model point towards future directions for research. In particular, the model assumes that everybody monitors the market at all times and makes the right inferences. It would be useful to see how the results of this paper would change in the presence of monitoring costs. This would also create the risk of limit orders being picked off, which does not occur in this model above and beyond information asymmetry.

Also, the model assumes that the private cost (or valuation) of the liquidity traders does not change. Similarly, informed traders do not use their information past the entry decision. While relaxing these assumptions arguably does not change the qualitative predictions of the model, it would probably lead to a richer model, which could also generate order cancellations.

A   Proofs of Results

PROOF OF PROPOSITION 1: Recall the recursive equation (5) for f_m:

$$f_m = \frac{1}{2}\left(\frac{1 - f_{m-1}}{2}f_{m-1} + \frac{1 + f_{m-1}}{2}f_m\right) + \frac{1}{2}\left(\frac{1 + f_{m+1}}{2}f_{m+1} + \frac{1 - f_{m+1}}{2}f_m\right) - \frac{\varepsilon}{2}.$$

When $m = 1$ one gets: $f_1 = \frac{1}{2}\left(\frac{1-a_1}{2}a_1 + \frac{1+a_1}{2}f_1\right) + \frac{1}{2}\left(\frac{1+f_2}{2}f_2 + \frac{1-f_2}{2}f_1\right) - \frac{\varepsilon}{2}$, which implies

$$f_1 = \frac{f_2(1 + f_2) + a_1(1 - a_1)}{2 + f_2 - a_1}.$$

Since the seller is a monopolist, he can choose an ask price a_1 that maximizes f_1 above. The solution is

$$a_1^* = 2 + f_2 - \sqrt{2(1 + f_2)}.$$

Substituting a_1^* in the formula for f_1, one gets $f_1 = 2a_1^* - 1 = 3 + 2f_2 - 2\sqrt{2(1 + f_2)}$, and so

$$a_1^* = \frac{1 + f_1}{2}.$$

One can check that indeed $f_2 < f_1 < f_0 = a_1 < 1$.   □

23

PROOF OF NUMERICAL PROPOSITION 1: The one-sided market with patient sellers and impatient buyers is a Markov system with transition matrix P, where P_i^j is the probability of transition from state i to state j, with $i, j = 0, 1, \ldots, M-1, M$. Then P can be written as

$$
P = \begin{bmatrix}
P_0^0 & P_1^0 & 0 & 0 & \cdots & 0 & 0 \\
P_0^1 & 0 & P_2^1 & 0 & \cdots & 0 & 0 \\
0 & P_1^2 & 0 & P_3^2 & \cdots & 0 & 0 \\
0 & 0 & P_2^3 & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \cdots & 0 & P_M^{M-1} \\
0 & 0 & 0 & 0 & \cdots & P_{M-1}^M & P_M^M
\end{bmatrix}.
\tag{11}
$$

To determine P_i^j, use the recursive equation (5) for f_m:

$$
f_m + \frac{2\varepsilon}{2 + f_{m+1} - f_m} = f_{m+1} \frac{1 + f_{m+1}}{2 + f_{m+1} - f_m} + f_{m-1} \frac{1 - f_{m-1}}{2 + f_{m+1} - f_m}.
$$

This means that for $m = 1, 2, \ldots, M-1$,

$$
P_{m+1}^m = \frac{1 + f_{m+1}}{2 + f_{m+1} - f_{m-1}}, \qquad P_{m-1}^m = \frac{1 - f_{m-1}}{2 + f_{m+1} - f_{m-1}}.
$$

For $m = 1$, $P_1^0 = 1$ and $P_0^0 = 0$. For $m = M$, $P_M^M = \frac{1}{2 - f_{M-1}}$ and $P_{M-1}^M = \frac{1 - f_{M-1}}{2 - f_{M-1}}$.

To calculate the distribution of the bid-ask spread or price impact, one needs to know the stationary probability that the system is in state m. Denote this by x_m. Consider the row vector X with entries x_m. From the theory of Markov matrices, one knows that $XP = X$.

Solving for X, one gets

$$x_0 = x_0 P_0^0 + x_1 P_0^1,$$

$$x_1 = x_0 P_1^0 + x_2 P_1^2,$$

$$\vdots$$

$$x_{M-1} = x_{M-2} P_{M-1}^{M-2} + x_M P_{M-1}^M,$$

$$x_M = x_{M-1} P_M^{M-1} + x_M P_M^M.$$

Starting with an arbitrary value of $x_0 > 0$, this generates a recursive system of equations in x_m. The resulting vector is then normalized so that the components sum to one. These are the stationary probabilities x_m. One can numerically check that they increase with state m, which is to be expected since the limit order book is resilient, so it is more likely that the system is in a state m with higher m. $\square$

For reference, one can write down the recursive equations that the expected utility $f_{m,n}$ satisfies in the canonical equilibrium in state (m,n) with m limit sellers and n limit buyers (recall that λ^U is the arrival rate of uniformed traders, λ^I is the arrival rate of informed traders, and P_j^i is the market-estimated probability that an informed trader of type $i \in \{PI, II\}$ (patient informed, impatient informed) makes the decision $j \in \{SMO, SLO, BLO, BMO, NO\}$ (sell market order, sell limit order, buy limit order, buy market order, or no order)):

$$
\begin{aligned}
f_{m,n} =\ & \frac{\lambda^U}{4\lambda^U + 2\lambda^I}\left(\frac{1 - a_{m,n}}{2} f_{m-1,n} + \frac{1 + a_{m,n}}{2} f_{m,n}\right) \\
& + \frac{\lambda^U}{4\lambda^U + 2\lambda^I}\left(\frac{1 + f_{m+1,n}}{2} f_{m+1,n} + \frac{1 - f_{m+1,n}}{2} f_{m,n}\right) \\
& + \frac{\lambda^U}{4\lambda^U + 2\lambda^I}\left(\frac{1 + b_{m,n}}{2} f_{m,n-1} + \frac{1 - b_{m,n}}{2} f_{m,n}\right) \\
& + \frac{\lambda^U}{4\lambda^U + 2\lambda^I}\left(\frac{1 - f_{m,n+1}}{2} f_{m,n+1} + \frac{1 + f_{m,n+1}}{2} f_{m,n}\right) \\
& + \frac{\lambda^I}{4\lambda^U + 2\lambda^I}\left(P_{BMO}^{PI} f_{m-1,n} + P_{SLO}^{PI} f_{m+1,n} + P_{SMO}^{PI} f_{m,n-1} + P_{BLO}^{PI} f_{m,n+1} + P_{NO}^{PI} f_{m,n}\right) \\
& + \frac{\lambda^I}{4\lambda^U + 2\lambda^I}\left(P_{BMO}^{II} f_{m-1,n} + P_{SLO}^{II} f_{m+1,n} + P_{SMO}^{II} f_{m,n-1} + P_{BLO}^{II} f_{m,n+1} + P_{NO}^{II} f_{m,n}\right) \\
& - \frac{r}{4\lambda^U + 2\lambda^I}.
\end{aligned}
\tag{12}
$$

PROOF OF PROPOSITION 3: The only difficult point is to show that the utility of an informed patient trader from submitting a sell limit order (SLO) is

$$u^{PI}_{SLO} = \hat{v} + f_{m+1,n} + k(v - \hat{v}), \qquad \text{with} \quad k = \frac{3\lambda^U + 4\lambda^I}{4\lambda^U + 4\lambda^I}$$

if v is sufficiently low. This follows from the recursive formula (12), except that the informed trader has better information about what the future informed traders are likely to do. Assume also that all the probabilities such as $\frac{1-a_{m,n}}{2}$ are close to $\frac{1}{2}$, since in Section 3.3 one saw that the bid and ask prices are small with very large probability.

Given that v is assumed sufficiently low, it is quite likely that the next informed trader submits a sell market order. This implies that the expected utility from waiting with a sell limit order, $f^*_{m,n}$, satisfies

$$(2\lambda^U + 2\lambda^I)f^*_{m,n} = \lambda^U \tfrac{1}{2}\big(f^*_{m-1,n} + f^*_{m+1,n} + f^*_{m,n-1} + f^*_{m,n+1}\big) + \lambda^I\big(f^*_{m,n-1} + f^*_{m,n-1}\big) - r.$$

Here we use f^* instead of f because the informed trader has superior information about what will happen in the limit order book. In particular, since prices converge on average to the fundamental value v, the informed traders expects all the $f^*_{i,j}$ to lose value by exactly $\hat{v} - v$. The only exception is in state $(m-1, n)$, since the patient trader has a limit order executed at the ask price of $a_{m,n} = f_{m-1,n}$. It follows that $f^*_{m,n} = f_{m,n} - \frac{3\lambda^U + 4\lambda^I}{4\lambda^U + 4\lambda^I}$. The analysis should be really done at $(m+1, n)$ instead of (m, n), so the desired formula is obtained. $\square$

PROOF OF COROLLARY 1: To prove the existence of the lower cutoff $\frac{-s_{m,n}}{k}$, compare the utility of patient informed trader (PI) from a sell market order (SMO) with the utility from a sell limit order (SLO), as in Proposition 3. The difference is given by

$$u^{PI}_{SMO} - u^{PI}_{SLO} = g_{m,n-1} - f_{m+1,n} - k(v - \hat{v}).$$

This is positive if and only if $v < \hat{v} + \frac{f_{m+1,n} - g_{m,n-1}}{k}$. As we saw in Section 3.3, one can approximate $f_{m+1,n}$ with $a_{m,n} = f_{m-1,n}$, while $g_{m,n-1} = b_{m,n}$. So SMO is preferred to SLO if

and only if

$$v \; < \; \hat{v} + \frac{s_{m,n}}{k}.$$

To show that existence of the second cutoff, one must compare a *SLO* with a *BLO*. The profit from *SLO* is given by $\big(\hat{v} + f_{m+1,n} + k(v - \hat{v})\big) - v$, while the profit from *BLO* is given by $v - \big(\hat{v} + g_{m,n+1} + k(v - \hat{v})\big)$. Having *SLO* preferred to *BLO* is then equivalent to $v \; < \; \hat{v} + \frac{1}{2(1-k)}(f_{m+1,n} + g_{m,n+1})$. Since the mid-point $p_{m,n}$ is approximately equal to $\frac{f_{m+1,n}+g_{m,n+1}}{2}$, one obtains that *SLO* is preferred to *BLO* if and only if

$$v \; < \; \hat{v} + \frac{p_{m,n}}{1 - k}.$$

$\square$

References

[1] ADMATI, A., AND P. PFLEIDERER (1988): "A Theory of Intraday Patterns: Volume and Price Variability," *Review of Financial Studies*, 1, 3–40.

[2] BIAIS, B., P. HILLION, AND C. SPATT (1995): "An Empirical Analysis of the Limit Order Book and the Order Flow in the Paris Bourse," *Journal of Finance*, 50, 1655–1689.

[3] BIAIS, B., D. MARTIMORT, AND J-C. ROCHET (2000): "Competing Mechanisms in a Common Value Environment," *Econometrica*, 68, 799–837.

[4] CHAKRAVARTY, S., AND C. HOLDEN (1995): "An Integrated Model of Market and Limit Orders," *Journal of Financial Intermediation*, 4, 213–241.

[5] COHEN, K., S. MAIER, R. SCHWARTZ, AND D. WHITCOMB (1981): "Transaction Costs, Order Placement Strategy, and Existence of the Bid-Ask Spread," *Journal of Political Economy*, 89, 287–305.

[6] DEMSETZ, H. (1968): "The Cost of Transacting," *Quarterly Journal of Economics*, 82, 33–53.

[7] EASLEY, D., AND M. OHARA (1987): "Price, Trade Size, and Information in Securities Markets," *Journal of Financial Economics*, 19, 60–90.

[8] EVANS, M., AND R. LYONS (2002): 'Order Flow and Exchange Rate Dynamics," *Journal of Political Economy*, 110, 170–180.

[9] FARMER, D., P. PATELLI, AND I. ZOVKO (2003): "The Predictive Power of Zero Intelligence in Financial Markets," Discussion Paper, September.

[10] FOUCAULT, T. (1999): "Order Flow Composition and Trading Costs in a Dynamic Limit Order Market," *Journal of Financial Markets*, 2, 99–134.

[11] FOUCAULT, T., O. KADAN, AND E. KANDEL (2005): "Limit Order Book as a Market for Liquidity," *Review of Financial Studies*, 18, 1171–1217.

[12] GLOSTEN, L. (1994): "Is the Electronic Open Limit Order Book Inevitable?," *Journal of Finance*, 49, 1127–1161.

[13] GLOSTEN, L., AND P. MILGROM (1985): "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders," *Journal of Financial Economics*, 14, 71–100.

[14] GOETTLER, R., C. PARLOUR, AND U. RAJAN (2005): "Equilibrium in a Dynamic Limit Order Market," *Journal of Finance*, 60, 2149–2192.

[15] HANDA, P., AND R. SCHWARTZ (1996): "Limit Order Trading," *Journal of Finance*, 51, 1835–1861.

[16] HARRIS, L., AND J. HASBROUCK (1996): "Market versus Limit Orders: the Superdot Evidence on Order Submission Strategy," *Journal of and Quantitative Analysis*, 31, 213–231.

[17] HOLLIFIELD, B., R. MILLER, AND P. SANDAS (2004): "Empirical Analysis of Limit Order Markets," *Review of Economic Studies*, 71, 1027–1063.

[18] HOLLIFIELD, B., R. MILLER, P. SANDAS, AND J. SLIVE (2006): "Estimating the Gains from Trade in Limit Order Markets," working paper.

[19] HUANG, R., AND H. STOLL (1997): "The Components of the Bid-Ask Spread: A General Approach," *Review of Financial Studies*, 10, 995–1034.

[20] KYLE, ALBERT (1985): "Continuous Auctions and Insider Trading," *Econometrica*, 53, 1315–1335.

[21] O'HARA, M. (1995): *Market Microstructure Theory*, Blackwell.

[22] PARLOUR, C. (1998): "Price Dynamics in Limit Order Markets," *Review of Financial Studies*, 11, 789–816.

[23] PARLOUR, C., AND D. SEPPI (2003): "Liquidity-Based Competition for Order Flow," *Review of Financial Studies*, 16, 301–343.

[24] ROCK, K. (1996): "The Specialist's Order Book and Price Anomalies," Discussion Paper.

[25] ROŞU, I. (2008): "A Dynamic Model of the Limit Order Book," Working Paper, University of Chicago, February.

[26] SANDAS, P. (2001): "Adverse Selection and Competitive Market Making: Empirical Evidence from a Limit Order Market," *Review of Financial Studies*, 14, 705–734.

[27] SEPPI, D. (1997): "Liquidity Provision with Limit Orders and a Strategic Specialist," *Review of Financial Studies*, 10, 103–150.