The Rodney L. White Center for Financial Research

Liquidity Supply and Demand: Empirical Evidence from the Vancouver Stock Exchange

Burton Hollifield
Robert A. Miller
Patrik Sandas
Joshua Slive

08-01

The Wharton School
University of Pennsylvania

The Rodney L. White Center for Financial Research

The Wharton School
University of Pennsylvania
3254 Steinberg Hall-Dietrich Hall
3620 Locust Walk
Philadelphia, PA 19104-6367

(215) 898-7616
(215) 573-8084 Fax
http://finance.wharton.upenn.edu/~rlwctr

The Rodney L. White Center for Financial Research is one of the oldest financial research centers in the country. It was founded in 1969 through a grant from Oppenheimer & Company in honor of its late partner, Rodney L. White. The Center receives support from its endowment and from annual contributions from its Members.

The Center sponsors a wide range of financial research. It publishes a working paper series and a reprint series. It holds an annual seminar, which for the last several years has focused on household financial decision making.

The Members of the Center gain the opportunity to participate in innovative research to break new ground in the field of finance. Through their membership, they also gain access to the Wharton School's faculty and enjoy other special benefits.

Members of the Center
2000 – 2001

Directing Members

Ford Motor Company Fund
Geewax, Terker & Company
Morgan Stanley
The Nasdaq Stock Market, Inc.
The New York Stock Exchange, Inc.
Twin Capital Management, Inc.

Members

Aronson + Partners
Credit Suisse Asset Management
Exxon Mobil Corporation
Goldman, Sachs & Co.
Merck & Co., Inc.
The Nasdaq Stock Market Educational Foundation, Inc.
Spear, Leeds & Kellogg

Founding Members

Ford Motor Company Fund
Merrill Lynch, Pierce, Fenner & Smith, Inc.
Oppenheimer & Company
Philadelphia National Bank
Salomon Brothers
Weiss, Peck and Greer

Liquidity Supply and Demand: Empirical Evidence from the Vancouver Stock Exchange*

Burton Hollifield
GSIA
Carnegie Mellon University

Robert A. Miller
GSIA
Carnegie Mellon University

Patrik Sandås
The Wharton School
University of Pennsylvania
and
CEPR

Joshua Slive
Faculty of Commerce
University of British Columbia

First Draft: November, 2000
Current Draft: March, 2001

*We would like to thank the Carnegie Bosch Institute at Carnegie Mellon University, the Social Science and Humanities Research Council of Canada and the Rodney L. White Center for Financial Research at Wharton for providing financial support and the Vancouver Stock Exchange for providing the data set. Discussions with Pierre Collin–Dufresne have been very helpful to us. All errors are ours.

Abstract

We analyze the costs and benefits of providing and using liquidity in a limit order market. Using a large and comprehensive data set which details the complete histories of orders and trades on the Vancouver Stock Exchange, we are able to model the order flow and measure market liquidity as it changes over time. We accomplish this by constructing a measure of the expected net payoffs to demanding or supplying liquidity, and using our data on order arrivals and placement decisions to make inferences about the traders' demand for liquidity and the cost of entering orders in the market. Our results show that liquidity demand is indeed time varying, and is related to several key observable measures of market characteristics. Furthermore, we find evidence of unexploited profit opportunities in the market, perhaps implying that traders do not continuously monitor the market for profitable trades.

Keywords: Liquidity, Limit Order Markets, High Frequency Data, Market Microstructure

Introduction

Successful financial markets offer a liquid marketplace for investors to trade. Black (1971) defines a market as liquid if any quantity of the security can be bought or sold immediately, if small amounts can be bought or sold immediately at prices close to the current market price, and if larger quantities can be transacted over time at prices that are on average close to the current market price. In different market architectures, liquidity is provided by different market participants. Liquidity may be supplied by a specialist, floor brokers, competing dealers, traders placing limit orders or various combinations of the above. In limit order markets, much of the liquidity is provided by traders placing limit orders into a centralized order book, and liquidity is used up by traders placing market orders. A liquid limit order market has a large volume of limit orders in the book, a small bid/ask spread, a relatively large quantity of shares offered close to the bid and ask quotes, and a limit order book which rebounds quickly after a market order is submitted.

In this paper we use a large data set consisting of the order submissions and trades from a particular limit order market, the Vancouver Stock Exchange. We use our detailed data set to construct empirical measures of the profitability of supplying liquidity and combine these estimates with the order flow to estimate the demand for liquidity in this market. We use our estimates to study how variations in the arrival rates of orders and trades are driven by changes in the relative profitability of supplying or demanding liquidity, and to measure the costs of providing liquidity in this market. We find empirical evidence that there are unexploited profit opportunities to submitting limit orders in our data. We also find evidence that the distribution of liquidity demand is state–dependent in our data, so that relative changes in the profitability of providing and consuming liquidity cannot explain all the variation in the timing and composition of the order flow in this market.

Our empirical approach is as follows. In our data, we observe the timing of order submissions, the market or limit orders chosen by the traders, and the timing of trades or cancellations of limit and market orders. We use this data to estimate the conditional probability that an order executes within a fixed amount of time, and we form estimates of the value of the asset, conditional upon a trade occurring. We use these estimates to compute the expected profits from submitting various orders. By assuming that traders submit orders to maximize their expected profits from the order submissions and pay a fixed order entry cost, we can use the order submissions and the time between order submissions to make inferences on the traders' demand for immediacy, the fixed costs of submitting orders and the underlying arrival rates of traders. In our model, traders with a low demand for immediacy typically submit limit orders and so supply liquidity to the market. Traders with more extreme demands for immediacy typically submit market orders and so demand liquidity. We use empirical variations in the relative profitability of limit orders and market orders along with variations in the relative proportion of order choices and the timing of order choices to make inferences on traders' demand for immediacy, the cost of submitting orders and the arrival rates of traders.

There is previous empirical evidence documenting that liquidity providers react to market conditions. Biais, Hillion, and Spatt (1995), provide empirical evidence from the Paris Bourse that the limit order book often has no limit orders directly around the best bid quotes. However, traders on the Bourse react to a large spread by submitting limit orders in rapid succession. Using data from the NYSE, Goldstein and Kavajecz (2000) document dramatic shifts in traders willingness to place limit orders in the order book during extreme market movements. Harris and Hasbrouck (1996) show that the profitability of limit and market orders vary with different market conditions on the NYSE. Foster and Viswanathan (1993) find interday and intraday variation in the costs of transacting on the NYSE. Sandås (2000) shows that the depth in the limit order book is time–varying in the Stockholm Stock Exchange. Hollifield, Miller and Sandås (1999) show that changes in the relative profitability of limit and market orders are important for explaining the empirical variation in order submission rates in the Stockholm Stock Exchange. Coppejans, Domowitz and Madhavan (2000) show that depth in the limit order book for the OMX futures contract changes through time, implying that liquidity provision changes dynamically. In our work, we provide further evidence of variation in liquidity supply and demand in a different limit order market, the Vancouver Stock Exchange. In addition, our empirical approach directly models how order profitability drives order submissions, and thus liquidity in the market.

Our empirical analysis builds upon the familiar trade–off facing traders in limit order markets. Limit orders offer a better price than market orders, but face execution and adverse selection risk. Execution risk captures the probability that an order submitted away from the best quote may fail to execute. Adverse selection risk, or the risk of being picked off, captures the likelihood that a limit order will execute in the future when the underlying value of the asset has moved against the trader. Our approach is an effort to determine how competitive the market for liquidity is without making strong a priori assumptions. In a perfect competition setting the limit order book is characterized by zero-expected profits as shown in Glosten (1994), or Seppi (1997). These models do not directly address timing of orders since it focuses on the equilibrium in a single period game. However, if we were to make predictions about timing based on a model such as in Glosten (1994), we would expect there to be a strong negative relationship between the time between orders and the profitability of being in the order book. Competition between the passive liquidity suppliers ensures that new limit orders are submitted in rapid succession whenever profit opportunities arises, for example, after a trade or an innovation in the asset value. In contrast, Foucault (1999) analyzes a dynamic model of a limit order market *without* any profit-seeking liquidity suppliers, i.e., all traders are actively interested in trading. In these models, the arrival rate of orders is independent of the state of the order book. Only the composition of the order flow itself depends on the state of the order book. Thus, with respect to the timing of orders these models provide two extreme predictions for the relationship between order arrival rates and the profitability of submitting limit orders.[1] Our strategy is to use the data to make inferences on the expected profits for different order strategies and to use the observed arrival rates to determine what best describes the market outcomes. Thus, we neither assume that a zero-expected profit condition holds nor that the order flow is independent of the state of the order book.[2]

Our empirical analysis uses the variation in the time between order submissions and trades to make inferences on the profitability and costs of order submissions. Easley, Kiefer and O'Hara (1997) relate the timing of trades to the properties of the spread using data from the NYSE. There is a recent and growing literature that exploits high–frequency financial data to study the timing of orders and trades using a variety of statistical techniques. Examples include Engle and Russell (1998) and Engle (2000), who develop and estimate autoregressive conditional duration models, Engle and Lange (1997) who estimate a time varying market depth measure, Coppejans and

[1]Cohen, Maier, Schwartz and Whitcomb (1981) and Parlour (1997) also analyze the implications of the trade–off between market and limit orders, in settings where the marginal limit order does not earn zero expected profits.

[2]Glosten (2000) provides examples that illustrate how under plausible assumptions that equilibrium the order book does not need to satisfy a zero-expected profit condition.

Domowitz (1999) who estimate different conditional duration models in limit order data, Hasbrouck (1999) who uses multivariate point processes to study time deformation, and Lo, MacKinlay, and Zhang (2000) who estimate duration models for the timing of limit order executions. We add to these studies by relating the variation in timing to the variation in estimates of the profitability of supplying and demanding liquidity.

Our empirical technique allows us to ask how potential liquidity suppliers respond to changes in market conditions: When are they likely to enter and supply liquidity? Given that we model jointly the demand for liquidity and the profitability of supplying liquidity we can also answer questions about liquidity demand. For example, are periods of high trading volume associated with more frequent arrival of all types of traders? Is an increase in the order flow driven by an increase in the profitability of submitting orders? Is higher volume associated with a shift in the distribution of liquidity demand toward traders with higher demand for immediacy? An increase in traders demanding immediacy will increase the quantity of market orders submitted, leading to an increase in the profitability of submitting limit orders, and thus increasing the supply of liquidity. There are a number of different theoretical models that feature this type of feedback effect in different environments. Examples include Pagano (1989), who provides a model of endogenous liquidity where more liquid markets attract more volume. Admati and Pfleiderer (1988) study a model where discretionary liquidity traders cluster intertemporally, and Dow (1999) analyzes a model with multiple equilibria, each featuring a different level of liquidity.

The plan of the paper is as follows. In the next section, we discuss some of the basic institutional features of the market we are studying, the Vancouver Stock Exchange, and some stylized facts for our data set. We then present the theoretical framework, and a discussion of the econometric techniques we apply to our dataset. Next, we provide our empirical estimates and discuss our results. The final section concludes. All proofs are contained in a technical appendix.

1 Description of the Market and the Data

1.1 The Market Institution

In 1989, the Vancouver Stock Exchange became the first North American equity exchange to convert to completely computerized trading using an electronic limit order book referred to as the Vancouver Computerized Trading system (VCT).[3] This trading system is similar to limit order market systems used on other exchanges, for example, the Paris Bourse and the Toronto Stock Exchange. Each trading day starts with a pre–opening period which is followed by an opening auction at 6:30 a.m. Pacific Time. Trading is continuous after the opening auction until 1:30 p.m. Our discussion of the institutional setting is focused on the continuous trading period, and on the rules that were in effect during our sample period of 1989 through 1993.

In this type of trading system there are no floor traders, market makers or specialists with special quoting obligations or trading privileges. The most important source of liquidity is the limit orders in the order book.[4] There are two groups of market participants who provide liquidity—investors and brokers. Investors place orders in the order book via brokers. Brokers at forty–five exchange member firms are directly connected to the electronic trading system via VCT terminals. The brokers can trade on their own account as well on behalf of outside investors. Thus, even if a broker is not obligated to make a market in a given stock, he may choose to do so. This choice is likely to be determined by the profitability of supplying liquidity in different market conditions or in different stocks.

All limit orders in the order book face the same priority rules which determine how orders are matched to produce trades. Thus, in this respect the two groups of potential liquidity suppliers, investors and brokers, compete on a level playing field. A limit order or a market order which is only partially filled, is assigned an entry time and placed in the order book. In the order book, all orders are prioritized first by price and second by the time of submission. A limit order may be "good 'til canceled" or it may be a so called "day order" that expires at the end of the trading day. Irrespective of the expiration condition, an order can be canceled at any time. All order prices must

[3] In 1999, after the end of our data sample, the Vancouver Stock Exchange was involved in an amalgamation of Canadian equity trading and became a part of the Canadian Venture Exchange (CDNX). The new exchange uses the same trading system as the Vancouver Stock Exchange.

[4] Some of the stocks may be cross–listed and off–exchange trading, subject to some rules described below, is also an alternative to the VCT system.

be a multiple of a fixed minimum price unit known as the *tick size.* The tick size varies between one cent for prices below $3.00, five cents for prices between $3.00 and $4.99, and twelve and a half cents for prices at $5.00 and above. Orders submitted to the main order book known as the *board lot book* must also be for a size that is a multiple of a fixed round or board lot size. The board lot size varies between 100 and 1000 shares depending on the price level.

The VSE allows traders to submit *partially disclosed* or *hidden* orders where part of the volume is not visible on the book. The trader must disclose at least 1000 shares or 50% of the true volume of the order, whichever is less. The undisclosed part of the volume retains price priority, but not time priority. Once the disclosed volume is used up, an amount equal to the initially disclosed volume is automatically disclosed and made available for trades. Traders can also submit *on–stop orders.* These orders have a limit price that is better than the current best quotes. Thus, if the order was submitted as a regular limit order, it would trade immediately. However, the order is not submitted to the the board lot book until a trade occurs at or better than the on–stop price. In our data set, both hidden and on–stop order show up when their respective submission conditions are triggered.

The VSE, like the Paris Bourse, handles "market orders" in a way that differs from, say, the practice at the NYSE.[5] The order described as a *market order* in the VCT system is actually a limit order that automatically is assigned a price that ensures that it is matched with the best order on the other side of the market. There is a so called *one tick* rule which limits how far the market order can "walk up the book." This rule states that the market order will trade against all orders at the best limit price, and, if volume remains, will also trade against any limit order with a price no more than *one tick* from the first trade price. Any remaining volume becomes a limit order at a price one tick from the first trade price. Hence, a large *market order* is not necessarily guaranteed full and immediate execution at the VSE.

One rationale for using the market order is that it allows a trader who does not monitor the market continuously to place an order that will trade only within one tick of the best available price irrespective of the state of the book. This type of order will also expose, and possible be matched with, any hidden order quantities at the best quote or the one tick away price level. On

[5] See Biais, Hillion and Spatt (1995), pp. 1660–1661, for a discussion of how market orders are handled on the Paris Bourse.

the other hand, it is clear that an active trader can accomplish the exact same payoffs by using a limit order with the appropriate limit price. Furthermore, a trader who demands immediacy can use an appropriately priced marketable limit order and walk up the book. Although traders with different costs of monitoring the book may have different preferences for marketable limit orders versus market orders, it is impossible to distinguish the payoffs from these two strategies in our data. Henceforth, we will follow standard terminology and refer to orders that express the highest demand for immediacy as *market orders* and those orders that do not result in immediate execution as *limit orders.*

The VSE, like most limit order markets, offers traders a large amount of real time information about the trading process. The entire book is visible to all brokers via their VCT terminals. This is a *market–by–order* system where each individual order is visible. This contrasts with *market–by–price* systems where the traders see only the total quantity available at each price. The information is updated almost in real time as new orders enter the system and trades take place. A burst in trading activity, for example, shortly after the open or before the close, may create short delays as the computer system processes the order flow.

All brokers with a VCT terminal can observe a code that identifies the member firm that submitted a given order on the book. Investors who are not members of the exchange have access to book information via commercial vendors who sell an information feed from the exchange. The external information feed includes the five best bid and ask levels with the corresponding total order quantities, but not the identification codes of the brokers. In addition, it includes the individual order quantities of the ten best orders on each side of the market. Both types of market participants, brokers and outside investors, have access to very detailed information about the contents of the order book and the trading process. This makes it easier for a market participant who monitors the market actively to detect profitable order placement opportunities.

Most of the order submission and trading activity goes through the main order book, the board lot book. However, some trading activity takes place outside the board lot book. Some of the stocks are cross–listed on other exchanges. There is also a *special terms book* for orders which have special restrictions such as delayed delivery and minimum fills, and an odd–lot book for small orders. Exchange members can also cross orders internally by submitting a *cross order* to the market. A cross order can be entered with a price that is equal to the current bid or ask quote or lies between the best quotes. If the price matches the best bid or ask price, then only 50% of the order will be crossed, and the remainder will be taken up by the book. However, if the stock is cross–listed or the value of the transaction exceeds $100,000, then 100% of the order will be crossed. This rule allows members to partially circumvent time priority when they have orders on both sides of the market.

1.2 The Data

Our data set was obtained directly from the Vancouver Stock Exchange in the form of *audit tapes*. This is essentially the raw data that is output by the VCT system and used by the exchange as the historical record of the trading process. It contains all stocks on the VSE and all transaction over a period starting in 1989 and ending in 1993. In this paper, we focus on only one stock, Eurus Resource Inc., for a three month period that includes 63 trading days, starting in May 28, 1990. Eurus Resource is a mining company specializing in precious metals.

The data is extremely rich and it allows us to reconstruct complete histories of *individual* orders submitted to the market using the following procedure. Each "event" such as a trade, cancellation or change in the status of the order generates a record in the data. The record includes an original order entry time that allows us to pinpoint when an order was initially submitted. This information, combined with data on the contents of the opening order book, allows us to identify the timing of all changes in the order book. We use an iterative procedure to reconstruct the trades' order placement decisions and some of the information available to them at the time of the decision. First, we identify the times of all changes. Second, we reconstruct the changes in the limit order book. Third, we extract the individual order histories, including initial submission and every future order fill or cancellation, and the corresponding order books. For a small fraction of the orders—less than one percent—we observe inconsistencies between the inferred order histories and the rules of the exchange. We drop these orders from our analysis.

Our data set is extremely detailed, but there are some important limitations. First, we cannot separate the trades that a broker makes on his own behalf from those he makes on behalf of his customers. Second, since the investors in our data set are anonymous, we cannot link separate orders submitted by the same investors, for example, as part of a dynamic order placement strategy. Third, traders with access to the VCT screens observe the identify of a member firm that submits an order. We do not have this information. The first two limitations cause us to focus on how a trader with an *unobserved* reservation price solves a *one–shot* order placement problem. We use our detailed data set to construct proxies for the traders' information sets. Of course, as econometricians, we seldom observe agents' complete information sets and the identification codes is only one example of a piece of information that we lack. Nevertheless, our data set, like other data sets from limit order markets, provides a very detailed characterization of the trading environment and the agents' decisions.

Our overall sample contains 10,608 order submissions over a period of 63 days, starting in May 28, 1990. There were 3 hidden orders submitted during the sample period. Approximately 4.4% of the submitted orders do not appear on the board lot book. These orders include odd–lots, special terms and on–stop orders. We drop these orders from any further empirical analysis.

Table 1 presents some descriptive statistics on the daily level of trading activity and the order flow for our sample stock. For each daily variable we report the mean, the standard deviation and the 5th, 25th, 50th, 75th and 95th percentiles of the distribution. The average closing mid–quote is $2.52 and for most of the sample period the stock is traded with a minimum tick size of one cent. The open–to–close and close–to–open returns show that the stock is relatively volatile both during the trading day and overnight. Traders submitting limit orders that remain outstanding for some period of time are therefore likely to face a substantial risk of being picked off due to changes in the value of the stock. The average daily trading volume is roughly 118,000 shares or $313,000. The daily average bid–ask spread is computed as a time series average of the average daily bid–ask spreads. The spread is on average 4.4 cents and half of the observations are between 3.6 and 5.2 cents. The average available depth is measured as the time series average of the average daily number of shares at the best bid and ask quotes. On average approximately 6,800 shares are offered at the best quotes, but the standard deviation of 4,540 suggests that even the average daily depth varies substantially over time.

Panel B of Table 1 presents summary statistics on the different types of orders that traders submit. There is a total of 10,608 orders in our current sample. On average we observe about 68 market orders, split between 53 fully and 15 partially filled market orders, per day. On average there are 101 limit order submissions during a trading day. Most orders last for only one day and of the orders that last for more than a day, only a small fraction stay in the system for more than 48 hours. On average there are two orders per day that stay for more than two trading days. Of course, this reflects to some extent the fact that many orders were simply filled soon after they were submitted or canceled. For our purposes, the important fact is that most of the uncertainty about the payoffs of a given order is resolved within two trading days.

Table 2 presents a more detailed breakdown of the characteristics of different types of orders. The orders are divided into 21 buy and sell order categories according to the aggressiveness of the order price. All market, marketable limit orders and partially filled market orders are aggregated into a market buy and a market sell order category. Buy and sell limit orders within 19 cents of the prevailing ask or bid quote at the time of order submission are reported on separate rows for each one cent price increment. Finally, orders placed at 20 cents or more from the best quotes are aggregated into one category for buy and sell orders, respectively. The first column reports the number of orders in each of the categories. There is a tendency for orders to cluster at 5 or 10 or 15 cents from the quotes, but traders occasionally use all prices. The next seven columns report statistics on the order size measured in thousand of shares. While there are differences across the different categories, it is striking how similar the order sizes are overall. For example, the median order size is 2,000 shares for all market orders and limit orders submitted within 10 cents of the best quotes. The inter–quartile range is also relatively narrow, for most order types it is about 2,000 shares.

The last four columns reports statistics on the overall performance of the orders. The first two columns report the average and the standard error for the fill–ratio of the order. The fill–ratio is calculated as the fraction of the original order that was filled within two trading days. We use a cut–off of two trading days since few orders last for more than two days. Accordingly we do not use new orders submitted during the two last trading days of the sample in our calculations. In general, the fill–ratio falls as we move away from the quotes, as we would expect. There are several non–monotonicities, but this is not surprising given that we do not have very precise estimates of many of the fill–ratios due to the small number of observations. In our estimation, we will aggregate these bins into a smaller number of choice categories for more precise estimates of the fill–ratios. The last two columns give the average time until an order is filled (in minutes) with the corresponding standard errors. Orders submitted further away from the quotes tend to be filled later, but there are large variations here as well. This is likely because the number of observation is small in some bins, and we are looking at unconditional averages. It is likely that limit orders further away from the best quotes are mainly submitted at times when the expected payoff from such a strategy is relatively high.

Table 3 reports descriptive statistics on the order book. These statistics are computed for the time series of observed order books. Each change in the order book generates one observation. The first half of the table reports statistics for the order quantities. On average, each of the five best bid or ask levels offer between 3,750 and 5,600 shares, which is equivalent to about one and a half to two times the average order sizes. Thus, in order to obtain a complete fill immediately, larger orders will typically need to go deeper into the order book than the best quotes. The distribution of order book quantities is skewed toward larger quantities with a median order quantity that is roughly 1,500 shares below the mean order quantity. The second half of the table reports the corresponding statistics for the order book prices levels relative to the mid–quote. On average the distance between adjacent levels in the order book is more than the tick size. Both the order book quantities and the prices vary substantially as is evident form the 5th and 95th percentiles of the distributions. This implies that the liquidity available in the order book varies quite dramatically through time. For example, a very tight order book would exhibit a difference between adjacent price levels equal to the minimum of one cent. On the other hand, the 95th percentile shows that occasionally these difference are equal to five cents. In our empirical analysis we are asking to what extent these types of variations in the observed liquidity can be explained by variations in the profitability of submitting limit orders. Is it the case that the order book offers less liquidity when the expected fill–ratio is smaller or the risk of being picked off is greater for a limit order?

Figures 1–4 offer more evidence on how the activity levels and the available liquidity, the spread and the depth, vary over our sample period. Figure 1 plots the daily time series of the total number of orders submitted (bars) and the daily closing price (solid line). Interestingly, most active trading days tend to be associated with large changes in the closing price. Our empirical techniques will allow us to ask whether such bursts in activity are due to an increase in the trader demand for immediacy, large shifts in the common value of the asset or some combination of the two. Figures 2 and 3 show the daily average bid–ask spread (in dollars) and the total depth at the best bid and ask quotes (shares). Figure 4 shows how the number of order submission of different types varies over the sample. The daily averages of the bid–ask spread and the depth are computed based on all distinct order book observations on a given day. These plots suggest that the available liquidity, at least at the best quotes, is time varying. Measures of trading activity such as the number of orders submitted is negatively correlated with the bid–ask spread and positively correlated with the available depth. We are interested in understanding whether some of the observed time variation in liquidity can be attributed to changes in the profitability of supplying liquidity.

Table 4 reports the results obtained by estimating a Cox proportional hazard rate model to the timing between order submissions[6]. The conditioning variables that we use are the dollar bid/ask spread, the order book quantity at the best bid, the order book quantity at the best ask, the order book quantity at bid levels 2 through 5, the order book quantity at ask levels 2 through 5, a dummy variable for time $\leq$ 8:00AM, a dummy variable for time $\geq$ 12:00 trading volume during last half hour and the standard deviation of transaction prices over last half hour. The point estimates and standard errors provide evidence that except for the afternoon dummy, all of these variables are useful at predicting variation in the time–between order submissions. Thus, our data provides evidence that there is variation in the durations between order submissions, and that this variation is related to the size of the limit order book, and to measure of lagged market activity. In the next section of the paper, we present a simple model of optimal order choice which relates traders demand for immediacy and measures of the profitability of submitting various order types to the conditional probability of observing different types of orders and the time–between order submissions.

2 Theoretical Framework

In this section, we describe the theoretical framework that we will apply in our dataset. We start with a description of our assumptions about the arrival rates and preferences of traders in our

[6] See Lancaster (1990) and Kiefer (1988) for a definition of the proportional hazard model.

model and then characterize the optimal order submission strategy in this environment. We then describe the implications of this model for the timing of order decisions and the probability of observing different order submission choices. Finally, we show how we estimate the parameters of the model and test various economic hypotheses on the arrival rates of traders and demands for liquidity using our data.

We assume that over the time interval $[t, t+\Delta t]$, the probability that one potential trader arrives in the market is given by $\lambda_t \Delta t$, where the subscript t represents information available at time t, and Δt is small. The conditional probability of entry, λ_t is interpreted as the intensity that traders will consider participating in the market. If this intensity is large, then there are many traders who could potentially submit orders in the market, and if this intensity is small, then there are few traders who can submit orders. For example, if there are many holders of the asset, then we would expect a higher value for λ_t, and if the arrival of traders is clustered in time, then this intensity will vary to reflect this clustering. Empirically, we will relate λ_t to information that predicts trading profits in the market.

Once a trader arrives in the market, he observes the electronic limit order book, which is a queue of unfilled limit orders. At this point, the trader can decide to enter an order into the electronic system. An order can either be a market order, or a limit order. Given that the trader observes the current limit order book when submitting his order, he can determine the price that results in immediate execution, and we will refer to such an order as a market order. We use the decision indicator variables d^s_{kt}, d^b_{lt} for $k = 0, 1, \ldots, K$ and $l = 0, 1, \ldots, L$. and quantities q_t to denote the trader's decision at time t. If $d^s_k = 1$, then the trader submits a sell limit order at the price k ticks above the current market order price, and q_t denotes the number of shares in the order. If the trader submits a market sell order order, then $d^s_{0t} = 1$. If the trader submits a market buy order, then $d^b_{0t} = 1$ and if the trader does not submit any order, then $d^s_{kt} = 0, d^b_{lt} = 0$ for all k and l. In order to reduce notational clutter, we will often drop the t subscript from these variables.

We now describe our assumptions on the traders' preferences in this model. All traders are assumed to be risk neutral, and maximize the expected utility of their order submission strategies. At time t, if trader j enters into the market, then we assume that his valuation for the asset is given by v_{jt}. Following Tauchen and Pitts (1983) and Hollifield, Miller and Sandås (1999), we decompose

v_{jt} into a common component and a private component:

$$v_{jt} = y_t + u_{jt}. \tag{1}$$

We interpret v_{jt} as the expected utility that the trader achieves from holding one unit of the asset. The random variable y_t represents the common value of the asset at time t, equal to the markets' expectation of the liquidation value of the security at time t. The common value is a stochastic process, and changes randomly over time as the market learns new information. However, since y_t is the conditional expectation of the securities liquidation value, it is a martingale relative to the market's information information set. That is,

$$y_t = E_t\left[y_{t+\Delta t}\right], \ \forall t, \text{ and } \Delta t > 0, \tag{2}$$

where the subscript t denotes conditioning on the market's information set at time t.

An important assumption in our model is that at time t, all traders agree on this value. However, since this value is stochastic, traders who enter into the market in the future will have an informational advantage relative to current traders regarding the common value of the asset. It is in this sense that limit orders are exposed to picking off risk in our model.

The private component of traders' valuation is drawn *i.i.d.* across traders from the continuous distribution

$$\text{Prob}\left(u_{jt} \leq u \middle| \text{Information at } t\right) = G_t\left(u\right), \tag{3}$$

with continuous density $g_t(\cdot)$. This distribution is conditional on information available at time t. We interpret u_{jt} as a measure of the traders' demand for immediacy or liquidity. Traders with extreme values of u_{jt} have a high desire to trade the asset immediately, and traders with private values close to zero have no particular reason to trade, unless the limit order book presents them with profits from doing so. In our model, the trading opportunities in the limit order book determines the profitability of various order submission strategies for traders with different valuations for the asset. If a trader submits an order to buy or sell the asset, then he must pay a fixed cost of C per share to submit the order. This cost is the same irrespective of the type of order submitted.

We now present the traders' payoff function and characterize the optimal order submission decision. Suppose that the trader with value $v = y + u$ submits a buy order of size q, l ticks away from the ask at price p_l. Here, we drop that t subscripts to reduce notational clutter. Define $\tilde{Q}_\tau$ as the cumulative quantity of the order executed in τ periods and let $\bar{\mathcal{T}} < \infty$ be the maximum life of the order. Using this notation,

$$d\tilde{Q}_\tau \equiv \tilde{Q}_\tau - \tilde{Q}_{\tau-}$$

is the quantity of the order that transacts at time τ and $\tilde{Q}_{\bar{\mathcal{T}}}$ is the total quantity of the order that executes by the maximum life of the order.

The surplus that the trader receives from execution of $d\tilde{Q}_\tau$ shares of the security in τ periods at price p_k is equal to

$$d\tilde{Q}_\tau \left(\tilde{y}_\tau + u - p_l\right) = d\tilde{Q}_\tau \left(v - p_l\right) + d\tilde{Q}_\tau \left(\tilde{y}_\tau - y\right),$$

where $\tilde{y}_\tau$ is the common value of the security in τ periods. The term $d\tilde{Q}_\tau \left(v - p_l\right)$ is equal to the surplus that a trade of size $d\tilde{Q}_\tau$ would earn upon immediate execution at price p_l. The term $d\tilde{Q}_\tau \left(\tilde{y}_\tau - y\right)$ is equal to the number of shares transacted in τ periods multiplied by the change in the common value, and so captures the picking off risk in the order. Integrating over the possible execution times for the order, and including the costs of submitting the order, qC, the realized payoff from submitting the order is equal to:

$$\int_{\tau=0}^{\bar{\mathcal{T}}} \left(v - p_l\right) d\tilde{Q}_\tau + \int_{\tau=0}^{\bar{\mathcal{T}}} \left(\tilde{y}_\tau - y\right) d\tilde{Q}_\tau - qC. \tag{4}$$

The expected payoff to the trader from submitting the order is equal to the expected value of equation (4), conditional on the trader's information. Taking expectations, the expected payoff is equal to

$$\begin{aligned}
E_t\left[U(p_l, q, \text{sell})|v\right] &= E_t\left[\int_{\tau=0}^{\bar{\mathcal{T}}} \left(v - p_l\right) d\tilde{Q}\tau \,\middle|\, q\right] + E_t\left[\int_{\tau=0}^{\bar{\mathcal{T}}} \left(\tilde{y}_\tau - y\right) d\tilde{Q}_\tau \,\middle|\, q\right] - qC \\
&= qE_t\left[\frac{\tilde{Q}_\tau}{q} \,\middle|\, q\right]\left(v - p_l\right) + E_t\left[\int_{\tau=0}^{\bar{\mathcal{T}}} \left(\tilde{y}_\tau - y\right) d\tilde{Q}_\tau \,\middle|\, q\right] - qC \\
&= q\psi_{lt}^b(q)\left(v - p_l\right) - q\xi_{lt}^b(q) - Cq. \qquad (5)
\end{aligned}$$

Here,

$$\psi^b_{lt}(q) \equiv E_t \left[\left. \frac{\tilde{Q}_{\bar{\mathcal{T}}}}{q} \right| q \right]$$

is the expected fill ratio for the order, which is equal to the expected fraction of the order that eventually transacts up to the cancellation time of the order, $\bar{\mathcal{T}}$, conditional on the information that the trader has at the time of submission and the order price chosen, p_l. If the order is a market order, then the expected fill ratio, $\psi^b_{0t}(q)$ is equal to one. We assume that as the order price moves away from the quotes, the expected fill ratio drops.

The first term in the trader's objective function, $q\psi^b_{lt}(q)\left(\tilde{v} - p_l\right)$, is equal to the expected number of shares that will eventually transact multiplied by the current surplus per share for certain execution of the order at price p_l. Thus, this quantity measures the trade–off in the order submission problem between the order price chosen and the expected number of shares transacted.

We refer to

$$q\xi^b_{lt}(q) \equiv qE_t \left[\left. \int_{\tau=0}^{\bar{\mathcal{T}}} \left(\tilde{y}_\tau - y\right) \frac{\tilde{d}Q_\tau}{q} \right| q \right]$$

as the expected picking off risk associated with the order. This variable measures the covariance of changes in the common value of the asset with the amount of the order that transacts. Thus, this variable measures the risk of the limit order executing when the common value of the asset moves against the limit order. The final term in equation (5) is Cq, and this is equal to the costs of submitting the order. This cost is the same for all order prices chosen, and increase with the size of the order. The expected payoff to a trader choosing a sell order at price p_k is defined similarly.

The trader chooses the order submission strategy which maximizes his expected payoffs. In our presentation of this problem, we have assumed that the quantity choice is exogenous to the price choice for the trade. We make this assumption because of the empirical results in Table 3 that there is very little variation in the order sizes across order submissions. We can, however, extend the model allow for the endogenous choice of quantity and a similar characterization to Proposition 1 results, with a suitable re-interpretation of the variables.

Formally, the trader chooses $d^s_k \in \{0,1\}$ for $k = 0, \ldots, K$ and $d^b_l \in \{0,1\}$ for $l = 0, \ldots, L$ to maximize

$$\sum_{k=0}^{K} d_k^s E_t\left[U(p_k, q, \text{sell})|v\right] + \sum_{l=0}^{L} d_l^b E_t\left[U(p_l, q, \text{buy})|v\right], \tag{6}$$

subject to the constraint that only one price is chosen. Let $d_{kt}^{*s}(v,q)$ and $d_{lt}^{*s}(v,q)$ be the optimal strategy. If the trader finds it optimal not to submit an order, then $d_{kt}^{*s}(v,q) = 0$ for $k = 0,\ldots,K$ and $d_{lt}^{*b}(v,q)$ for $l = 0,\ldots,L$.

Proposition 1 gives the optimal order submission strategy for a trader who enters the market with valuation $v = y + u$ who desires to trade q units of the asset, faced with a limit order book providing trading opportunities summarized by $\psi_{kt}^s(q)$, $\xi_{kt}^s(q)$, for $k = 0,1,\ldots,K$ and $\psi_{lt}^b(q)$, $\xi_{lt}^b(q)$, for $l = 0,1,\ldots,L$. We assume that as the limit order price moves from the spread, the conditional fill ratio decreases monotonically. Hollifield, Miller and Sandås (1999) show how to test this assumption.

Proposition 1 *Optimal Order Placement Strategy*

Suppose that a trader arrives at the market with valuation for the security $v = y + u$, desiring to trade q units of the security, where the trading opportunities in the limit order book can be summarized by $\psi_{kt}^s(q)$ and $\xi_{kt}^s(q)$ for all sell order choices k, with a similar definition on the buy side. The set of prices that could be optimal for some trader given this book equal $p_0^s < p_1^s < \ldots < p_K^s$ on the sell side, with p_0 equal to the bid price on the limit order book, with $p_0^b > \ldots > p_L^b$ defined similarly for the buy side. Define

$$\theta_{kt}^s(q) = p_k - \frac{(p_{k+1} - p_k)\,\psi_{k+1t}^s(q) + \left(\xi_{kt}^s(q) - \xi_{k+1t}^s(q)\right)}{\psi_{kt}^s(q) - \psi_{k+1t}^s(q)},\ \ k = 0,1,\ldots,K-1, \tag{7}$$

$$\theta_{Kt}^s(q) = p_{Kt}^s - \left(\frac{\xi_{Kt}^s(q) + C}{\psi_{Kt}^s(q)}\right), \tag{8}$$

$$\theta_{lt}^b(q) = p_l + \frac{(p_l - p_{l+1})\,\psi_{l+1t}^b(q) + \left(\xi_{l+1t}^b(q) - \xi_{lt}^b(q)\right)}{\psi_{lt}^b(q) - \psi_{l+1t}^b(q)},\ \ l = 0,1,\ldots,L-1, \tag{9}$$

$$\theta_{Lt}^b(q) = p_L^b + \left(\frac{-\xi_{Lt}^b(q) + C}{\psi_{Lt}^b(q)}\right). \tag{10}$$

The optimal decision rule is given by

$$d_{0t}^{*s}(v,q) = \begin{cases} 1, & v \le \theta_{0t}^s(q) \\ 0, & \textit{else}, \end{cases} \tag{11}$$

$$d_{kt}^{*s}(v,q) = \begin{cases} 1, & v \in \left(\theta_{kt}^s(q), \theta_{k+1t}^s(q)\right], \\ 0, & \textit{else}, \end{cases} \tag{12}$$

$$d^{*b}_{lt}(v,q) = \begin{cases} 1, & v \in \left(\theta^b_{l+1t}(q), \theta^b_{lt}(q)\right], \\ 0, & else, \end{cases} \tag{13}$$

$$d^{*b}_{0t}(v,q) = \begin{cases} 1, & v \geq \theta^b_{0t}(q) \\ 0, & else. \end{cases} \tag{14}$$

The functions $\theta^s_{kt}(q)$ and $\theta^b_{lt}(q)$ define a set of threshold valuations. They are obtained by solving for the valuation that makes a trader just indifferent between submitting an order at two adjacent price points. An implication of this model is that the associated sequence of thresholds form a monotone sequence. Hollifield, Miller and Sandås (1999) show how to test this restriction. In this paper, we will assume that this restriction holds.

We now describe the implications of the optimal order placement strategy for the conditional probability that limit and market orders are submitted. Recall that we have assumed that at time t, traders' valuations are given by $v_{jt} = y_t + u_{jt}$, with $u_{jt} \sim G_t(\cdot)$, where $G_t(\cdot)$ is a continuous distribution. We have also assumed that that y_t is common knowledge at time t. Using the optimal order placement strategy given in Proposition 1 along with the assumption that a trader enters into the market with probability $\lambda_t \Delta t$, the conditional probability that we observe a market sell order between t and Δt is equal to

$$\begin{aligned} \Pr\left(\text{Market sell over } t, t+\Delta t\right) &= \Pr\left(d^{*s}_{0t}(v,q) = 1 | \text{Trader arrives over } t, t+\Delta t,\right) \\ &\quad \times \Pr\left(\text{Trader arrives over } t, t+\Delta t\right) \\ &= \Pr\left(v_{jt} \leq \theta^s_{0t}(\Omega_t, q)\right) \lambda_t \Delta t \\ &= \Pr\left(y_t + u_{jt} \leq \theta^s_{0t}(q)\right) \lambda_t \Delta t \\ &= \Pr\left(u_{jt} \leq \theta^s_{0t}(q) - y_t\right) \lambda_t \Delta t \\ &= G_t\left(\theta^s_{0t}(q) - y_t\right) \lambda_t \Delta t. \end{aligned} \tag{15}$$

Similarly, the probability that any particular limit sell order is submitted between t and $t + \Delta t$ is

$$\begin{aligned} \Pr\left(d^{*s}_{kt}(v,q) = 1\right) &= \Pr\left(d^{*s}_{kt}(v,q) = 1 | \text{Trader arrives over } t, t+\Delta\right) \\ &\quad \times \Pr\left(\text{Trader arrives over } t, t+\Delta t\right) \\ &= \Pr\left(\theta^s_{kt}(q) - y_t \leq u_{jt} \leq \theta^s_{k+1t}(q)\right) \lambda_t \Delta t \\ &= \left[G_t(\theta^s_{kt}(q) - y_t) - G_t(\theta^s_{k+1t}(q) - y_t)\right] \lambda_t \Delta t, \end{aligned} \tag{16}$$

with similar expressions for the buy market and limit orders.

The probability that no order submission is observed between t and $t+\Delta t$ equals the probability that no trader arrives, plus the probability that a trader who arrives does not find it optimal to place an order. This is equal to

$$\Pr(\text{No order over } t, t+\Delta) \quad = \quad [1-\lambda_t \Delta t] + \left[G_t(\theta^b_{Lt}(q) - y_t) - G_t(\theta^s_{Kt}(q) - y_t)\right] \lambda_t \Delta t \quad (17)$$

The optimal order submission strategy is depicted in Figure 5 for a state where traders find it optimal to submit market order and limit orders away from the bid and ask prices, so that $K = 2$, and $L = 2$. Here, we plot the trader's private valuation for the asset, v, against the price chosen. The upper curve of the plot gives the density of the distribution of private valuations, $g(\cdot)$. The thresholds minus the common value, $\theta^s_k(q) - y$ and $\theta^b_l(q) - y$ partition the private valuations into intervals, and traders within each interval make the same order choice. For example, a trader with a valuation less than $\theta^s_0(q)$ will find it optimal to submit a sell market order. A trader with valuation between $\theta^s_0(q) - y$ and $\theta^s_1(q) - y$ will find it optimal to submit a limit sell order at the next price above the bid price on the limit order book and the mass of traders who will submit a limit sell order at the next price above the bid price is given by the area under the density of private valuations between in this area marked by diagonal lines in the plot. Traders with valuations $\theta^s_2(q) - y$ are indifferent between submitting a sell limit order at the 2^{nd} highest price above the bid price not entering any order. Similarly, traders with valuations equal to $\theta^b_2(q) - y$ are indifferent between submitting a limit buy order at the second highest limit price below the market ask price. Traders with valuations between $\theta^s_2(q) - y$ and $\theta^b_2(q) - y$ will not find it optimal to submit an order.

We are interesting in applying this model to study how trading opportunities, the arrival rates of the traders and the traders' preferences influence the supply and demand of liquidity in the market through the traders' optimal order submission strategies. To that end, the next proposition characterizes the effects of changes in the arrival rates, trading opportunities, the common value and the distribution of liquidity demand on the supply and demand for liquidity.

Proposition 2 *Comparative Statics*

1. *Changing the common value:*

$$\begin{aligned} \frac{\partial Pr(\text{Sell order over } t, t+\Delta t)}{\partial y_t} &< 0, \\ \frac{\partial Pr(\text{Buy order over } t, t+\Delta t)}{\partial y_t} &> 0. \end{aligned} \tag{18}$$

2. *Consider two dates, t_1 and t_2 where the conditional fill ratios, conditional picking off risk, common values and arrival rates of the traders are all the same. If $G_{t_1}(\cdot)$ second order stochastically dominates $G_{t_2}(\cdot)$, then*

$$Pr\left(\text{Any order observed}|t_2\right) \geq Pr\left(\text{Any order observed}|t_1\right),$$

   and

$$Pr\left(\text{Market order observed}|t_2\right) \geq Pr\left(\text{Market order observed}|t_1\right).$$

3. *Variance in the common value. Consider two times , t_1, t_2 such that the trading opportunities and current level of the common value are the same. If $var_{t_1}(y_{t_1+\Delta t}) < var_{t_2}(y_{t_2+\Delta t})$, then*

$$Pr\left(\text{Any order observed}|t_2\right) \geq Pr\left(\text{Any order observed}|t_1\right),$$

   and

$$Pr\left(\text{Market order observed}|t_2\right) \geq Pr\left(\text{Market order observed}|t_1\right).$$

4. *Arrival rate of the traders:*

$$\frac{\partial Pr\left(\text{No order observed over } t, t+\Delta t\right)}{\partial \lambda_t} < 0. \tag{19}$$

5. *Change in the conditional fill–ratio for the most aggressive limit order:*

$$\begin{aligned} \frac{\partial Pr\left(\text{Sell limit order at } K \text{ observed over } t, t+\Delta t\right)}{\partial \psi^s_{Kt}(q)} &> 0, \\ \frac{\partial Pr\left(\text{Sell order at any other price } t, t+\Delta t\right)}{\partial \psi^s_{Kt}(q)} &= 0. \end{aligned} \tag{20}$$

   Similar results hold for the buy side, and for changes in other conditional fill–ratio.

6. *Change in the conditional picking off risk:*

$$\begin{aligned} \frac{\partial Pr\left(\text{Sell limit order at } K \text{ observed}\right)}{\partial \xi^s_{Kt}(q)} &< 0, \\ \frac{\partial Pr\left(\text{Sell limit order at any other price }\right)}{\partial \xi^s_{Kt}(q)} &> 0. \end{aligned} \tag{21}$$

   Similar results hold for the buy side, and for changes in other conditional picking off risk.

Proposition 2 provides the empirical implications of the trader's optimal order choice on the provision of liquidity, through the submission of the limit orders, and on the consumption of liquidity, through market order submissions.

The first part of the proposition says that an increase in the common value will lead to in increase in the probability of submission of buy orders and decrease the likelihood of sell orders. The second part of the proposition says that an increase in the dispersion of liquidity demand will weakly increase the probability that any order is submitted, and will weakly increase the probability that market orders are submitted. Thus, an increase in the dispersion of private valuations, holding everything else constant, will tend to increase the order submissions activity in the market, and tend to lead to an increase in the submission of market order that consume liquidity in the market. This increase in the probability of order submissions will decrease the time between order submissions. Increasing the volatility of the common value, holding all else equal will also decrease the time between order submissions. Note, however, that this comparative statics results holds the conditional picking off risk constant. In particular, if both the volatility of the common value and the conditional picking off risk increase simultaneously, then the total effect of the change is ambiguous, as increasing the conditional picking off risk will tend to reduce the profits in submitting limit orders and so increase the time between order submissions. Increasing the arrival rates of traders, $\lambda(\cdot)$, will also tend increase the level of activity in the market, and thus decrease the time between order submissions. However, an increase in the arrival rate of traders will not change the relative proportion of the type of orders submitted, *conditional* upon observing an order submission, while an increase in the dispersion of liquidity demands will change both the time between submissions and the relative proportion of orders submitted, *conditional* upon submission. Our empirical work uses this fact to estimate changes in the arrival rates and the distribution of liquidity demand.

Increasing the profitability of limit orders, through increases in the conditional fill ratios or decreases in the conditional picking off risk terms, will also increase the supply of liquidity in the market. If traders believe that there will be an increase in the relative probability of market order submissions in the future, we would expect that the conditional fill ratio would tend to increase and that the conditional picking off risk would tend to drop. In the above circumstances, our model predicts an increase in the supply of liquidity through an increase in the probability of limit order submissions. In our empirical work, we will study how much of the observed variation in the speed of limit order submissions can be explained by variation in the profitability in supplying liquidity, and how much can be explain by exogenous variation in the arrival rates of traders.

We now describe our empirical strategy to estimate the model. Our data set consists of a series of order submission times, the outcomes of those orders, the order book, and and a series of conditioning variables at various times. Letting t_i denote the time of the i^{th} order entry, $d^s_{kt_i}$, $k = 0, 1, \ldots, K$ and $d^b_{lt_i}$, $l = 0, 1, \ldots, L$ denote the observed decision indicators. Then, the conditional log–likelihood function for the data is given by

$$\begin{aligned}&\sum_{i=1}^{I} \Big\{ d^s_{0,t_i} \ln G_{t_i}\left(\theta^s_{0t_i} - y_{t_i}\right) + \sum_{k=1}^{K} d^s_{kt_i} \ln\left[G_{t_i}\left(\theta^s_{k+1t_i} - y_{t_i}\right) - G_{t_i}\left(\theta^s_{kt_i} - y_{t_i}\right)\right] \\ &+ \sum_{l=1}^{L} d^b_{lt_i} \ln\left[G_{t_i}\left(\theta^b_{l-1t_i} - y_{t_i}\right) - G_{t_i}\left(\theta^b_{lt_i} - y_{t_i}\right)\right] + d^b_{0t_i} \ln\left[1 - G_{t_i}\left(\theta^b_{0t_i} - y_{t_i}\right)\right] \\ &+ \ln \lambda_{t_i} - \int_{t_{i-1}}^{t_i} \lambda_t \left(1 - \left[G_t(\theta^b_{Lt} - y_t) - G_t(\theta^s_{Kt} - y_t)\right]\right) dt \Big\},\end{aligned} \quad (22)$$

where I is equal to the total number of order entries in our data, and we drop the argument q from the thresholds to reduce notational clutter.[7]

The first two rows of equation (22) provide the contributions of the order entry decisions to the log–likelihood, while the final row provides the contributions of the time between order entries to the log-likelihood. The first term on the final row is the log of the conditional probability of a trader entering the market at time t_i, and the second term is equal to the log of the probability of observing a time interval of length $t_i - t_{i-1}$. An important feature of tick-by-tick data such as ours is duration dependence. For example, Engle (2000) and Engle and Russell (1998) find evidence that the conditional hazard rate for transactions and order displays autocorrelation and duration dependence. Our model allows for such effects in two ways. First, the conditional probability of a trader arriving, λ_t can depend on the time since the last transaction, and can be conditioned on lagged durations. Secondly, our model implies that the conditional hazard rate is time–varying, due to the effect of the term $\left(G_t\left(\theta^b_{Lt} - y_t\right) - G_t\left(\theta^s_{Kt} - y_t\right)\right)$ in the likelihood function. This term measures the probability that a trader who arrives at the market does not find it profitable to submit an order, and this depends on the picking off risk, fill ratio and level of the common value. If either the profitability of submitting orders, or the common value, or the distribution of liquidity

[7] Equation (17) implies that the hazard rate for ending a spell of no order submissions is equal to

$$\lambda_t \left(1 - \left[G_t(\theta^b_{Lt} - y_t) - G_t(\theta^s_{Kt} - y_t)\right]\right).$$

demand, $G_t(\cdot)$ changes during a period where there is no activity in the market, then our model can exhibit duration dependence, even with an *i.i.d.* arrival rate of traders.

The log–likelihood function in equation (22) depends on the common value at the time order order submission, the functions θ which depend on the conditional fill ratio and conditional picking off risk, the arrival rates of the traders, λ_t and the distribution of liquidity demand, $G_t(\cdot)$. We now explain how we deal with these terms.

We parameterize the common value in terms of observable variables, we parameterize the markets' information set in terms of observable book variables, and we parameterize the distribution of private valuations and arrival rates of the traders. Our data also allows us to infer the limit order book at the time of each order submission. For each order submitted, we can compute the realized fill ratio by following the order for the two days after it is submitted, and we know how much of the order is filled at time $t + \tau$. We use these realized fill ratios and the path of the common value to form nonparametric estimators for the terms making up the threshold functions. We now explain these steps in more detail.

In our model, we have assumed that the common value is a martingale relative to the market's information set. This implies that the common value has a unit root. In order to form an estimate of the common value, we assume that there is a vector of factors, f_t such that the common value is a linear combination of these factors,

$$y_t = \beta' f_t, \tag{23}$$

and that the factors follow non-stationary processes. We apply our empirical model to the Eurus Inc, a mining company specializing in precious metals. Our factors are high frequency futures prices for the Canadian dollar, a US Treasury bill, silver, gold, crude oil, the S &P 500 index, and a 30 year Treasury bond. Thus, we believe that with these factors, equation (23) is a reasonable assumption for this firm. In order to estimate β to form the common value, we assume that the bid quote at time t, is cointegrated with the common value, so that

$$\begin{aligned} p^s_{0t} &= y_t + \nu_t \\ &= \beta' f_t + \nu_t, \end{aligned} \tag{24}$$

where p^s_{0t} is the bid quote at time t and ν_t is a stationary process, This implies that we can use a cointegrating regression between p^s_0 and the factors to estimate β. Engle and Granger (1987) show that this estimator of β is super–consistent, and De Jong (2000) provides conditions in which using a first–stage cointegrating regression in a second–stage non–linear model results in consistent and asymptotically normal second–stage estimators. These conditions hold in our case. Letting $\hat{\beta}$ denote the estimate of β obtained through the cointegrating regression, our estimate of the common value equals

$$\hat{y}_t = \hat{\beta}' f_t. \tag{25}$$

We now turn to estimating $\theta^s_{kt}(q)$ and $\theta^b_{lt}(q)$. These functions depend on the conditional fill ratios and picking off risk terms for each order at each information set, as well as the order entry cost, C. To estimate the conditional fill ratios and picking off risk terms, we make a standard rational expectations assumption. That is, at every information set, we assume that the traders' have rational expectations about the conditional fill ratios and conditional picking off risk for each order, and so compute their expectations using the non-parametric regressions of the realized execution history of each of the orders, conditional a set of I observable state variables, $X'_t \equiv [1, X_{1t}, X_{2t}, \ldots, X_{It}]$ to capture the traders' information sets. Hotz and Miller (1993) and Manski (1991) shows how this method leads to consistent estimators for discrete choice models in an *i.i.d.* environment, and Ahn and Manski (1993) show how to compute standard errors in this environment. We extend these results to deal with the time–series nature of our data.

Given our assumption that the traders' information set can be represented by the observable state variables X_t, we assume that the private values distribution $G_t(\cdot)$ is normal, with mean zero and standard deviation, σ_t satisfying

$$\frac{1}{\sigma_t} = S' X_t, \tag{26}$$

where $S \equiv [S_0 : S_1]$, S_0 is a scalar and S_1 is a vector. If the distribution of liquidity demand is constant, then $S_1 = 0$. The arrival rate of the traders, λ_t is parameterized as

$$\lambda_t = \exp(L' X_t), \tag{27}$$

where $L \equiv [L_0 : L_1]$, L_0 is a scalar and L_1 is a vector. If the arrival rate of the traders is constant, then $L_1 = 0$. We note here that this parameterization is the simplest specification of the conditional hazard rates. We can extend our model to allow for more elaborate hazard rates.

In the appendix, we use assumptions (24) through (27) to compute an approximation to the conditional log–likelihood function, equation (22) that we use to estimate the parameter vector (S, L, C), and we provide the asymptotic standard errors for this procedure.

3 Empirical Results

In this section, we describe the empirical results we obtained from estimating the model in our data. First, we discuss our estimation of the common value of the Eurus Inc., using a cointegrating regression. Next, we provide estimates of the fill ratios, picking off risk terms and the thresholds. In the final section, we present the estimates obtained from estimating the model using the observed timing of decisions and order choices in our data.

Table 5 shows the futures prices that we use to proxy for our factors in estimating the common value of the security. The futures prices are provided by the Futures Data Institute, and are sampled tick–by–tick. We align these prices to the time of each event in our equity market data set by using the last available observation of the futures price. Since the futures markets and the VSE have different opening hours, there are periods of time when no recent observation of the futures prices can be obtained. We drop order submissions that occur during these times from our data set. In particular, we drop all new order submissions occurring after 12:30 p.m. Pacific Time. For orders submitted before 12:30 that actually transact between 12:30 and the closing time of the market 1:30, we use the opening value of the futures contracts the next day to infer changes in the common value. We also drop order submissions on days when the futures markets are closed and on the immediately prior days.

Although this is surely not a perfect set of proxies for the common value, these futures prices do a good job of explaining a portion of the price variations accounted for by public information. Stocks on the VSE during our sample period are primarily small resource based firms, with a high exposure to fluctuations in resource prices. In addition, because of the high exposure of the Canadian economy as a whole to fluctuations in resource prices, these futures prices can proxy for the overall health of the Canadian economy, which will in turn influence the health of the individual VSE companies. The Canadian dollar exchange rate series will, in particular, proxy for the output prices for the majority of the firms on the VSE, since output prices are almost always quoted in U.S. dollars.

Table 6 shows that the cointegration assumptions accurately describe our data. The first column gives evidence from an Augmented Dickey–Fuller test that each of the factor series and the bid price series possess a unit root. The second column shows the coefficients of the cointegrating regression estimated by ordinary least squares. The third column shows the long–term implied coefficients obtained from an error correction model estimation. The error correction model cointegration test, as described in Banerjee, Dolado, and Mestre (1998), finds evidence that the factors are cointegrated with the bid price series.

We now turn to the results obtained from estimating the theoretical model. Given the evidence in Table 2 that dividing the possible limit order into prices 1 cent apart would leave us with few observations of each price choice, we decided to combine prices close to each other into categories. Thus, we applied our model to the trader's choice of market orders and limit orders at a price between 0 and 5 cents from the quotes, 6 and 10 cents from the quotes, and a price greater than 10 cents from the quotes. For each of these limit order categories, we computed the unconditional fill ratios and picking off risk terms. The picking off risk terms are computed using estimates of the common value obtained from the co-integration procedure described above. That is, for each order submitted, we calculate the change in the common value between the time of each transaction associated with the order times the number of shares transacted.

Table 7 contains the estimates of the fill ratios, picking off risk terms and estimates of associated thresholds for each category. The unconditional fill ratios are estimated fairly precisely, and satisfy the requirement that the farther away the order is from the quotes, the lower the associated fill ratio. For example, a sell order submitted at a price 5 cents higher than the market as has a 0.5477 fill ratio, on average. The average picking off risk terms are not estimated with much precision, and all the point estimated are small. This suggests that on average, there is little relationship between the picking off risk and the order price chosen, and that limit orders face little adverse selection risk in this market, on average. The fourth row of the table contain unconditional estimates of the thresholds in our model, θ_k^s and θ_l^b, computed according from equations (7) through (9), less the average deviation from the price and the common value $\bar{p}_{kt}^s + \theta_k^s - \bar{y}_t$ and $\bar{p}_{kt}^b + \theta_k^2 - \bar{y}_t$, where the bar denotes sample average and $\bar{y}_t$ is the average common value in our sample.

Table 8 contains the parameter estimates we obtained from estimating the theoretical model, using out estimates of the fill ratios and picking off risk terms along with the time series of the common value and order prices in the limit order book. In our estimation, we approximate the integral in the log–likelihood function in equation (22) with a sum, and we set the time step in the approximation, Δt, equal to 15 minutes in this approximation[8]. We estimated various specifications of the model. Panel A provides the results from estimating a base model, where we impose the constraint that the variance liquidity demand and the arrival rate of traders is constant. In panel B of the table, we allow the depth in the order book at the time of order entry to enter into the conditional volatility of liquidity demand and the arrival rate of the traders. Panel C uses the trading volume over the last hour as the state variable, and Panel D uses the volatility of common value returns over the last hour. The depth variable was chosen to proxy for measures of competition in the limit order book, the volume variable was chosen to measure changes in trading activity and the volatility measure was chosen to provide a proxy for the information flow and trading activity. For each of the state variables, we allow the state variable to influence the dispersion in liquidity demand and the arrival rate of trading, by themselves and in combination. Thus, for each state variable, we estimated 3 different specifications.

Panel A of Table 8 reports the parameter estimates for the base model. The estimates for the precision of liquidity demand of 2.6126 corresponds to a standard deviation of of liquidity demand distribution of roughly 38 cents. Our estimate of the arrival rate corresponds to an average time between arrivals of approximately 175 seconds, which is somewhat higher than the sample average of 146 seconds. Finally, our estimate of the entry cost is -15.44 cents. This point estimate implies that all traders, on average, find it profitable to enter orders into the limit order book. In particular, this estimate implies that aggressive limit orders, far away from the quotes, appear to be profitable. Given that limit orders appear to be profitable in the data and that there are often long time gaps between order submissions, our estimate implies that liquidity suppliers do not constantly monitor

[8] We experimented with changing the time step, as small at 10 seconds and as large as 20 minutes. Our estimates are relatively insensitive to changes in Δt.

profit opportunities in the order book. In this base specification of the model, all of the empirical variation in the time between orders and the relative proposition of order types is driven by changes in the conditional profitability of providing and consuming liquidity.

Panel B of Table 8 provides the parameter estimates for a specification of the model that allows the dispersion of liquidity demand and the arrival rate to depend on the order book depth, both individually and jointly. The point estimate of the effect on the precision, S_t, is positive and estimated precisely, and the estimated effect of the depth on the arrival rate is positive, but insignificant. This implies that when the order book depth is high, there is a smaller dispersion of the liquidity demand in the market. In our theoretical model, an increase in the precision of liquidity demand leads to a decrease in the probability of observing an order submission, and an increase in the probability observing a limit order, *conditional* on observing any order submission, holding the profitability of the orders fixed. Empirically, the conditional fill ratio decreases when the depth of the limit order book increases so that limit orders are less profitable as the depth of the limit order book increases. Our estimates imply that when the limit order book is deep, the decrease in profitability of placing limit orders is offset by a decrease in the dispersion of liquidity demand, so that the probability of observed a limit order conditional upon observing any order increases, even though each limit order submission itself becomes less profitable.

The parameter estimates in panel C show that increased lagged volume leads to an increase in the precision of liquidity demand, implying that when lagged volume is high, the standard deviation of liquidity demand is lower. Thus, we find evidence that the periods of high trading volume tend to attract further supply of liquidity, as more traders who find it profitable to submit limit orders enter the market. In addition, lagged trading volume has a weakly positive effect on the arrival rate of the traders. When lagged volume is allow to effect the arrival rate and volatility of liquidity demand, it has a positive effect of the arrival rate but an insignificant effect on liquidity demand.

Panel D contains the parameter estimates when lagged common value volatility is used as a state variable. We find that high volatility in the common value has a weakly positive effect on the precision of liquidity demand and an insignificant effect on the arrival rate of traders. Thus, there is some evidence that periods of high volatility may reduce the demand for liquidity, perhaps because of the increased uncertainty about the value of the asset.

Overall, we find economically reasonable parameter estimates for the volatility of liquidity demand and the arrival rates of the traders. However, in all specifications that we considered, we estimate the costs of order submissions to be negative. We interpret this finding as evidence that, on average, there are unexploited profit opportunities in the limit order book in this market. Our estimates also imply that the arrival rates of traders, λ_t is relatively constant, while the distribution of liquidity demand is state dependent.

4 Conclusions

In this paper, we present a unifying model of liquidity supply and demand, which incorporates into a single framework the adverse selection risk and execution risk inherent in a limit order market. The decision of an individual trader, arriving at a limit order market and choosing a one–shot trading strategy, is solved to obtain the optimal order–submission strategy. The order–submission may be a market order which uses up liquidity available in the book, a limit order which supplies liquidity to the book, or no order at all. The optimal strategy depends on the distribution of the observable common–value, the private liquidity demand of the trader, and the stochastic arrival rate of traders, as well as a fixed cost of submission.

This parsimonious model of the limit–order submission decision allows us to make several empirically testable predictions about the supply and demand for liquidity. All else being equal, either increasing the dispersion of private valuations or increasing the arrival rates of traders tends to increase activity in the market. In addition, the supply of liquidity is positively related to the expected fill ratio and negatively related to the degree of adverse–selection risk.

Our empirical work based on this model is able to decompose observed variability in market activity into a component caused by changes in profitability of supplying liquidity to the market and a component caused by exogenous changes in the arrival rate of traders. We estimate our model using a rich and detailed set of data which includes the histories of all orders and trades on the Vancouver Stock Exchange. The common value of the stock is modeled using a set of exogenous factors, while the distribution of the traders' private valuation and the arrival rates of traders are estimated using a semi–parametric maximum–likelihood technique.

Using these techniques, we are able to obtain economically reasonable parameter estimates for the volatility of liquidity demand and the arrival rate of traders. Our results provide evidence that greater book depth and higher trading volume each shift the conditional probability of a limit order submission more than predicted by the change in profitability alone. This leads to an increase in the supply of liquidity. Furthermore, our estimate of order submission cost is significantly negative under all specifications. This implies that traders are not able to exploit all profit opportunities available on the book, perhaps because the cost of continuously monitoring the book is high.

Overall, this theoretical and empirical model provides a valuable means to examine the underlying costs and available profits present in limit order markets. Our technique is particularly illuminating, in that it allows us to decompose observed variations in market activity into the underlying economic forces that affect the order submission strategy. In this way, we gain a better understanding of how the optimal behavior of traders is influenced by factors inside and outside the market.

References

Admati, A., and P.Pfleiderer, 1988, 'A Theory of intraday patterns: Volume and price variability', *Review of Finacial Studies*, 1, 3-40.

Ahn, H., and C. Manski, 1993, 'Distribution theory for the analysis of binary choice under uncertainty with nonparametric estimation of expectations', *Journal of Econometrics*, 56, 291–321.

Amemiya, T., 'Advanced Econometrics', 1985, Harvard University Press, Cambridge, MA.

Banerjee, A., J. Dolado, J. Galbraith, and D. Hendry, 1993, 'Co–Integration, Error Correction and the Econometric Analysis of Non–Stationary Data', Oxford University Press, .

Banerjee, A., J. Dolado, and R. Mestre, 1998, 'Error–Correction mechanism tests for cointegration in a single–equation framework', *Journal of Time Series Analysis*, 19, 267–283.

Biais, B., P. Hillion and C. Spatt, 1995, 'An empirical analysis of the limit order book and order flow in the Paris Bourse', *Journal of Finance*, 50, 1655-1689.

Black, F., Towards a fully automated exchange, part I, *Financial Analysts Journal*, 27, 29–34.

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.

Copeland, T. and D. Galai, 1983, 'Information effects and the bid-ask spread', *Journal of Finance* 38, 1457-1469.

Coppejans, M., and I. Domowitz, 1999, 'Screen information, trader activity, and bid—ask spreads in a limit order market', working paper, Pennsylvania State University.

Coppejans, M., I. Domowitz, and A. Madhavan, 2000, 'Liquidity in an automated auction', working paper, Pennsylvania State University.

De Jong, R.M., 'Nonlinear estimation using estimated cointegrating regressions', *Journal of Econometrics*, forthcoming.

Demsetz, H., 1968, 'The cost of transacting', *Quarterly Journal of Economics*, 82, 33-53.

Dow, J., 1999, 'Self–sustaining liquidity in an asset market with asymmetric information', working paper, London Business School.

Easley, D., N. Keifer and M. O'Hara, 1997, 'One day in the life of a very common stock', *Review of Financial Studies*, 10, 805–835.

Easley, D. and M. O'Hara, 1992, 'Time and the process of security price adjustment', *Journal of Finance*, 47, 577-605.

Engle, R., 2000, The econometrics of ultra—high frequency data, *Econometrica* 68, 1-12.

Engle, R. and C.W.J. Granger, 1987, 'Co–integration and error correction: Representation, estimation and testing', *Econometrica*, 55: 251–76.

Engle, R. and J. Russell, 1998, 'Autoregressive conditional duration: A new model for irregularly spaced transaction data', *Econometrica*, 66, 1127-1162.

Engle, R. and J. Lange, 1997, 'Measuring, forecasting and explaining time varying liquidity in the stock market', working paper, University of California, San Diego.

Ericsson, N. and J. MacKinnon, 1999, 'Distributions of error correction tests for cointegration', working paper number 655, The Federal Reserve Board, International Finance Discussion Papers.

Foster, D., and S. Viswanathan, 1993, 'Variations in trading volume, return volatility, and trading costs: Evidence from recent price formation models', *Journal of Finance* 48, pp. 187-211.

Foucault, T., 1999, 'Order flow composition and trading costs in a dynamic limit order market', *Journal of Financial Markets*, 99-134.

French, K., and R. Roll, 1986, 'Stock return variances: The arrival of information and the reaction of traders', *Journal of Finacial Economics* 17, 5-26.

Glosten, L. 1994, 'Is the electronic limit order book inevitable?', *Journal of Finance*, 49, 1127-1161.

Glosten, L., 2000, 'Competition, design of exchanges and welfare', working paper, Columbia University.

Goldstein, M. and K. Kavajecz, 2000, 'Liquidity provision during circuit breakers and extreme market movements', Rodney White Center #001-00, University of Pennsylvania.

Grossman, S., and M. Miller, 1988, 'Liquidity and market structure', *Journal of Finance*, 43, 617-633.

Hamao, Y. and J. Hasbrouck, 1995, 'Securities trading in the absence of dealers: trades and quotes on the Tokyo Stock Exchange', *Review of Financial Studies*, 8, 849–878.

Harris, L. and J. Hasbrouck, 1996, 'Market vs. limit orders: the SuperDot evidence on order submission strategies', *Journal of Financial and Quantitaive Analysis*

Hausman, J, A. Lo, and A.C. MacKinlay, 1992, 'An ordered probit analysis of transaction stock prices', *Journal of Financial Economics*, 31, 319-379.

Hollifield, B., R. A. Miller, and P. Sandås, 1999, 'Empirical analysis of limit order markets', working paper, Carnegie Mellon University.

Hotz, J., and R. A. Miller, 1993, 'Conditional choice probabilities and the estimation of dynamic models', *Review of Economic Studies*, 60, 497–529.

Kiefer, N., 1988, 'Economic duration data and hazard functions', *Journal of Economic Literature*, 26, 646–679.

Lancaster, T. 1990,'The Econometric Analysis of Transition Data', Cambridge University Press.

Lo, A., C. MacKinlay and J. Zhang, 'Econometric models of limit order executions', forthcoming, *Journal of Financial Economics.*

Manski, C., 1991, 'Nonparametric estimation of expectations in the analysis of discrete choice under uncertainty', in *Nonparametric and semiparametric methods in econometrics and statistics*, W. Barnett, J. Powell and G. Tauchen, ed, Cambridge University Press.

Newey, KW.K. and K.D. West, 1987, 'A simple, positive semi–definite, heteroskedasticity and autocorrelation consistent covariance matrix', *Econometrica*, 55, 703–708.

Parlour, C., 1998, 'Price dynamics in limit order markets', *Review of Financial Studies*, 789-816.

Pagano, M., 1989, 'Trading volume and asset liquidity', *Quarterly Journal of Economics*, 104, 255-274.

Pagano, M., 1989, 'Endogenous stock market thinness and stock price volatility', *Review of Economic Studies*, 56, 269-288.

Robinson, P.M, 1989, 'Hypothesis testing in semiparametric and nonparametric models for economic times series', *Review of Economics Studies*, 60, 511–534.

Sandås, P., 2000, 'Adverse selection and competitive market making: empirical evidence from a limit order market', forthcoming, *Review of Financial Studies*

Seppi, D., 1997, 'Liquidity provision with limit orders and a strategic specialist', *Review of Financial Studies*, 103-150.

Tauchen, G. and M. Pitts, 1983, 'The price variability–volume relationship on speculative markets', *Econometrica*, 51, 485–505.

A Proofs

To prove Propositon 1, we start with a preliminary lemma.

Lemma 1 *Revealed preference implies that for the two buyer with valuations v and v', $v' > v$, respectively, who optimally choose to submit buy orders at prices l and l' we have:*

$$(v - v')(\psi^b_l - \psi^b_{l'}) \geq 0.$$

Proof of Lemma 1
By optimality,

$$q\psi^b_l(q)(v - p_l) - q\xi^b_l(q) - qC \geq q\psi^b_{l'}(q)(v - p_{l'}) - q\xi^b_{l'}(q) - qC,$$

$$q\psi^b_{l'}(q)(v' - p_l) - q\xi^b_{l'}(q) - qC \geq q\psi^b_l(q)(v' - p_l) - q\xi^b_l(q) - qC.$$

Multiplying the second inequality buy -1 and adding and rearranging yields:

$$(v - v')(\psi^b_l - \psi^b_{l'}) \geq 0$$

■

Proof of Proposition 1
For a buyer who is indifferent between submitting a buy order at the lowest possible price p^b_L and not entering an order the threshold valuation $\theta^b_L(q)$ solves:

$$q\Psi^b_L(q)(\theta^b_L(q) - p^b_l) - q\xi^b_L(q) - Cq = 0$$

$$\theta^b_L(q) = p^b_L + \frac{-\xi^b_L(q) + C}{\Psi^b_L(q)}$$

Lemma 1 shows that the traders order placement strategies are monotone in their valuations. Thus, if all buy order prices $p^b_0 > \ldots > p^b_L$ are optimal for some buyers then the corresponding threshold valuations form a monotone sequence $\theta^b_0(q) > \ldots > \theta^b_L(q)$ that divides the valuation into $L + 1$ intervals. The optimal decision rule is a direct mapping from these $L + 1$ valuation intervals into the L possible prices $p^b_0 \ldots p^b_L$ and the no-order decision.

Analogous arguments apply to the sell side. ■

Proof of Propositon 2

The comparative statics follow up inspection of the definitions of the thresholds and the conditonal choice probability. The comparative statics with respect to a second order stochastic dominance shift in $G_(\cdot)$ and the conditional volatility of the common both follow because the mapping from valuations to choices is strictly concave. ■

B Econometrics Appendix

In this section, we provide details on the kernel estimator used to compute the conditional fill ratios and winners curse terms, the likelihood function for our estimator, and the formula for the standard errors for the estimators.

We use a product of three univariate Nadarya-Watson kernels to estimate the functions $\psi(\cdot)$ and $\xi(\cdot)$. Let $X_t \equiv (Qsm_t, VOL_t.VTY_t)$ denote the vector of state variables. The estimators we use are defined as

$$
\begin{aligned}
\hat{\psi}_k^s(X_t) &= \frac{\sum_{i=1}^N \mathcal{K}(\frac{QSm_{i,1}-QSm_{t,1}}{h_{QSm}})\mathcal{K}(\frac{VOL_{i,2}-VOL_{t,2}}{h_{VOL}})\mathcal{K}(\frac{VTY_{i,3}-VTY_{t,3}}{h_{VTY}})\frac{Q_{\mathcal{T}}}{q}}{\sum_{n=1}^N \mathcal{K}(\frac{QSm_{i,1}-QSm_{t,1}}{h_{QSm}})\mathcal{K}(\frac{VOL_{i,2}-VOL_{t,2}}{h_{VOL}})\mathcal{K}(\frac{VTY_{i,3}-VTY_{t,3}}{h_{VTY}})}, \\
\hat{\xi}_l^s(X_t) &= \frac{\sum_{i=1}^N \mathcal{K}(\frac{QSm_{i,1}-QSm_{t,1}}{h_{QSm}})\mathcal{K}(\frac{VOL_{i,2}-VOL_{t,2}}{h_{VOL}})\mathcal{K}(\frac{VTY_{i,3}-VTY_{t,3}}{h_{VTY}})\int \Delta\hat{y}_\tau \frac{dQ_\tau}{q}}{\sum_{n=1}^N \mathcal{K}(\frac{QSm_{i,1}-QSm_{t,1}}{h_{QSm}})\mathcal{K}(\frac{VOL_{i,2}-VOL_{t,2}}{h_{VOL}})\mathcal{K}(\frac{VTY_{i,3}-VTY_{t,3}}{h_{VTY}})},
\end{aligned}
$$

for $k = 1, \ldots, K$ with similar estimates on the buy side. Here $\mathcal{K}(\cdot)$ is a higher order Gaussian kernel function, with bandwidth sequence $\hat{\sigma}h_X^T$, where $\hat{\sigma}$ is the sample standard deviation of the associated conditioning variable, X. In our case, the kernel is of order 6, and the bandwidth sequence satisfies $constant \times T^{-\kappa_1}$ where $\frac{1}{12} < \kappa_1 < \frac{1}{5}$.

To derive the conditional log–likelihood function, we define the variables,

$$
\begin{aligned}
\nu_t &= p_{0t}^s - \beta' f_t, \\
m_{kt}^s &= -\frac{(p_{k+1} - p_k)\,\psi_{k+1}^b(X_t, q) + \left(\xi_k^s(X_t, q) - \xi_{k+1}^s(X_t, q)\right)}{\psi_k^s(X_t, q) - \psi_{k+1}^s(X_t, q)}, \;\; k = 0, 1, \ldots, K-1, \\
m_{Kt}^s &= -\left(\frac{\xi_K^s(X_t, q)}{\psi_K^s(X_t, q)}\right), \\
m_{lt}^b &= \frac{(p_l - p_{l+1})\,\psi_{l+1}^b(X_t, q) + \left(\xi_{l+1}^b(X_t, q) - \xi_l^b(X_t, q)\right)}{\psi_l^b(X_t, q) - \psi_{l+1}^b(X_t, q)}, \;\; l = 0, 1, \ldots, L-1, \\
m_{Lt}^b &= \left(\frac{-\xi_L^b(X_t, q)}{\psi_L^b(X_t, q)}\right).
\end{aligned}
$$

Using these definitions,

$$
\begin{aligned}
\theta_k^s(X_t, q) - y_t &= m_{kt}^s + p_{kt}^s - p_{0t}^s + \nu_t \\
&\equiv z_{kt}, \;\; k = 0, 1, \ldots, K-1,
\end{aligned}
$$

$$
\begin{aligned}
\theta_K^s(X_t, q) &= m_{Kt}^s - C\frac{1}{\psi_K^s(X_t, q)} + p_{Kt}^s - p_{0t}^s + \nu_t \\
&= z_{Kt}(C)
\end{aligned}
$$

with similar expressions for the buy side. To approximate the conditional log likelihood function, we divide time into steps of length Δt and we have $T_{\Delta t}$ such intervals in our data. We record a decision indicator over each time interval, where $d_t^s = d_t^b = 0$ during an interval $[t, t+\Delta t]$ where no order submission is observed. Using this notation, the conditional log likelihood is approximated by

$$\mathcal{L}(B) = \sum_{t=1}^{T} L_t(B), \tag{A1}$$

where

$$\begin{aligned} L_t(B) = {} & d^s_{0,t} \ln \Phi\left(S' X_t \hat{z}^s_{0t}\right) + \sum_{k=1}^{K-1} d^s_{kt} \ln\left[\Phi\left(S' X_t \hat{z}^s_{kt}\right) - \Phi\left(S' X_t \hat{z}^s_{k-1t}\right)\right] \\ & + d^s_{Kt} \ln\left[\Phi\left(S' X_t \hat{z}^s_{Kt}(C)\right) - \Phi\left(S' X_t \hat{z}^s_{K-1t}\right)\right] \\ & + d^b_{Lt} \ln\left[\Phi\left(S' X_t \hat{z}^b_{L-1t}\right) - \Phi\left(S' X_t \hat{z}^b_{Lt}(C)\right)\right] \\ & + \sum_{l=1}^{L-1} d^b_{lt} \ln\left[\Phi\left(S' X_t \hat{z}^b_{l-1t}\right) - \Phi\left((S' X_t \hat{z}^b_{lt}\right)\right] + d^b_{0t} \ln\left[1 - \Phi\left(S' X_t \hat{z}^b_{0t}\right)\right] \\ & + d^n_t \ln\left[(1 - \exp(L' X_t)\Delta t) + \exp(L' X_t)\Delta t\left(\Phi\left(S' X_t \hat{z}^b_{Lt}(C)\right) - \Phi\left(S' X_t \hat{z}^s_{Kt}(C)\right)\right)\right] \\ & + (1 - d^n_t) \ln\left[\Delta t \exp(L' X_t)\right]. \end{aligned} \tag{A2}$$

Here $d^n_t = \left(1 - \sum_{k=0}^{K} d^s_{kt} - \sum_{l=0}^{K} d^b_{lt}\right)$ is an indicator variable that is equal to one if no order submission is observed between t and $t + \Delta t$, $B \equiv (S, L, C)$ is the vector of parameters, $\Phi(\cdot)$ is the cumulative standard normal distribution function, and the hat over top of a variable means that is is pre–estimated, either through a cointegrating regression for $\hat{\nu}_t$ or through first–stage kernel regressions for $\hat{m}^s_{kt}$. Here, we approximate the continous time hazard in equation (22) with an approximating discrete time process, with time steps Δt. Our estimates maximize a trimmed version of this conditional log–likelihood function,

$$\hat{B}_T \equiv \arg\max_{B \in \mathcal{B}} \sum_{t=1}^{T} I(X_t \in X) L_t(B), \tag{A3}$$

where, $I(X_t \in X)$ is a trimming indicator for X_t in trimming set X in the interior of the support of X_t.

To compute the standard errors of the estimators, define the derivatives of the conditional likelihood function, evaluated at the true parameters, $B_0, \beta, \psi(X), \xi(X)$

$$\begin{aligned} L_{tB} &= I(X_t \in X) \frac{\partial L_t(B_0)}{\partial B}, \\ L_{tBB} &= I(X_t \in X) \frac{\partial^2 L_t(B_0)}{\partial B \partial B'}, \\ L_{tB\nu} &= I(X_t \in X) \frac{\partial^2 L_t(B_0)}{\partial B \partial \nu'}, \\ L_{tB\psi} &= I(X_t \in X) \frac{\partial^2 L_t(B_0)}{\partial B \partial \psi'}, \\ L_{tB\xi} &= I(X_t \in X) \frac{\partial^2 L_t(B)}{\partial B \partial \xi'}, \end{aligned}$$

and

$$
\begin{aligned}
\mathcal{I}_B &= E\left[L_{tB}\right], \\
\mathcal{J} &= E\left[L_{tB\nu}\right],
\end{aligned}
$$

and

$$
\begin{aligned}
\mu_\psi(X_t) &= E\left[L_{tB\psi} | \, X_t\right] \\
\mu_\xi(X_t) &= E\left[L_{tB\xi} | \, X_t\right].
\end{aligned}
$$

Also, define the following functions of the error terms in the kernel estimates of the the conditional fill ratios and winners curse terms and conditional probabilities of different order choices as:

$$
\begin{aligned}
\varphi_{kt} &= I(d_{kt}^s = 1) \frac{\left(\int_{\tau=0}^{\bar{\mathcal{T}}} \frac{dQ_{t+\tau}}{q_t} - \psi_{kt}^s(X_t) \right)}{\Pr(d_{kt}^s = 1 | X_t)} \\
\rho_{kt} &= I(d_{kt}^s = 1) \frac{\left(\int_{\tau=0}^{\bar{\mathcal{T}}} (\beta' \Delta f_{t+\tau}) \frac{dQ_{t+\tau}}{q_t} - \xi_{kt}^s(X_t) \right)}{\Pr(d_{kt}^s = 1 | X_t)},
\end{aligned}
$$

for the $k = 1, \ldots, K$ sell decisions, similarly for the buy decisions. Let φ_t and ρ_t denote the respective vectors these functions.

Under the appropriate regularity conditions, and applying the results in De Jong (2000), Ahn and Manski (1993) and Hollifield, Miller and Sandås (1999), this estimator is consistent, and asymptotically normal,

$$
\sqrt{T}\left(\hat{B}_T - B_0\right) \overset{d}{\rightsquigarrow} \mathcal{N}(0, \Sigma), \tag{A4}
$$

with

$$
\Sigma = \mathcal{I}_B^{-1}[I_1 : \mathcal{J} : I_2] \mathcal{A} [I : \mathcal{J} : I]' \mathcal{I}_B^{-1}, \tag{A5}
$$

where I_1 and I_2 are appropriately dimensioned identity matrices and $\mathcal{A}$ is the long run variance–covariance matrix of the vector

$$
T^{-1/2} \left[\sum_t L_{tB}, \, \sum_t \nu_t, \, \sum_t (\mu_\psi(X_t)' \varphi_t + \mu_\xi(X_t)' \rho_t) \right]. \tag{A6}
$$

The matrix $\mathcal{A}$ is estimated consistently using a standard Newey–West procedure, $\mathcal{I}_B$ and $\mathcal{J}$ are estimated with the appropriate sample averages and $\mu_\psi(X_t)' \varphi_t + \mu_\xi(X_t)' \rho_t$ is estimated using the a second stage kernel estimate of the conditional choice probabilities, conditional on X_t along with the residuals from the kernel estimators of the conditional fill ratios and winner's curse terms.

Table 1: **Summary Statistics on the Order Flow.**

Item	Mean	S.D.	Percentiles				
			5th	25th	50th	75th	95th
Panel A: Price and Volume Statistics							
Closing Midquote $	2.52	0.55	1.47	2.31	2.43	2.91	3.41
Open-to-Close Return %	-1.82	12.54	-10.49	-4.61	-0.79	0.96	13.15
Close-to-Open Return %	9.92	75.54	-5.26	-0.8	0.09	2.96	8.26
Trading Volume—1000 shares	118.44	96.01	16.26	49.03	89.10	157.17	303.14
Trading Volume—$1,000	313.50	279.57	34.89	115.69	250.26	424.88	924.27
Bid-ask Spread $	0.044	0.011	0.029	0.036	0.043	0.052	0.062
Bid + Ask Depth—1,000 shares	6.78	4.54	3.34	4.34	5.10	7.40	14.72
Panel B: Daily Statistics on Order Types 10,608 observations							
# Market Orders–complete fill	53.08	40.39	11.25	26.50	47.00	64.50	121.30
# Market Orders–partial fill	14.65	11.48	2.30	6.00	12.00	19.75	38.05
# Limit Orders	100.84	75.18	23.00	49.00	89.00	128.75	226.90
# day orders	157.22	119.97	38.95	73.25	137.00	197.50	358.30
Orders lasting $>$ 1 day	4.02	3.37	0	2.00	4.00	6.00	10.35
Orders lasting $<$ 1 day	1.30	1.86	0	0	0	2.00	5.35
Orders finished 2^{nd} day	0.29	0.73	0	0	0	0	2.00
Order lasting $>$ 2 days	1.98	2.28	0	0	1.00	3.00	6.00

Statistics for the daily closing mid-quote, open-to-close and close-to-open mid-quote returns are reported on the first three rows. The trading volume measured in thousands of shares and dollars is presented on the fourth and fifth rows, followed by the dollar bid-ask spread and the sum of the depth at the best bid and ask quotes. The statistics reported for the spread and the depth, for example, the standard deviation, is computed for the time series of daily averages spreads and depths. The daily averages are computed as equally weighted averages across each distinct order book observed. The lower half of the table reports statistics on the daily numbers of different types of orders. Markets orders are separated depending on whether they were filled completely or partially. Limit orders are categorized based on how long they remain outstanding, for example, orders that remain outstanding for more than one trading day after submission.

Table 2: **Descriptive Statistics on the Order Flow**

order price distance	N	Order size: Mean	S. D.	Percentiles 5th	25th	50th	75th	95th	Fill-ratio Mean	S.E.	Time-to-Fill minutes Mean	S.E.
Market buy orders including partial-fills												
	2334	2.58	2.67	0.4	1.0	2.0	3.0	7.5	0.962	0.003	2.1	0.25
Limit buy orders by distance from ask quote $ (N=3011)												
-0.01	128	2.56	2.21	0.8	1.0	2.0	3.0	5.0	0.586	0.042	5.9	1.31
-0.02	251	2.50	1.74	0.5	1.0	2.0	3.0	5.0	0.672	0.029	12.5	1.92
-0.03	319	2.55	2.18	0.5	1.0	2.0	3.0	5.0	0.570	0.027	16.2	3.77
-0.04	363	2.57	3.05	0.5	1.0	2.0	3.0	5.0	0.513	0.026	19.2	6.21
-0.05	770	2.48	2.43	0.5	1.0	2.0	3.0	5.6	0.596	0.017	31.3	8.27
-0.06	119	2.82	2.05	1.0	1.3	2.0	3.9	5.0	0.315	0.042	24.2	9.31
-0.07	152	2.58	1.94	0.5	1.0	2.0	3.0	5.0	0.354	0.037	22.6	4.63
-0.08	122	2.59	1.79	0.9	2.0	2.0	3.0	5.0	0.251	0.039	14.3	3.54
-0.09	106	2.57	1.73	0.9	1.0	2.0	3.0	5.0	0.303	0.043	95.6	79.16
-0.10	267	2.83	2.89	0.6	1.0	2.0	3.0	10.0	0.339	0.029	108.9	50.10
-0.11	27	2.47	1.63	0.5	1.0	2.0	3.0	5.1	0.185	0.076	10.4	5.12
-0.12	32	2.39	1.23	1.0	2.0	2.0	3.0	5.0	0.312	0.083	13.0	2.28
-0.13	29	2.71	2.40	0.5	1.0	2.0	3.0	10.0	0.288	0.082	39.2	17.01
-0.14	32	2.57	2.16	0.7	1.6	2.0	3.0	8.4	0.219	0.074	22.5	10.16
-0.15	87	3.05	2.43	0.5	1.4	2.0	3.9	10.0	0.243	0.045	594.5	394.22
-0.16	13	1.85	0.77	1.0	1.0	2.0	2.5	3.0	0.102	0.079	2.8	2.11
-0.17	10	2.30	1.40	0.5	1.0	2.0	3.0	5.0	0.100	0.100	575.6	575.66
-0.18	9	2.78	1.79	1.0	1.0	2.0	5.0	5.0	0.111	0.074	0.6	0.567
-0.19	19	3.26	2.54	1.0	2.0	2.0	5.0	9.1	0.211	0.096	11.9	9.29
≤-0.20	156	3.28	2.94	1.0	2.0	2.0	4.0	10.0	0.103	0.024	113.4	59.14

This table reports statistics on the total number of marker buy and limit buy orders submitted. The statistics reported include the number of orders in each price category, the mean order size, the standard deviation of the order size, the 5th, 25th, 50th, 75th and 95th percentiles of the order size, the average fill-ratio and time-to-fill with corresponding standard errors. The fill-ratio is computed for each order as the fraction of the initial order size that is filled within two trading days from the initial submission. The time-to-fill is a weighted average of the number of minutes until the order is filled. The market order category includes all orders that were at least partially filled immediately. The limit orders are sorted into twenty categories depending on the difference between the order price and the best ask quote at the time the order was submitted. The last category $\leq -\$0.20$ includes all limit buy orders with prices at least 20 cents below the prevailing ask quote.

Table 2: **(continued) Descriptive Statistics on the Order Flow**

order price distance	N	Order size: Mean	S.D.	Percentiles 5th	25th	50th	75th	95th	Fill-ratio Mean	S.E.	Time-to-Fill Minutes Mean	S.E.
Market sell orders including partial-fills												
	1933	2.39	2.36	0.2	0.9	1.7	3.0	7.2	0.972	0.121	1.1	8.2
Limit sell orders by distance from ask quote $ (N=3330)												
0.01	156	2.60	2.09	0.4	1.0	2.0	3.8	5.7	0.654	0.465	37.5	337.6
0.02	306	2.24	1.71	0.5	1.0	2.0	3.0	5.0	0.607	0.477	41.2	349.1
0.03	329	2.21	2.11	0.5	1.0	2.0	2.6	5.0	0.523	0.492	23.1	139.0
0.04	369	2.30	1.83	0.4	1.0	2.0	3.0	5.0	0.466	0.487	54.8	423.4
0.05	809	2.29	2.29	0.3	1.0	2.0	3.0	5.0	0.548	0.488	57.3	423.5
0.06	105	2.19	1.34	0.5	1.0	2.0	2.8	5.0	0.385	0.482	20.9	58.7
0.07	134	2.28	1.73	0.4	1.0	2.0	3.0	5.0	0.357	0.476	38.0	138.7
0.08	126	2.23	2.28	0.5	1.0	2.0	2.5	5.0	0.331	0.464	82.6	417.0
0.09	98	2.50	2.68	0.7	1.0	2.0	2.2	5.0	0.390	0.482	61.9	217.6
0.10	341	2.13	1.57	0.4	1.0	2.0	3.0	5.0	0.283	0.445	101.4	669.2
0.11	23	1.69	1.35	0.3	0.8	1.0	2.0	4.3	0.348	0.487	19.6	52.2
0.12	32	1.58	1.22	0.5	0.7	1.0	2.0	4.9	0.312	0.471	187.5	572.2
0.13	33	1.66	1.03	0.4	1.0	1.5	2.0	3.9	0.212	0.415	11.4	33.5
0.14	32	2.86	2.83	0.5	1.0	1.8	5.0	5.9	0.281	0.457	67.3	209.8
0.15	130	1.87	1.52	0.2	1.0	1.4	2.0	5.0	0.292	0.457	203.7	878.4
0.16	6	2.30	1.88	0.2	0.6	2.0	4.0	5.0	0.500	0.548	12.3	16.9
0.17	11	2.51	1.58	1.0	1.1	2.0	3.9	5.0	0.045	0.151	25.6	84.9
0.18	8	2.15	2.00	0.2	1.0	1.0	3.5	6.0	0.250	0.463	52.4	135.4
0.19	21	1.96	1.62	0.3	0.6	1.1	3.0	5.0	0.238	0.436	439.5	1310.2
$\geq$0.20	261	1.84	1.96	0.2	0.6	1.0	2.0	5.0	0.149	0.356	297.1	1582.4

This table reports statistics on the total number of marker sell and limit sell orders submitted. The statistics reported include the number of orders in each price category, the mean order size, the standard deviation of the order size, the 5th, 25th, 50th, 75th and 95th percentiles of the order size, the average fill-ratio and time-to-fill with corresponding standard errors. The fill-ratio is computed for each order as the fraction of the initial order size that is filled within two trading days from the initial submission. The time-to-fill is a weighted average of the number of minutes until the order is filled. The market order category includes all orders that were at least partially filled immediately. The limit orders are sorted into twenty categories depending on the difference between the order price and the best bid quote at the time the order was submitted. The last category $\geq$ \$0.20 includes all limit sell orders with prices at least 20 cents above the prevailing bid quote.

Table 3: **Descriptive Statistics on the Order Books**

Item	Mean	S.D.	Percentiles				
			5th	25th	50th	75th	95th
Panel A: Order Book Quantity by Level—1,000 shares							
1st-bid-level	4.20	4.91	0.40	1.30	2.50	5.00	14.00
2nd-bid-level	5.59	6.05	0.90	2.00	3.40	7.00	17.90
3rd-bid-level	5.16	5.43	0.90	2.00	3.00	6.00	15.30
4th-bid-level	4.58	4.96	1.00	2.00	3.00	5.00	13.00
5th-bid-level	4.03	3.82	0.50	2.00	3.00	5.00	12.00
1st-ask-level	3.75	4.96	0.30	1.00	2.00	4.70	11.70
2nd-ask-level	5.08	5.28	0.60	2.00	3.40	6.60	14.50
3rd-ask-level	5.09	4.53	0.80	2.00	3.60	7.00	13.90
4th-ask-level	4.91	4.31	0.50	2.00	3.30	6.90	12.83
5th-ask-level	4.23	3.91	0.50	2.00	3.00	5.31	12.00
Panel B: Order Book Prices Relative to Mid-Quote $							
1st-bid-level	-0.02	0.02	-0.00	-0.02	-0.02	-0.03	-0.05
2nd-bid-level	-0.06	0.15	-0.02	-0.03	-0.05	-0.07	-0.10
3rd-bid-level	-0.11	0.27	-0.03	-0.05	-0.07	-0.12	-0.15
4th-bid-level	-0.16	0.38	-0.04	-0.06	-0.10	-0.17	-0.23
5th-bid-level	-0.22	0.46	-0.06	-0.08	-0.12	-0.22	-0.43
1st-ask-level	0.02	0.02	0.00	0.02	0.02	0.03	0.05
2nd-ask-level	0.05	0.11	0.02	0.03	0.05	0.08	0.10
3rd-ask-level	0.08	0.22	0.03	0.05	0.08	0.12	0.15
4th-ask-level	0.10	0.36	0.04	0.07	0.11	0.17	0.23
5th-ask-level	0.11	0.51	0.06	0.09	0.14	0.23	0.35

The first half of the table presents the mean, standard deviation, 5th, 25th, 50th, 75th, 95th percentiles for each order book variable. The first five rows include the total order book quantities at the best through the fifth best price levels in the order books. The next five rows provide the corresponding numbers for the ask side. The bottom half of the table presents the same set of statistics on the difference between the best through the fifth best bid or ask price level in the order book and the prevailing mid-quote.

Table 4: **Estimation Results for Cox Proportional Hazard Model.**

Variable	Coefficient	Std. Err.	p–value
Spread	-.2098942	0.0729606	0.004
Bid quantity	5.75e-06	7.86e-07	0.000
Ask quantity	2.02e-06	8.16e-07	0.013
Total bid quantity	6.51e-07	7.45e-08	0.000
Total ask quantity	9.67e-07	1.34e-07	0.000
Morning	.2710226	.0102266	0.000
Afternoon	-.0010992	.0119811	0.927
Volume	2.85e-06	5.10e-08	0.000
Volatility	1.216014	.0349104	0.000
LR $\chi^2(9)$	7055.21		
p–value	(0.0000)		

The conditioning variables are: Spread: dollar bid/ask spread, Bid quantity: quantity in the order book at the best bid, Ask quanity : quantity in the order book at the best ask, Total bid quantity: quantity in the order book at bid levels 2 through 5, Total ask quantity : quantity in the order book at ask levels 2 through 5, Morning: a dummy variable for time $\leq$ 8:00AM, Afternoon : a dummy variable for time $\geq$ 12:00 PM, Volume: Trading Voluume during last half hour, Volatility : the standard deviation of transaction prices over last half hour. The model is estimated using maximum likelihood.

Table 5: **Definitions of Factors**

CD	Canadian Dollar Futures Price
TB	US Treasury Bill Futures Price
GC	Gold Futures Price
SV	Silver Futures Price
CL	Crude Oil Futures Prices
SP	S&P 500 Index Futures
US	US 30 Year Treasury Bond Futures
price	Order submission prices

The data is obtained from the Futures Data Institute. Canadian dollar, US Treasury Bill, and S&P 500 Index futures are from the Chicago Mercantile Exchange. The US Treasury bond (30-year) futures are from the Chicago Board of Trade. The silver and gold futures are from the New York Mercantile Exchange.

Table 6: **Estimation of the Common Value**

	ADF	OLS	EC
Price	-1.6695		
CD	-2.0938	-16.514	-18.256
		(0.6406)	(4.4341)
CL	-1.6168	-0.213	-0.192
		(0.0047)	(0.0378)
GC	-1.9114	0.039	0.038
		(0.0010)	(0.0068)
SP	-2.8049	-0.010	-0.014
		(0.0008)	(0.0058)
SV	-2.4587	-0.035	-0.034
		(0.0006)	(0.0044)
TB	0.7705	-0.705	-0.992
		(0.0468)	(0.4408)
US	-2.9849	0.030	0.078
		(0.0069)	(0.0587)
R^2		0.7623	
Number of Observations	7980		
Start Date	90/05/28		
Cointegration test			-12.55

The data analyzed is three months of order submissions and changes following the specified start date. Factor prices are aligned to the times of the order submissions using the most recent available quote. ADF is Augmented Dickey-Fuller t–tests for a unit root in each series. All tests fail to reject the presence of a unit root at the 2.5% significance level (critical value -3.12) and most fail to reject at the 10% level (critical value -2.57). The lag order for the ADF test was set using the highest significant lag of the differences. OLS is the regression coefficients (standard errors in parenthesis) for the cointegrating regression estimated by ordinary least squares on demeaned variables. All coefficients are significant at the 1% level. The regression as a whole is also significant at the 1% level. EC is implied long term coefficients from the Error Correction Estimation of the model. Standard errors in parenthesis are calculated as suggested in Banerjee et al (1993). The error correction test finds significant cointegration at the 1% level. In order to remove the autocorrelation in the residuals, ten leads and ten lags of the factors are included in the error correction regression along with one lags of the price series.

Table 7: **Estimates of Conditional Fill Ratios, Picking Off Risk and Thresholds**

	Order Choice Categories							
	Sell Orders				Buy Orders			
Item	Mkt	Lim +1	Lim +2	Lim +3	Lim -3	Lim -2	Lim -1	Mkt
Avg. fill-ratio	1.0000	0.5477	0.3349	0.1750	0.1676	0.3265	0.6205	1.0000
	–	(0.0124)	(0.0182)	(0.0177)	(0.0204)	(0.0180)	(0.0127)	–
Picking off risk	0	0.0015	-0.0020	0.0016	0.0003	0.0012	0.0013	0
	–	(0.0007)	(0.0014)	(0.0016)	(0.0015)	(0.0015)	(0.0010)	–
Avg threshold	-0.0630	-0.0301	0.0698	0.1003	-0.0612	0.0040	0.0420	0.1399

This table reports statistics for the fill-ratios, the winner's curse terms and the thresholds with and without the common value adjustment. The first two rows present the average fill-ratio for the eight order categories we consider with the standard errors directly below. The next two rows present the average picking off risk terms and associated standard errors. On the last row we present the threshold valuations adjusted for the common value relative to the quotes. The threshold of $-0.063 for a market sell order implies that traders with a private valuation or demand for immediacy below $-0.063 would find a market sell order optimal.

Table 8: **Estimation Results for the Theoretical Model**

	S_0	S_1	L_0	L_1	C
Panel A: Base Specification					
Constant liquidity	2.6126		-5.1736		-0.1544
and arrivals	(0.0325)		(1.1091)		(0.0144)
Panel B: State variable, Order book depth					
Varying liquidity	2.4691	0.0254	-5.1732		-0.1558
distribution	(0.0467)	(0.0059)	(1.1106)		(0.0145)
Varying arrival	2.6113		-5.4058	0.0427	-0.1545
rates	(0.0325)		(1.5907)	(0.1262)	(0.0144)
Varying arrival	2.5295	0.0148	-5.4021	0.0421	-0.1537
and liquidity	(0.0468)	(0.0059)	(1.5998)	(0.1298)	(0.0144)
Panel C: State variable, Trading volume over last hour					
Varying liquidity	2.5243	0.0026	-5.1729		-0.1575
distribution	(0.0460)	(0.0010)	(1.1178)		(0.0146)
Varying arrival	2.6088		-5.5754	0.0165	-0.1526
rates	(0.0325)		(1.4185)	(0.0101)	(0.0143)
Varying arrival	2.6402	-0.0010	-5.5766	0.0166	-0.1528
and liquidity	(0.0454)	(0.0010)	(1.4206)	(0.0102)	(0.0143)
Panel D: State variable, Common value volatility					
Varying liquidity	2.5900	0.1102	-5.1735		-0.1546
distribution	(0.0349)	(0.0618)	(1.1105)		(0.0145)
Varying arrival	2.6122		-5.1662	-0.0332	-0.1543
rates	(0.0325)		(1.1553)	(1.2445)	(0.0145)
Varying arrival	2.5876	0.1189	-5.1655	-0.0359	-0.1546
and liquidity	(0.0350)	(0.0623)	(1.1548)	(1.2737)	(0.0145)

This table reports estimation results for six specifications of the model. Panel A reports estimates using no state variables, in Panel B the state variable equals to the the depth in the order book as measured by the total quantity of orders at the 5 best prices in the order book , in Panel C the state variable is the trading volume over the last and in Panel D the state variable is the volatility of common value over the last hour. For each specification, we report estimates allowing the state variable to enter the liquidity distribution and the arrival rates of the traders, both individually and jointly. The parameter estimates are reported along with corresponding standard errors reported in parentheses.

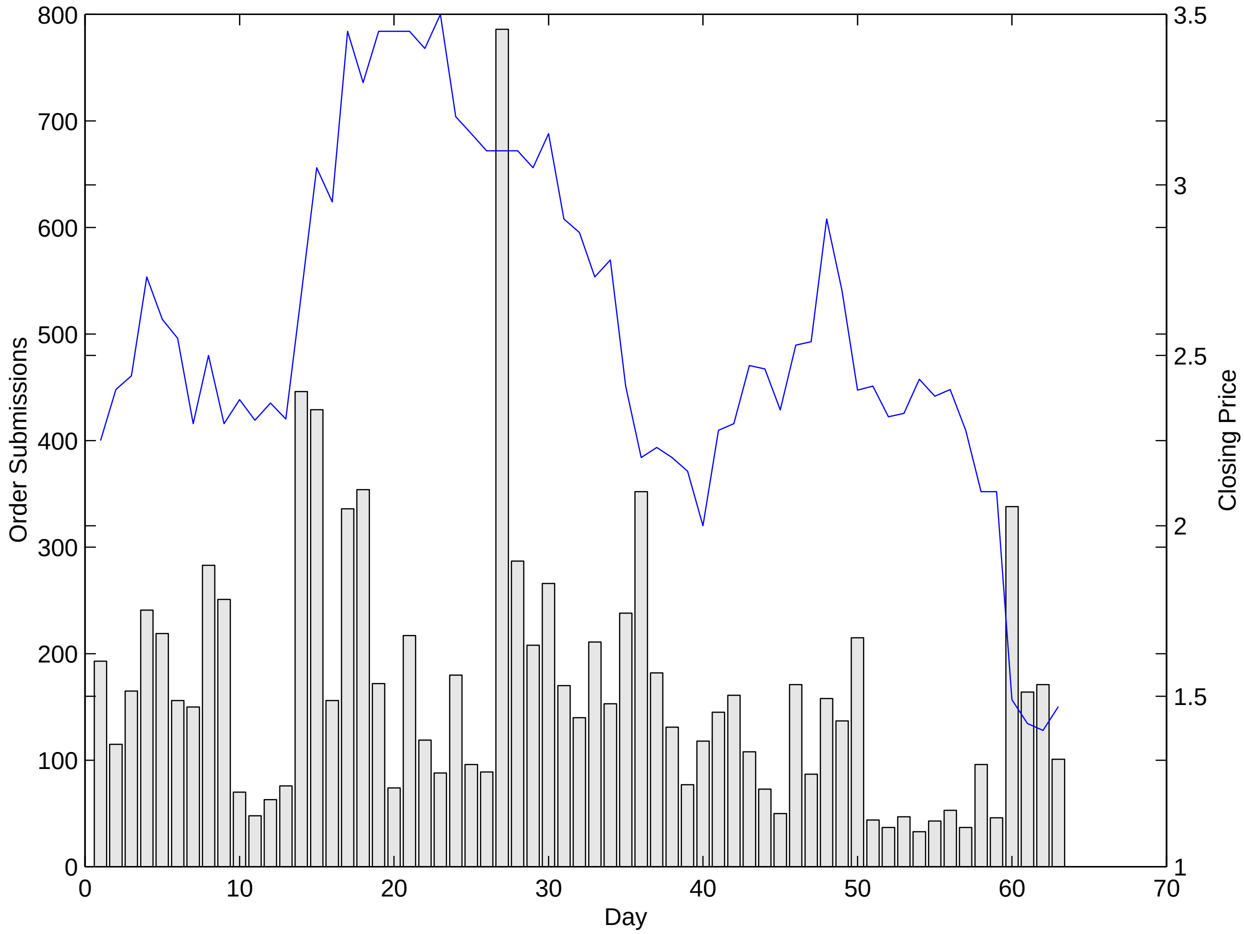

Figure 1: The solid line plots the daily closing price (right-hand scale) and the bars represent the daily number of order submissions (left-hand scale) over the 63 trading days.

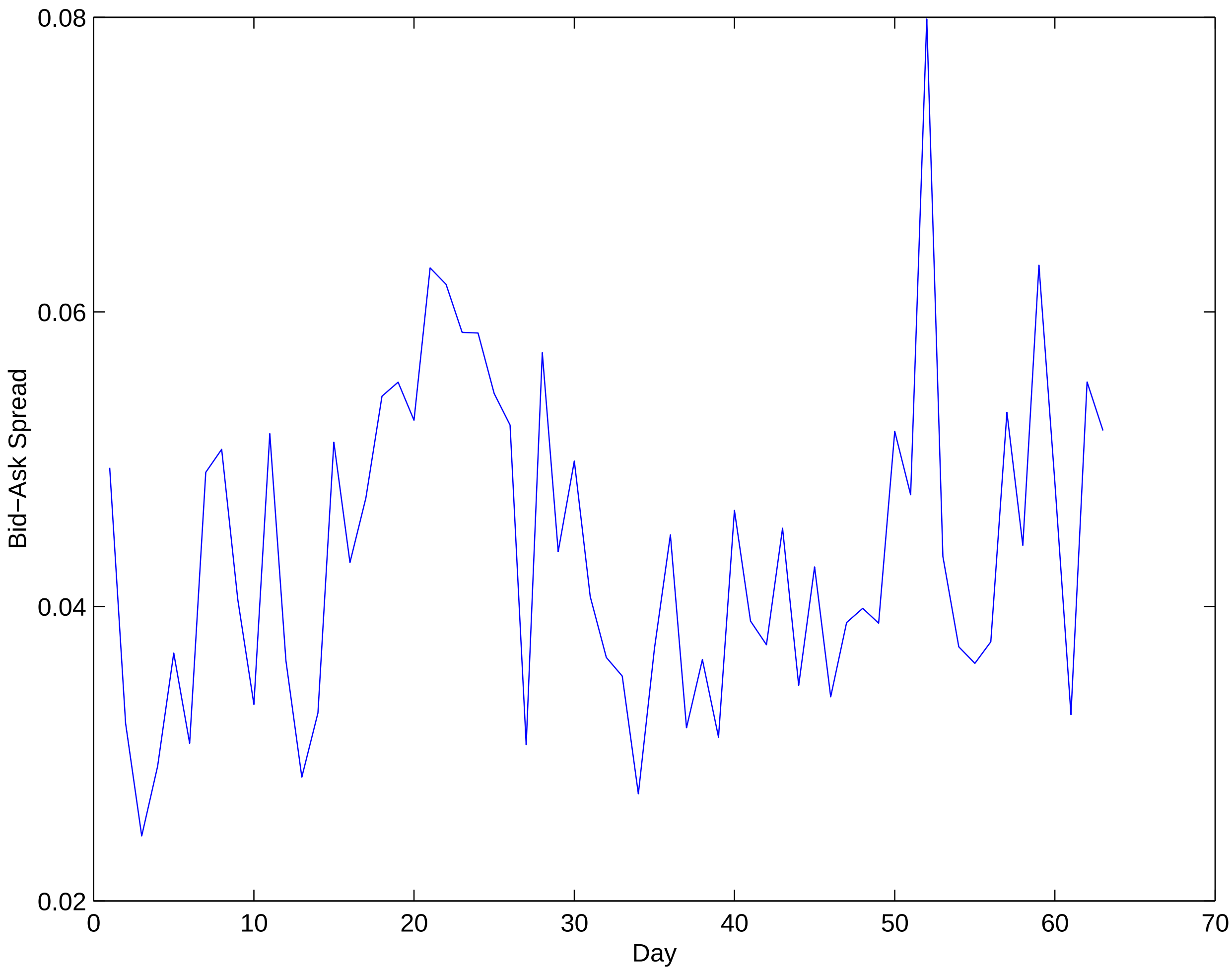

Figure 2: The daily average bid-ask spread (measured in dollars) over 63 trading days.

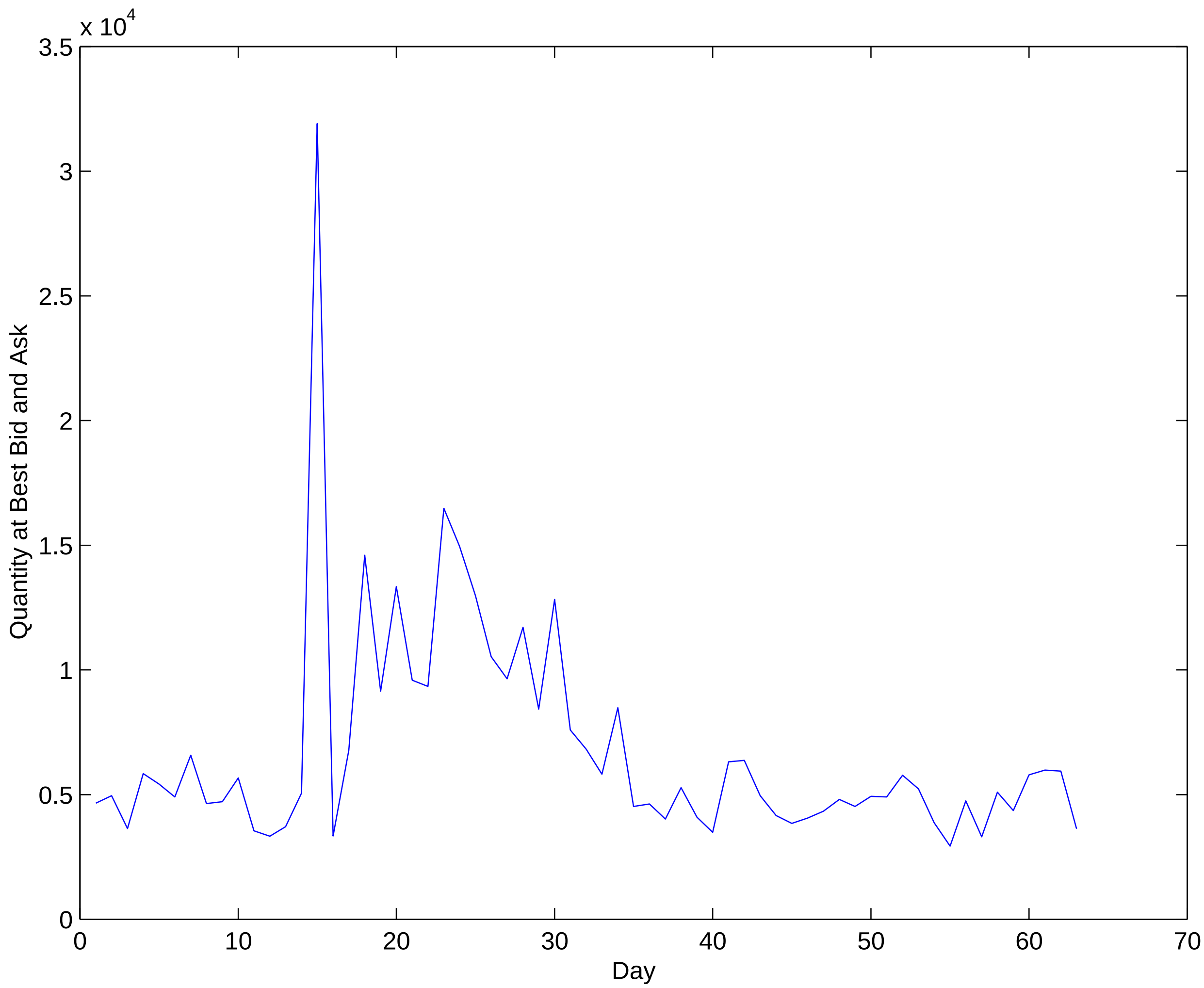

Figure 3: The daily average sum of the bid and the ask depth (number of shares) over 63 trading days.

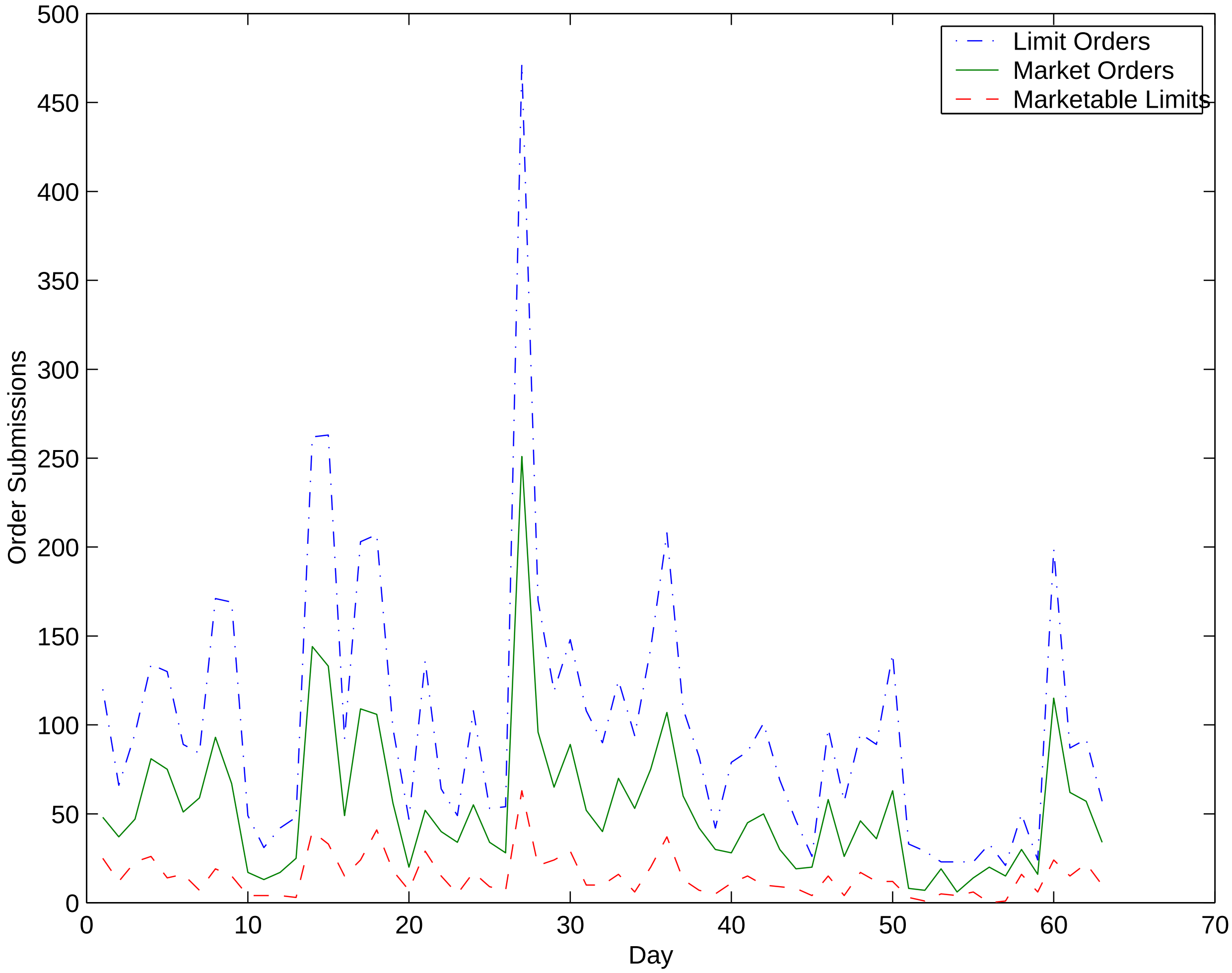

Figure 4: Daily number of order submitted by type: market orders (solid), limit orders (dash-dot) and marketable limit orders (dash). Marketable limit orders have a buy or sell price that matches or exceeds the prevailing ask or bid quote, respectively, and thus they will typically be filled partially or in full immediately.

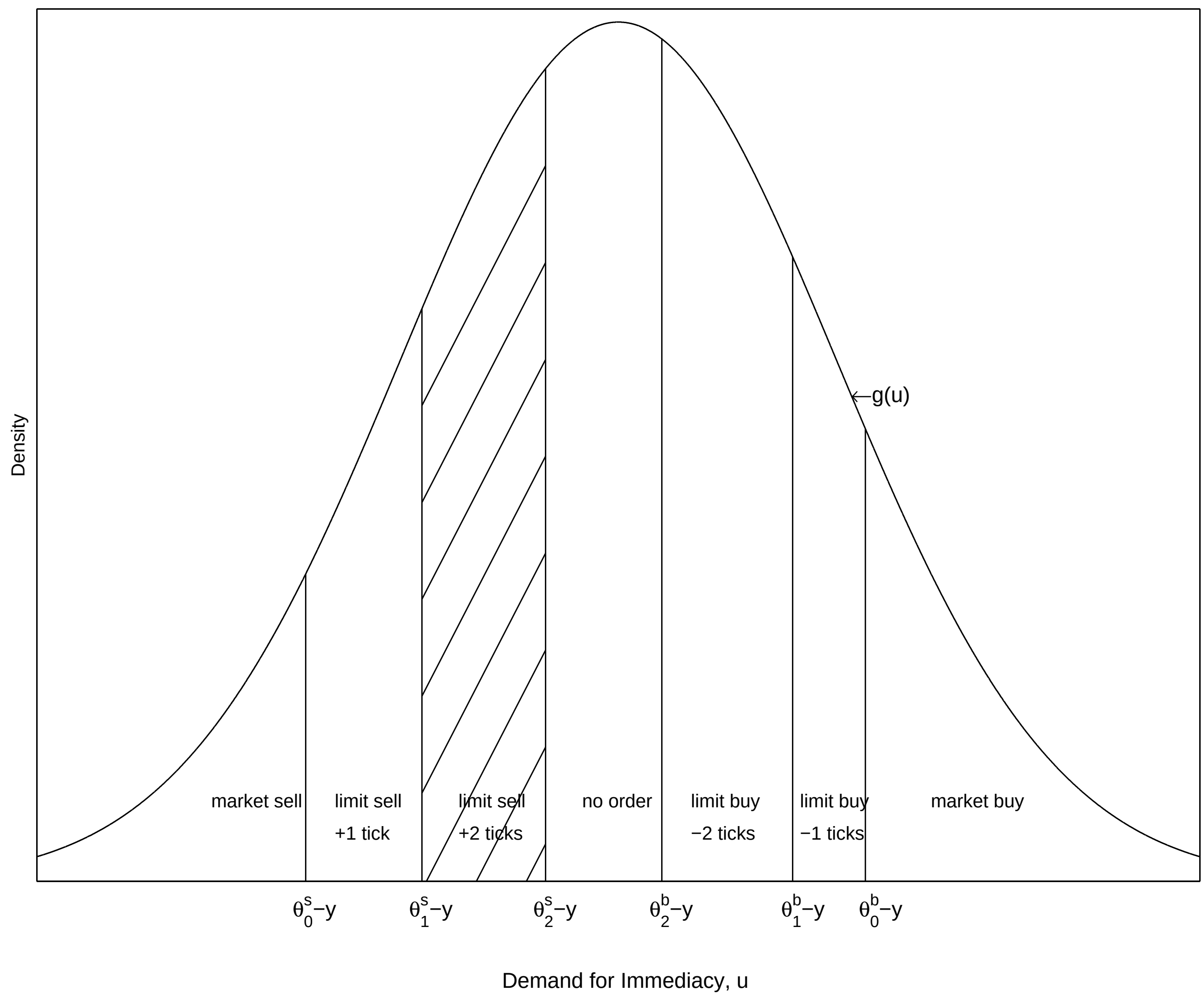

Figure 5: The graph illustrates how the probabilities of observing different order choices are determined in our model. There are two types of limit orders that a trader would consider submitting (K=L=2). The threshold valuations θ_i^s and θ_i^b $i = 0, 1, 2$ are determined as outlined in Proposition 1. Then the probability of observing a given order choice *conditional* on a trader arriving is given by the area under the probability distribution function between two adjacent threshold valuations. For example, the probability of observing a limit sell order at two ticks from the best bid quote is given by the probability that the trader has a demand for immediacy u that falls in the interval $[\theta_1^s - y, \theta_2^s - y)$, which corresponds to the area with diagonal lines in the graph.