

What Pieces of Limit Order Book Information Matter in Explaining Order Choice by Patient and Impatient Traders?

Veredas, David; Pascual, Roberto

Postprint / Postprint

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

www.peerproject.eu

Empfohlene Zitierung / Suggested Citation:

Veredas, D., & Pascual, R. (2009). What Pieces of Limit Order Book Information Matter in Explaining Order Choice by Patient and Impatient Traders? *Quantitative Finance*, 9(5), 527-545. <https://doi.org/10.1080/14697680802616704>

Nutzungsbedingungen:

Dieser Text wird unter dem "PEER Licence Agreement zur Verfügung" gestellt. Nähere Auskünfte zum PEER-Projekt finden Sie hier: <http://www.peerproject.eu>. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

gesis
Leibniz-Institut
für Sozialwissenschaften

Terms of use:

This document is made available under the "PEER Licence Agreement". For more information regarding the PEER-project see: <http://www.peerproject.eu>. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Mitglied der

Leibniz-Gemeinschaft

What Pieces of Limit Order Book Information Matter in Explaining Order Choice by Patient and Impatient Traders?

Journal:	<i>Quantitative Finance</i>
Manuscript ID:	RQUF-2007-0012.R1
Manuscript Category:	Research Paper
Date Submitted by the Author:	23-May-2008
Complete List of Authors:	Veredas, David; Universite Libre de Bruxelles, ECARES Pascual, Roberto; Universidad de las Islas Baleares, Business
Keywords:	Limit Orders, Market orders, Limit order book, Order aggressiveness, Order-driven markets
JEL Code:	G1 - General Financial Markets < G - Financial Economics

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

What Pieces of Limit Order Book Information Matter in Explaining Order Choice by Patient and Impatient Traders?*

Roberto Pascual
Department of Business
Universidad de las Islas Baleares
rpascual@uib.es

David Veredas
ECARES, Université Libre de Bruxelles
and
CORE, Université catholique de Louvain
dveredas@ulb.ac.be

This version: May 2008
(First version: December 2003)

* We wish to thank David Abad, Sonia Falconieri, Pierre Giot, Angelo Ranaldo, Andreu Sansó, José Vidal, and Klaus F. Zimmermann, and two anonymous referees for their suggestions on previous versions of this paper. We acknowledge the insightful comments of the participants at seminars at Universidad Carlos III de Madrid (Spain), Université Catholique de Louvain (Belgium), and Bolsa de Madrid (Spain), and participants at the Namur's Workshop on Automated Auction Markets (Belgium). This work was partially supported by the European Commission Project HPRN-CT-2002-00232 "MICFINMA: Microstructure of Financial Markets in Europe". Roberto Pascual also acknowledges the financial sponsorship of the Fulbright Grant and the Spanish Ministry of Education, Culture and Sports, and of the Spanish DGICYT projects BEC2001-2552-C03-03, SEJ2004-07530-C04-04, and SEJ2007-67895-C04-03. The first version of this paper was completed while Roberto Pascual was a Visiting Scholar at the New York University Salomon Center. David Veredas is also a member of ECORE, the recently created association between CORE and ECARES. Corresponding author: Roberto Pascual, Departamento de Economía de la Empresa, Universidad de las Islas Baleares, Ctra. de Valldemossa Km. 7.5, 07122 Palma de Mallorca, Baleares, Spain. Phone: +34-971-17-13-29, Fax: +34-971-17-23-89, e-mail: rpascual@uib.es. The scientific responsibility is assumed by the authors.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

What Pieces of Limit Order Book Information Matter in Explaining Order Choice by Patient and Impatient Traders?

ABSTRACT

In this paper, we extend the existing empirical evidence on the relationship between the state of the limit order book (LOB) and order choice. Our contribution is twofold: first, we propose a sequential ordered probit (SOP) model which allows studying patient and impatient traders' choices separately; second, we consider two pieces of LOB information, the best quotes and the book beyond the best quotes. We find that both pieces of LOB information explain the degree of patience of an incoming trader and, afterwards, its order choice. Nonetheless, the best quotes concentrate most of the explanatory power of the LOB. The shape of the book beyond the best quotes is crucial in explaining the aggressiveness of patient (limit order) traders, while impatient (market order) traders base their decisions primarily on the best quotes. Patient traders' choices depend more on the state of the LOB on the same side of the market, while impatient traders mostly look at the state of the LOB on the opposite side. The aggressiveness of both types of traders augments with the inside spread. However, patient (impatient) traders submit more (less) aggressive limit (market) orders when the depth of the own (opposite) best quote and the length of the own (opposite) side of the book increase. We also find that higher depth away from the best ask (bid) quote may signal that this quote is "too low (high)", causing incoming impatient buyers (sellers) to be more aggressive and incoming patient sellers (buyers) to be more conservative.

Key words: Limit orders, market orders, limit order book, order aggressiveness, order-driven markets.

1. Introduction

In order driven markets without designated market makers, liquidity relies on the submission of limit orders. Limit orders are stored on an electronic register, called the (open) limit order book (henceforth, LOB), to wait for execution. To appreciate how and why liquidity varies over time in these venues, it is first necessary to comprehend the forces leading an incoming investor to decide between submitting a limit order (liquidity supply) or a market order (liquidity demand). Some recent studies illustrate that informed and uninformed investors order choices also happen to be essential to understand how new information is incorporated into prices (e.g., Harris, 1998, Anand et al., 2005, Bloomfield et al., 2005). It has also been shown that order choice determines the speed and costs (quality) of execution (e.g., Harris and Hasbrouck, 1996; Lo et al., 2002, and Al-Suhaibani and Kryzanowski, 2000). Finally, the growing interest in the optimization of order submissions, an underlying foundation of algorithm trading (see O'Hara, 2007), also corroborates the relevance of this topic in microstructure research.

Quite a few recent theoretical studies have evaluated the role the LOB plays in order choice. In the dynamic models of Parlour (1998), Foucault (1999), Handa et al. (2003), Foucault et al. (2005), Goettler et al. (2005), and Rosu (2005), LOB dimensions, such as the quoted depth on both sides of the LOB and the inside spread, determine the choice between market and limit orders. This theoretical interaction between the state of the LOB and the traders' order choices has been extensively corroborated by the empirical literature (e.g., Biais et al., 1995; Griffiths et al., 2000; Al-Suhaibani and Kryzanowski, 2000; Rinaldo, 2004; Beber and Caglio, 2005, and Ellul et al., 2007).¹

¹ All this theoretical and empirical literature is revised in detail in the next section.

1
2
3 This paper is aimed to contribute to this literature. Our approach differs from
4 previous empirical studies in two main aspects. First, we propose a two-stage *sequential*
5 ordered probit (henceforth SOP) model, instead of the more commonly used ordered
6 probit model (e.g., Griffiths et al., 2000; Rinaldo, 2004), to model order aggressiveness.
7
8 The SOP model allows studying patient and impatient order choices separately. In the
9 first stage of our model, the trader defines herself as patient or eager, as she must choose
10 between removing, providing and consuming liquidity, without specifying a particular
11 type or order. In most of the previously cited theoretical models on order choice, the
12 degree of patience of an incoming trader is exogenously given (by factors such as her
13 degree of risk aversion, a higher or lower discount rate, her waiting costs, her liquidity
14 needs, or her access to perishable non-publicly known information). In equilibrium,
15 patient traders tend to submit limit orders while impatient traders tend to submit market
16 orders (e.g., Foucault et al., 2005, and Rosu, 2005). However, these models also
17 postulate that the state of the LOB may have an influence in reshaping their degree of
18 patience, leading to unanticipated order choices.² In our empirical analysis, we identify
19 patient and impatient traders based on posterior order choices. In this first stage of our
20 SOP model, we therefore evaluate the importance of the LOB in determining the
21 character of an incoming trader.
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45

46 In the second stage of the SOP model, the trader chooses among the different types
47 of orders available given their predetermined disposition. Thus, a patient trader chooses
48 among the different types of limit orders, either away from, at, or within the best quotes.
49
50 In this decision, the patient trader faces a direct tradeoff between a more favorable limit
51 price and a higher risk of non-execution. The impatient trader, though, has to fix the
52
53
54
55
56
57
58

59 ² Recent empirical and theoretical research (see Kaniel and Liu, 2006; Harris, 1998; Bloomfield et al.,
60 2005, and Anand et al., 2005) has shown that, under certain circumstances, traders that are presumed to
be more inclined to be impatient, such as informed traders, may choose to submit orders that a priori are
more appropriate for patient traders.

1
2
3 relative size of her order. She must choose between a market order that requests less
4 than available, all that is available, or more volume than available at the best opposite
5 quote. This decision determines the instantaneous price impact of the trade and the
6 average immediacy cost per share. Hence, in this second stage, we evaluate the
7 relevance of the LOB in the decision of choosing a particular type of order by each
8 particular type of trader.
9

10
11
12
13
14
15
16
17
18 Second, we contribute to the literature by providing empirical evidence on how the
19 state of the LOB beyond the best quotes affects order choices. We understand the LOB
20 as a set of two main pieces of information, the best quotes and the book beyond the best
21 quotes. This distinction is not arbitrary, since even some of the most pre-trade opaque
22 markets provide information about the best quotes. Additionally, most of the above-
23 cited theoretical papers focus on modeling order choice conditional on the best quotes.
24 The exception is Goettler et al. (2005). Using simulations, they postulate different
25 effects of both pieces of LOB information on order choice by both patient and impatient
26 traders. In this paper, we provide empirical evidence confirming many of their
27 simulation findings.³ Moreover, previous empirical papers about order aggressiveness
28 only consider the best quotes of the LOB. Relevant exceptions are Irvine et al. (2000)
29 and Cao et al. (2003). Irvine et al. (2000) show that a liquidity measure based on LOB
30 data is more informative about subsequent order flow than the traditional measures
31 based on the best quotes. Cao et al. (2003) examine the informativeness of the LOB
32 beyond its first level. They conclude that the best quotes lead price discovery and
33 provide a better estimator of the true value, but the information derived from the
34 secondary level of the book adds some “marginal” explanatory power on future returns.
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

³ Rosu (2005) also considers the whole LOB in his theoretical model. In this case, however, there are no clear cut predictions regarding best quotes versus the book beyond the best quotes.

1
2
3 Although our paper inevitably overlaps with Cao et al.'s (2003) paper at some point, the
4 motivations and methods differ since they focus on price discovery.
5
6
7

8
9 We use high frequency LOB data on the Stock Exchange Interconnection System of
10 the Spanish Stock Exchange (henceforth SSE). In this market, the LOB is available in
11 real time to all market participants and market orders are not restricted to the best
12 quotes. Hence, the shape of the entire LOB, and not only the best quotes, determines the
13 immediacy costs of a market order, as well as the expected time-to-execution of a limit
14 order.
15
16
17
18
19
20
21
22

23 Our main findings are as follows. Consistently with previous empirical studies, we
24 find that the state of the LOB matters in explaining the aggressiveness of incoming
25 traders. Additionally, we show that both the best quotes and the book beyond the best
26 quotes help to shape the degree of patience of incoming buyers and sellers, though most
27 of this effect is due to the best quotes. When we look at patient and impatient traders
28 separately, we report asymmetric patterns. Thus, incoming patient traders strongly rely
29 on the state of the book beyond the best quotes, while impatient traders base their
30 strategic decisions primarily on the prevailing best quotes. Both patient and impatient
31 traders condition the placement of their limit orders on the inside spread. However,
32 patient (impatient) traders' choices rely on the depth at and away from the best quote
33 and the length of the book on the same (opposite) side of the market. Consistently with
34 Goettler et al. (2005), our evidence shows that higher depth away from the best ask
35 (bid) quote may signal that this quote is "too low (high)", causing incoming impatient
36 buyers (sellers) to place market orders more aggressively and incoming patient sellers
37 (buyers) to place limit orders more conservatively. Finally, we show that our SOP
38 model provides better in sample and out of sample goodness-of-fit performance than the
39 traditional ordered probit model of order aggressiveness.
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 The paper proceeds as follows. In the next section, we briefly review the pertinent
4 literature. In section 3, we introduce the two-stage SOP model and discuss several
5 methodological details. In section 4, we describe the database, the sample, and the
6 market. In section 5, we present our empirical findings. Finally, in section 6, we
7 conclude.
8
9

10 11 12 13 14 15 16 **2. Literature review** 17

18
19 Cohen et al. (1981), Copeland and Galai (1983), and Handa and Schwartz (1996) set
20 the basic rationale for limit order trading as a trade-off between the non-execution costs
21 and the picking-off risk carried by limit orders, and the immediacy costs supported by
22 market orders. Other things being equal, the lower the non-execution risk and the higher
23 the immediacy costs, the more likely the incoming trader is to submit a limit order.
24 Moreover, an increase in adverse selection costs aggravates the winner's curse problem
25 faced by uninformed limit order traders.
26
27
28
29
30
31
32
33
34
35

36 Recent theoretical developments suggest that the state of the LOB may be critical to
37 order choice. Parlour (1998) proposes a dynamic equilibrium model in which the
38 thickness of a given side of the LOB produces a "crowding out" effect on the limit
39 orders on that side, because of the long expected time-to-execution. In contrast, traders
40 on the opposite side, anticipating the crowding out effect, are better submitting less
41 aggressive orders.
42
43
44
45
46
47
48
49
50

51 Foucault (1999) develops a dynamic limit order market in which the mix between
52 market and limit orders in equilibrium depends on the fundamental volatility of the
53 stock. When volatility increases, the risk of being picked-off by an informed trader
54 increases. Limit order traders respond posting higher ask prices and lower bid prices.
55
56
57
58
59
60

1
2
3 This makes market orders less attractive and, consequently, the proportion of limit
4 orders on the total order flow increases.
5
6

7
8 Foucault et al. (2005) develop a dynamic model for an order driven market populated
9 by strategic liquidity traders that differ in their waiting costs. In this model, order choice
10 is determined by the bid-ask spread, but not by quoted depth.⁴ For certain spread levels,
11 patient (low waiting costs) traders tend to place limit orders, while impatient (high
12 liquidity costs traders) tend to submit market orders. As the spread increases and
13 surpasses a certain cutoff level, however, all traders tend to choose limit orders to trade.
14
15 Moreover, liquidity suppliers are willing to offer larger spread improvements when the
16 spread is large.
17
18

19
20 Rosu (2005) proposes a continuous time version of the latter model where
21 competition between limit order traders generates endogenous undercutting. The whole
22 shape of the LOB is considered, not only the best quotes. By allowing market orders of
23 different sizes with probabilities which do not decrease too fast with order size, the
24 author shows that limit orders may cluster away from the best quotes. Moreover, traders
25 are allowed to dynamically cancel and resubmit their orders as the shape of the LOB
26 changes over time. When the LOB is full, “fleeting” orders may happen: a patient trader
27 submits a limit order inside the quotes and an opposite patient trader accepts it
28 immediately by canceling her limit order and placing a market order.
29
30

31
32 In Goettler et al. (2005), liquidity motivated traders with differing degree of patience
33 choose whether to buy or sell (or both), either using limit orders or market orders (or
34 both), given the state of the LOB and a consensus true value. Traders who are eager to
35 buy (sell) are allowed to submit limit orders above (below) the consensus value, which
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

⁴ This arises because limit order submitters are not allowed to submit orders at or away from the best quotes.

1
2
3 means that market orders may result in negative transaction costs. Because the
4 consensus value can change, limit orders suffer from picking-off risk. All else equal, the
5
6 higher the probability of execution and the lower the picking-off risk at some price, the
7
8 higher the number of limit orders to sell (buy) submitted at that price. Using
9
10 simulations, they find that, when the spread is wide, a thicker best bid increases
11
12 (decreases) the frequency of inside-the-quotes limit (market) orders to buy.
13
14 Additionally, higher depth above (below) the best ask (bid) quote may signal that the
15
16 best quote may be “too low (high)” resulting in a more aggressive (conservative) order
17
18 placement by incoming buyers. Symmetric patterns are predicted for sellers’ choices.
19
20
21
22
23
24

25 In the previous models, asymmetric information is not allowed. Handa et al. (2003)
26 extend Foucault’s (1999) model to allow for private information and a varying
27
28 proportion of buyers and sellers. In this model, the thickness of the bid and ask sides of
29
30 the LOB proxy for the proportion of high and low-valuation traders. Buy (sell)
31
32 competition is expected to increase with the proportion of high (low) valuation traders,
33
34 prolonging the expected time-to-execution of limit buy (sell) orders and thus making
35
36 more attractive the use of buy (sell) market orders.
37
38
39
40
41

42 Regarding empirical research, several papers have shown that limit (market) order
43
44 traders enter in the market when liquidity is scarce (plentiful). Thus, Biais et al. (1995),
45
46 Al-Suhaibani and Kryzanowski (2000), Beber and Caglio (2005), and Ellul et al.
47
48 (2007), among others, find that limit order submissions inside the quotes occur more
49
50 frequently when the spread widens, while market order submissions are more likely
51
52 when the spread is tight. Consistently, Ranaldo (2004) and Hall and Hautsch (2007)
53
54 show that order aggressiveness and trading intensity both decrease with the bid ask
55
56 spread. Similarly, Ahn et al. (2001) show that when transitory volatility, due to a
57
58
59
60

1
2
3 paucity of limit orders, arises from the ask (bid) side, investors submit more limit orders
4
5
6 to sell (buy) than market orders to sell (buy).
7

8
9 Quoted depth on both sides of the market has also been found to matter. Griffiths et
10
11 al. (2000) and Rinaldo (2004) find that traders become more aggressive when their own
12
13 (opposite) side of the book is thicker (thinner). Chakrabarty et al. (2006) conclude that
14
15 investors aimed to reduce the expected time-to-execution of their limit orders should
16
17 submit them when the depth on the own (opposite) side is smaller (greater). Likewise,
18
19 Ellul et al. (2007) document that when quoted depth is large traders are more likely to
20
21 “jump the queue” by submitting limit orders with prices improving existing quotes and
22
23 less likely to submit limit orders with prices equal or worse than current quotes.
24
25 Additionally, positive (negative) bid-ask depth imbalances may anticipate short-term
26
27 price increases (decreases), since they attract additional limit orders on the deep side
28
29 and additional market orders on the thin side, rising the imbalance.
30
31
32
33

34
35 Recent studies show that order choice is a more complex phenomenon than the
36
37 simple dichotomy of the uninformed patient trader who submits a limit order and
38
39 passively waits for execution and the impatient trader who submits a market order and
40
41 demands immediate execution. Beber and Caglio (2005) provide evidence that order
42
43 choice and its interaction with the state of the LOB vary in periods characterized by
44
45 higher probability of informed trading. Their findings suggest that informed traders may
46
47 act strategically during these periods, using passive limit orders to hide their
48
49 information. Hasbrouck and Saar (2007) evidence that over one-third of the limit orders
50
51 submitted in INET are cancelled within two seconds. These “fleeting” limit order
52
53 traders are not meant to be patient liquidity providers. Pardo and Pascual (2007) and De
54
55 Winnie and D’hondt (2007) both show that the detection of hidden volume at the best
56
57 quotes temporarily increases order aggressiveness on the opposite side of the market.
58
59
60

1
2
3 A related line of research suggests that the state of the LOB may have informational
4 content. Seppi (1997), Harris (1998), Moinas (2004), and Kaniel and Liu (2006) provide
5 conditions under which informed traders, usually modeled as impatient, may be willing
6 to use limit orders. Bloomfield et al. (2005) and Anand et al. (2005) provide
7 experimental and empirical evidence supporting these last models' predictions. More
8 recently, Foucault et al. (2006) develop a model in which limit order traders possess
9 private information on the occurrence of future information events ("volatility
10 information"), such as corporate announcements, and they adjust their order
11 submissions to the level of risk perceived. Pascual and Veredas (2006) provide
12 supporting evidence.
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27

28 **3. Methodological details**

31 *3.1. The two-stage sequential ordered probit model*

32
33 As previously indicated, we understand order choice as a two stage process. In the
34 first stage, the trader is characterized as patient or impatient, as she must choose
35 between removing, providing, and consuming liquidity. The degree of patience of the
36 incoming trader may be partially exogenous, but we allow it to be shaped by the state of
37 different pieces of LOB information. In the second stage, the patient or impatient trader
38 chooses among the different type of orders available according to their needs. This
39 second step decision is essentially determined by the transient market conditions that, in
40 the end, are nested in the state of the LOB.
41
42
43
44
45
46
47
48
49
50
51
52

53 An appropriate econometric model to analyze the sequential decision process
54 described above is the Sequential Ordered Probit (henceforth SOP) model. In the first
55 stage of the process, the degree of impatience of a given trader i arriving at the market is
56
57
58
59
60

a (linear) function of K observed variables, X_k , $k = 1, \dots, K$. The aggressiveness index A_i^* is represented as

$$A_i^* = \sum_{k=1}^K \beta_k^A X_{ik}^A + \varepsilon_i^A = Z_i^A + \varepsilon_i^A, \quad [1]$$

where ε_i is the error term and β_k^A is the slope coefficient associated with the k^{th} regressor (to be defined later). Since A_i^* is difficult (if not impossible) to observe, equation [1] is a latent regression. However, it is possible to infer about the degree of the traders' impatience by observing the specific orders submitted. Thus, the three categories of order aggressiveness considered in this first stage (liquidity withdrawal, liquidity provision, and liquidity consumption) represent a partition of the state space that allows mapping the latent variable into observable discrete values. Let A_i be an ordinal response variable such that

$$A_i = \begin{cases} 1 & \text{if } A_i^* \leq \delta_1^A \\ 2 & \text{if } \delta_1^A < A_i^* \leq \delta_2^A, \\ 3 & \text{if } A_i^* > \delta_2^A \end{cases} \quad [2]$$

with δ_1^A and δ_2^A being unknown thresholds, to be estimated along with the slope parameters. $A_i = 1$ means that the trader decides to leave the market and, accordingly, she withdraws liquidity by submitting a cancellation; $A_i = 2$ means the trader is willing to assume certain degree of non-execution risk and, consequently, she will latterly provide liquidity by submitting a limit order; finally, $A_i = 3$ means the trader is impatient and requires an immediate execution of their trading interests; at the end, she will therefore choose among the different types of market orders.

The probability that the incoming trader is of type j is

$$P_{ij}^A = P(\delta_{j-1}^A < A_i^* < \delta_j^A) = P(\delta_j^A - Z_i^A) - P(\delta_{j-1}^A - Z_i^A), \quad [3]$$

with $\delta_0^A = -\infty$ and $\delta_3^A = +\infty$. Assuming that the probability distribution of the error terms ε_i is normal, $\Pr(\delta_j^A - Z_i^A) = \Phi(\delta_j^A - Z_i^A)$, where $\Phi(\cdot)$ is the normal cumulative distribution function.⁵ The corresponding log-likelihood is

$$\ln L^A = \sum_{i=1}^n \sum_{j=1}^3 d_{ij}^A \ln P_{ij}^A, \quad [4]$$

where $d_{ij}^A = 1$ if $A_i = j$ and $d_{ij}^A = 0$ otherwise, maximized with respect to β_k^A , δ_1^A and δ_2^A .

In the second stage of the SOP model, we analyze order choice by patient and impatient traders separately. Suppose, for instance, that trader i is patient.⁶ The ordinal response variable L_i takes three possible values, given by the unobservable thresholds δ_1^L and δ_2^L ,

$$L_i = \begin{cases} 1 & \text{if } L_i^* \leq \delta_1^L \\ 2 & \text{if } \delta_1^L < L_i^* \leq \delta_2^L, \\ 3 & \text{if } L_i^* > \delta_2^L \end{cases}$$

where L_i^* represents the latent aggressiveness of the limit order trader,

$$L_i^* = \sum_{k=1}^K \beta_k^L X_{ik}^L + \varepsilon_i^L = Z_i^L + \varepsilon_i^L.$$

Whenever $L_i^* \leq \delta_1^L$ the trader submits a non-aggressive limit order to buy (sell) below (above) the best bid (ask), and $L_i = 1$; whenever $\delta_1^L < L_i^* \leq \delta_2^L$, the trader submits a limit

⁵ If we assume a logistic distribution, equations [1] and [2] will define an ordered *logit* model. Preliminary estimations provide identical results under either assumption. Since related studies such as Griffiths et al. (2000) and Rinaldo (2004) assume normality, we also base our analysis on the ordered probit model.

⁶ The following discussion is equivalent for an impatient trader that has to choose between three types of market orders, from lower to higher aggressiveness: a non-aggressive market order, a market-to-limit order, and an aggressive market order. In terms of notation, we would replace L by M .

order to buy (sell) at the current best bid (ask), and $L_i = 2$; if $L_i^* > \delta_2^L$, the trader seeks a short time-to-execution and submits a limit order inside the best quotes, and $L_i = 3$.

The probability that an incoming trader submits a, say, limit order at the quotes is $P_{i2}^A P_{i2}^L$ where P_{i2}^A is the probability of a patient trader arriving at the first stage ($A_i = 2$) and P_{i2}^L is the unconditional probability of the patient trader submitting a limit order at the quotes ($L_i = 2$) at the second stage. Let $d_{il}^L = 1$ if $L_i = l$ and $d_{il}^L = 0$ otherwise, for $l = 1, 2, 3$. Because P_{i2}^A does not depend on l and $\sum_{l=1}^3 d_{il}^L = 1$, the log-likelihood of the second stage for the patient traders is

$$\ln L^L = \sum_{i=1}^n \ln P_{i2}^A + \sum_{i=1}^n \sum_{l=1}^3 d_{il}^L P_{il}^L, \quad [5]$$

The first term on the RHS of [5] does not play any role in the maximization of the log-likelihood, since the latter is maximized with respect to the parameters on the second RHS term (β_k^L , δ_1^L and δ_2^L).

The SOP approach is conceptually different to the OP approach. The latter presumes that any trader arriving at the market, independently of their degree of patience, chooses among all the possible type of orders. That is,

$$O_i = \begin{cases} 1 & \text{if } O_i^* \leq \delta_1^O \\ 2 & \text{if } \delta_1^O < O_i^* \leq \delta_2^O, \\ \vdots & \\ 7 & \text{if } O_i^* > \delta_7^O \end{cases}$$

where $O_i^* = \sum_{k=1}^K \beta_k X_{ik} + \varepsilon_i^O = Z_i + \varepsilon_i^O$ and $P_{ik}^O = \Pr(O_i = k) = \Phi(\delta_k^O - Z_i) - \Phi(\delta_{k-1}^O - Z_i)$.

In this case, $O_i = 1$ is a cancellation, $O_i = 2$ is a non-aggressive limit orders, and so on and so forth. In contrast to the SOP model, an increase in X_{ik} , for example, the bid-ask

1
2
3 spread, will be predicted to affect in exactly the same manner to all types of traders
4
5 through a single parameter. With the SOP model however, an increase in the bid-ask
6
7 spread may affect differently the order choice of patient and impatient traders through
8
9 one parameter for the first stage and two parameters for the second stage.
10
11

12
13 The extra computational demand of the SOP model is marginal since the
14
15 optimization of the SOP model boils down to the optimization of two OP models.
16
17 Nonetheless, the OP and the SOP models are not nested because the estimated
18
19 probabilities of observing the different events in the second stage and, hence, the
20
21 likelihood function, depend on the probabilities of the first stage, as shown in equation
22
23 [5].
24
25
26
27

28 A general drawback of all these models is that, given a change in a regressor, it is
29
30 only the direction of change in the probabilities of the two extreme cases that can be
31
32 unambiguously determined from the sign of the associated coefficient.⁷ The more levels
33
34 or categories, the more limited the model. Thus, in an OP model with seven levels of
35
36 aggressiveness, as that suggested in previous studies, the sign of the coefficients may
37
38 fail to describe the direction of change in the probabilities of the five intermediate
39
40 categories. This problem is significantly lessened with the SOP model since each node
41
42 in the decision process involves only three possible outcomes and, thus, only one
43
44 category is left in each stage with an ambiguous probability change. We compute
45
46 marginal probabilities to clarify the direction of the change in probabilities. They
47
48 measure the expected increase in the probability of each category after a marginal
49
50 increase in a given regressor. In the first stage, the marginal effect of, say, the k th
51
52 regressor on the probability of A_i taking value m is,
53
54
55
56
57
58
59
60

⁷ See Borooah (2001, p. 23-24) and Greene (2003, p. 738-739) for further details on this issue.

$$\frac{\partial P(A_i = m)}{\partial X_{ik}^A} = \begin{cases} -\phi(\delta_1^A - Z_i^A) \beta_k^A & \text{if } m = 1 \\ [\phi(\delta_1^A - Z_i^A) - \phi(\delta_2^A - Z_i^A)] \beta_k^A & \text{if } m = 2 \\ \phi(\delta_2^A - Z_i^A) \beta_k^A & \text{if } m = 3 \end{cases} \quad [6]$$

where $\phi(\cdot)$ is the normal density function. For the second stage, marginal probabilities are defined analogously.

3.2. Order aggressiveness

We use a generally accepted categorization of order aggressiveness originally proposed by Biais et al. (1995) that has been used, with some minimum variations, in previous papers. From the most to the least aggressive category,

- C7: Buy (sell) orders that demand more volume than is available at the best prevailing ask (bid) and are allowed to walk up (down) the book.
- C6: Buy (sell) orders that demand more volume than is available at the best ask (bid) but are not allowed to walk up (down) the book.
- C5: Buy (sell) orders that demand less volume than is available at the best ask (bid).
- C4: Orders with prices lying between the best bid and offer.
- C3: Buy (sell) orders with prices equal to the best bid (ask).
- C2: Buy (sell) orders with prices below (above) the best bid (ask).
- C1: Cancellations.

Categories C5 to C7 imply total or partial immediate execution of the order and, therefore, consume liquidity. These are the possible choices we consider for an impatient trader in the second stage of the SOP model. Categories C2 to C4 imply non-immediate execution and, therefore, provide liquidity. These are the possible choices we

consider for a patient trader in the second stage of the SOP model. Category C1 includes orders that withdraw liquidity, and they are to be chosen exclusively in the first stage of the SOP model.

3.3. Pieces of LOB information

We consider two large pieces of LOB information. The *best quotes* (henceforth BQ) set includes the following variables, all of them defined with respect to the incoming order:

- SPR : Bid-ask spread.
- NS^l (NO^l): Pending number of orders on the same (opposite) side of the book.⁸

The *additional quotes* (hereafter AQ) set includes,

- NS^{25} (NO^{25}): Accumulated number of orders on the same (opposite) side of the book.
- LS^{l2} (LS^{25}): Distance, in ticks, between the best and the second best (the second best and fifth best) quotes on the same side of the book.
- LO^{l2} (LO^{25}): Distance between the best and the second best (the second best and fifth best) quotes on the opposite side of the book.

The information contained in the second set of variables could be thought of as consisting of two components: Redundant information already contained in the best quotes, and brand new information only present in the secondary levels. To isolate the latter piece of information, the AQ set is defined as the residuals of a linear regression of each of its components on the contemporaneous and lagged values of the variables in

⁸ We also considered the quoted depth as an alternative to the number of orders. The correlation between both proxies was very high, so that our main findings barely differed. In this final version, we only report the analysis based on the number of orders.

1
2
3 the BQ set. In this way, any explanatory power attributed to the secondary levels of the
4 book will be (linearly) unconnected with the information on the best quotes.
5
6
7

8 We consider three alternative specifications for the SOP model. The “Baseline”
9 model (BM) only includes the first lag of the dependent variable in the explanatory
10 variable set; the “Best Quotes” model (BQM) adds the BQ set of explanatory variables
11 to the BM model; the “Complete Book” model (CBM) adds the AQ set of explanatory
12 variables to the BQM. In the next sections, we will evaluate the relative performance of
13 each specification in each stage of the SOP model both in-sample and out-of-sample.
14 We will also compare the performance of the CBM-SOP model with an alternative
15 CBM-OP model.
16
17
18
19
20
21
22
23
24
25
26
27

28 3.4. Hypotheses 29 30

31 The choice of LOB variables described above is supported by the existing literature.
32 According to Foucault (1999), Foucault et al. (2005), and Rinaldo (2004), we
33 hypothesize that the wider the spread, the higher the frequency of patient traders on the
34 first stage of the SOP model. A wider spread increases the costs of trading aggressively.
35 In the second stage, we expect a lower proportion of the least aggressive market orders
36 (C5) in favor of most aggressive (C7) market orders as the inside spread increases. This
37 hypothesis builds on the fact that C7 traders bear higher immediacy costs than those
38 given by the inside spread, since they consume more than the depth available at the best
39 opposite quote. Contrarily, immediacy costs of C5 orders are entirely determined by the
40 spread. We therefore expect C5 orders to be more sensible than C7 orders to variations
41 in the inside spread. Regarding limit orders, according to Foucault (1999), Foucault et
42 al. (2005), and Goettler et al. (2005) we should expect limit order traders to become
43 more aggressive in their price improvements as the bid-ask spread widens. Therefore,
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 we hypothesize limit order traders to submit more C4 (inside the best quotes) limit
4
5 orders as the spread increases.
6
7

8
9 Regarding quoted depth at the best quotes, the crowding out effect of limit orders
10 predicted by Parlour (1998) sets the basis for our hypotheses. Additionally, in Handa et
11 al.'s (2003) framework higher imbalances between buyers and sellers are to be
12 associated with increased order aggressiveness. Rinaldo (2004), Ahn et al. (2001), and
13 Ellul et al. (2007), among others, provide supporting empirical evidence. Accordingly,
14 in the first stage of our SOP model, we expect an increase in the rate of arrival of
15 impatient (patient) buyers (sellers) when the quoted depth at the best bid quote
16 increases. Similarly, increases in the depth at the best ask is expected to encourage more
17 impatient (patient) sellers (buyers).
18
19
20
21
22
23
24
25
26
27
28
29

30 With respect to limit order traders in the second stage, a larger depth at the best quote
31 makes it more attractive to “jump the queue” by submitting inside the quote (C4) limit
32 orders and conversely less attractive to submit limit orders at the quote (C3) or away
33 from the best quote (e.g., Ellul et al., 2007). Traders with stale limit orders may also
34 engage in a “chasing” strategy (see Hasbrouck and Saar, 2007) consisting of canceling
35 the old limit orders and resubmitting them more aggressively. Moreover, according to
36 Parlour (1998), we expect patient traders on the thinner side of the market to submit less
37 aggressive limit orders, anticipating an increase in the arrival rate of market orders from
38 the opposite side.
39
40
41
42
43
44
45
46
47
48
49
50

51
52 Coppejans et al. (2004) provide evidence that the actual costs incurred by traders are
53 significantly lower than the costs that would be incurred under a raw strategy that would
54 ignore time-variations in liquidity, particularly for large trades. They conclude that
55 traders act strategically, taking advantage of periods of liquidity surpluses and avoiding
56 liquidity deficits. Their finding also agrees with traders fitting order size to the state of
57
58
59
60

1
2
3 the LOB, for example, by splitting orders into smaller (C2) ones, or by submitting
4 market-to-limit (C3) orders. Thus, in Mendelson and Tunca (2000), discretionary
5 liquidity traders adjust the order size along with the changing market depth, equalizing
6 trading costs across sizes. Similarly, De Winnie and D'hondt (2007) and Pardo and
7 Pascual (2007) both show that the detection of hidden volume on the best opposite
8 quote fosters order aggressiveness. Traders behave strategically, however, and just seize
9 the opportunity for depth improvement, without consuming more than available at the
10 best opposite quote. In accordance with the previous findings, we would expect the
11 available depth at the best quotes of the LOB to shape the size of the incoming market
12 orders, but we would not expect an increase in the proportion of aggressive market
13 orders. Namely, we hypothesize an increase in the submission of C5 market orders in
14 the second stage of our SOP model as NO^l goes deeper.

15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Goettler et al.'s (2005) model predicts that a deep book away from the best ask (bid) quote signals that the best quote price is "too low" ("too high"). Their simulation results suggest that the higher NS^{25} , the lower the frequency of both market orders and aggressive limit orders, and the higher the frequency of limit orders at or away from the best quote. Contrarily, higher NO^{25} may lead to more aggressive order submissions. Accordingly, we hypothesize NS^{25} (NO^{25}) to decrease (increase) the aggressiveness of both patient and impatient traders order choices in the second stage of the SOP model. Regarding the first stage of the SOP model, our prediction is more ambiguous. Order imbalances beyond the best quotes may reinforce Parlour's (1998) crowding out effect. Thus, increased competition in the form of depth below (above) the best bid (ask) may result in a higher frequency of impatient (patient) buyers (sellers). Deep book imbalances may also anticipate short-term price movements, as suggested by Kavajecz and Odders-White (2004). Since rising (falling) markets are expected to prolong

1
2
3 (shorten) the time-to-execution of limit buy (sell) orders, a thick bid (ask) book beyond
4
5 the best quotes may induce incoming buyers (sellers) to be more impatient and
6
7 incoming sellers (buyers) to be more patient.⁹ The simulations in Goettler et al. (2005),
8
9 however, point towards the opposite prediction (see Tables IV and VI in p. 2170). We
10
11 therefore treat the impact of NS^{25} and NO^{25} on the first stage of the SOP model as a
12
13 theory-free empirical issue.
14
15

16
17
18 “Length” measures are less common in microstructure research. Given the quoted
19
20 depth, the length of the LOB determines the expected price impact of large market
21
22 orders. Thus, LO^{l2} measures the demanded compensation for consuming more than
23
24 available at the best opposite quote. According to Glosten (1994), a lengthy ask (bid)
25
26 book may indicate that potential sellers (buyers) have higher upper (lower) tail
27
28 expectations. As long as this signaling effect outweighs the increased costs of trading, a
29
30 lengthy ask (bid) book may increase buy (sell) pressure (see Hall and Hautsch, 2007).
31
32 Similarly, a lengthy book on the ask side may suppose a longer time-to-execution for a
33
34 limit buy order, causing incoming buyers to be more aggressive and, consequently,
35
36 reducing the waiting costs of incoming patient sellers. Following this reasoning, we
37
38 expect the length of the book to reinforce the crowding out effect on the first stage of
39
40 the SOP model: the length of the ask (bid) side is expected to be positively correlated
41
42 with the frequency of impatient buyers (sellers) and positively correlated with the
43
44 probability of patient sellers (buyers). Regarding the second stage, all else equal, we
45
46 would expect a high LO^{l2} to discourage the most aggressive type of market order (C7)
47
48 traders because of the increased costs of transaction. A narrow LS^{l2} may signal a
49
50 crowded LOB, which may encourage patient traders to place their orders in front of the
51
52 queue so as to gain time precedence. In line with Glosten (1994), we would also expect
53
54
55
56
57
58
59
60

⁹ Beber and Caglio (2005) and Chan (2005) report a positive connection between momentum and order aggressiveness.

1
2
3 high values of LS^{l2} and LS^{25} to discourage aggressive limit order submissions since
4
5 they may signal a high perceived picking-off risk.
6
7

8 9 **4. Market background and data**

10
11 The Spanish Stock Exchange Interconnection System (henceforth SIBE) is the
12
13 electronic platform that, since 1995, holds all the Spanish stocks that achieve pre-
14
15 determined minimum levels of trading frequency and liquidity.¹⁰ Every order submitted
16
17 to the system is electronically routed to a centralized LOB.
18
19

20
21 The SIBE is organized as an order-driven market with a daily continuous trading
22
23 session, from 9:00 a.m. to 5:30 p.m., and two call auctions that determine the opening
24
25 and closing prices. During the continuous trading session, a trade takes place if an only
26
27 if an order hits the quotes. Pre-arranged trades are not allowed, and price-improvements
28
29 are impossible.¹¹ There are no market makers, and there is no floor trading. The market
30
31 is governed by a strict price-time priority rule. The minimum price variation (tick)
32
33 equals €0.01 for prices below €50 and €0.05 for prices above €50. The minimum trade
34
35 size is one share.
36
37
38
39

40
41 There are three basic types of orders: market, limit, and market-to-limit. Market
42
43 orders execute against the best prices on the opposite side of the LOB. Any excess that
44
45 cannot be executed at the best bid (ask) quote is executed at less favorable prices by
46
47 walking down (up) the book until the order is fulfilled. Market-to-limit orders are
48
49 limited to the best opposite-side quote at the time of entry. Any excess that cannot be
50
51 executed is converted into a limit order. Finally, limit orders are to be executed at the
52
53
54
55

56
57 ¹⁰ According to the Focus Monthly Bulletin of June 2006 of the World Federation of Exchanges
58 (www.world-exchanges.org), the Spanish market is the fourth European market in terms of market
59 capitalization (\$US 1,116,146 millions), right after the LSE (\$US 3,370,070 millions), Euronext (\$US
60 3,192,428 millions) and the Deutsche Bourse (\$US 1,412,118 millions). The SSE is also the 4th European
market in terms of total value of share trading.

¹¹ Because of the allowance of hidden limit orders, however, depth improvements are possible.

1
2
3 limit price or better. Any unexecuted part of the order is stored in front of the book at
4
5 the limit price. Market orders belong to the C5 to C7 categories of aggressiveness
6
7 previously defined, market-to-limit orders are C6 orders, and limit orders are C2 to C4
8
9 orders.¹² By default, all orders expire at the end of the session.¹³
10
11
12

13 Our database consists of all the updates of the 5 best bid/ask quotes of the LOB and
14
15 all trades executed from July to December 2000 (124 trading days) during the
16
17 continuous trading session. The LOB data includes quotes, disclosed depth, and the
18
19 number of orders supporting each quote. All the movements of the book are time-
20
21 stamped to the nearest hundredth of a second. The trading data is updated each time the
22
23 best quotes change. It details the trade price, the trade size, and the best quotes before
24
25 and after each trade. We match both files using an algorithm introduced by Pardo and
26
27 Pascual (2007). Once matched, we can classify all the LOB updates into one of the
28
29 seven categories of aggressiveness formerly defined. It is also easy to distinguish
30
31
32
33
34
35
36
37

38
39 ¹² Limit orders at a price equal to the best quote on the opposite side of the book and of smaller (larger)
40
41 size than the quantity available at that opposite quote cannot be distinguished in practice from market
42
43 (market to limit) orders. Therefore, we pool them as C5-market (C6-market to limit) orders. Similarly, we
44
45 put together limit orders that walk up or down the book and become totally fulfilled with C7-market
46
47 orders, and market-to-limit orders with size smaller than the available depth as C5-market orders. Limit
48
49 orders that walk up or down the book but are only partially executed represent less than 0.3% of all orders
50
51 submitted. These orders are also considered C7. By a limit order that walks up or down the book we mean
52
53 a limit order to buy (sell) which price is above (below) the prevailing best ask (bid) quote and which size
54
55 is larger than the depth available at the best opposite quote. This order will consume the best ask (bid)
56
57 quote on the book, and the excess will be executed at less favorable ask (bid) prices until it either fully
58
59 executes or reaches a quote above (below) the limit price.

60
¹³ For all type of orders, brokers may specify special conditions: (a) "Execute or eliminate" means the
order must be executed immediately. In case of partial execution, the unexecuted part must be eliminated
right away. (b) "Fill or kill" also requires instantaneous execution. In this case, partial execution is not
possible. The order must be fully executed or fully eliminated. (c) "Minimum execution" implies that at
least a given part of the order must be immediately executed. If the minimum execution is not possible,
the whole order must be eliminated. If the minimum execution is possible, the rest of the order must be
treated as a regular order. (d) Finally, partially undisclosed limit orders, known as "iceberg" orders, are
allowed. Our database does not identify orders submitted with special conditions. Hidden orders, for
example, are only detectable upon execution. In practical terms, orders with these conditions are therefore
undistinguishable from some of the seven categories of aggressiveness defined above. In this paper, we
consider the undisclosed depth when classifying orders. We classify all orders relative to the total depth
(disclosed plus undisclosed) available at the best quotes. Market orders with size larger than the disclosed
depth at the best-opposite quote on the book are classified as C7 if and only if they exhaust all the
available depth at that quote.

1
2
3 between buyer and seller initiated trades because the LOB always acts as the passive
4 side on every trade.
5
6

7
8 The SSE database has some drawbacks we consider convenient to mention. We
9 observe pictures of the whole book, but we do not have information order by order.
10 Therefore, we cannot follow the history of all the orders submitted. This is a limitation
11 because we can not compute the realized time-to-execution or time-to-cancellation of all
12 limit orders. We do not have LOB information about the opening and closing auctions.
13 We must therefore exclude these relevant periods for price discovery from our analysis.
14 The initial LOB each day, however, corresponds to the unexecuted limit orders coming
15 from the pre-auction phase. So, the role these orders play in order choice is considered.
16 Moreover, the SSE allows submitting iceberg orders. Recent empirical studies (see De
17 Winnie and D'hondt, 2007, and Pardo and Pascual, 2007) have shown that the state of
18 the LOB matters in choosing between disclosed and undisclosed limit orders. Moreover,
19 hidden volume, when revealed to the marketplace, affects order choice by increasing
20 order aggressiveness on the opposite side of the market. It would be therefore desirable
21 to extend our analysis by including iceberg orders either as an additional possible order
22 choice by incoming traders or as a conditional variable for order choice. Unfortunately,
23 we can only detect the presence of iceberg orders at execution and, even then, we
24 cannot always know the exact amount of the hidden volume. Thus, we leave this
25 possible extension for future research.
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

50
51 Our initial sample consists of all the stocks that were part of the official market
52 index, the IBEX-35, during 2000.¹⁴ We exclude one stock involved in a merger. Table I
53 provides descriptive statistics on the 36 stocks conforming the final sample. It is
54
55
56
57

58
59 ¹⁴ The IBEX-35 is composed of the 35 most liquid and active SIBE-listed stocks during the most recent
60 six-month control period. The composition is ordinarily revised twice a year. Extraordinary revisions are possible due to major events like mergers or new stock issues. During 2000, a total of 37 stocks were index constituents.

possible to appreciate huge differences between them in terms of both liquidity and activity. The most illiquid stock in the sample (ALB) shows a bid-ask spread which is 14.2 times larger (in number of ticks) than the most liquid stock (SCH), a quoted depth at the ask side (in number of shares) which is 60.5 times smaller, and a distance between the first and the fifth ask quotes in the book, also in number of ticks, which is 5 times larger. Similarly, the most active stock (TEF) presents an average daily volume (in shares) and an average daily number of trades which are, respectively, 157 and 86.7 times larger than the corresponding statistics for the least frequently traded stocks (ANA and AUM). Our sample is representative in the sense that it represents about 94% of the total volume traded and 64% of the total market capitalization.

[Table I]

In the next section, we evaluate the relative performance of each SOP model in section 3, both in-sample and out-of-sample. For in-sample analyses, we use data from July to November 2000. For out-of-sample analyses, we use December 2000 data.

5. Empirical findings

Table II provides summary statistics about order aggressiveness. We classify each update of the LOB into the 7 categories of aggressiveness defined earlier. We provide the distribution of orders by category over all the orders submitted (Panel A) and over buyer-initiated and seller-initiated orders (Panel B). The most frequent category is small market orders (C5), which represents 40.39% of all orders submitted. The least frequent category is that of large market orders (C7), which represent 3.75% of all orders.¹⁵ According to Panel A, 35.39% of the orders submitted provide liquidity, 50.21% consume liquidity, and 14.41% withdraw liquidity. We do not find remarkable

¹⁵ Similar patterns are reported by Rinaldo (2004), for the Swiss Stock Exchange, Griffiths et al (2000) for the Toronto Stock Exchange, Biais et al (1995) for the Paris Bourse, now Euronext, Al-Suhaibani and Kryzanowski (2000) for the Saudi Stock Market, and Beber and Caglio (2005) for the NYSE.

1
2
3 differences between buyer-initiated and seller-initiated orders. Even though our sample
4
5 period is biased by a bear phase, the IBEX constituents, for the most part, remained
6
7 relatively stable.
8
9

10
11 [Table II]
12

13
14 Table III summarizes the maximum likelihood estimation of the first stage of the
15
16 CBM SOP model for the 36 stocks. We estimate separated models for buyer-initiated
17
18 traders and seller-initiated traders. We also provide the marginal probabilities in [6]
19
20 evaluated at the mean of the explanatory variables.
21
22

23
24 [Table III]
25

26
27 We find a strong first order positive autocorrelation in order aggressiveness. That is,
28
29 the probability of observing an aggressive type of order increases if the immediately
30
31 preceding event on the same side of the market is also an aggressive order. This is
32
33 consistent with the “diagonal effect” reported by Biais et al. (1995), but it is at odds
34
35 with Parlour (1998), who predicts that submitting an aggressive order is less attractive
36
37 when the immediately preceding order at the same side is a limit order. Similar order-
38
39 by-order serial correlation in order type is provided by Ellul et al. (2007).¹⁶
40
41
42

43
44 The estimated coefficients for the BQ set are consistent with previous empirical
45
46 findings (using the OP model), and corroborate the hypotheses posted in section 3. A
47
48 wider spread increases the probability of the incoming trader being patient, consistently
49
50 with Foucault (1999), a finding supported by almost all the stocks in our sample.
51
52 Consistent with Parlour’s (1998) crowding out effect, we find that the thicker the book
53
54 on the buy (sell) side, the more impatient the incoming buyer (seller) becomes.
55
56 Marginal probabilities show that, as the bid-ask spread (depth at the own side) increases
57
58
59

60

¹⁶ Biais et al. (1995) argue that this diagonal effect may reflect imitative behavior by uninformed traders, order splitting, traders reacting to the same public information, or competition between liquidity providers.

1
2
3 the probability of the most impatient type of trader decreases (increases), while the
4
5 probability of the most patient types increases (decreases).
6
7

8
9 Regarding depth at the best opposite quote, we find a statistically significant and
10
11 negative relationship, consistent with Parlour's (1998) prediction that a thicker book on
12
13 the sell (buy) side will decrease the impatience of new buyers (sellers). Marginal
14
15 probabilities, however, show that sellers' behavior is at odds with them expecting a
16
17 crowding out effect. This finding might be a result of the bear market affecting our
18
19 sample period. However, Beber and Caglio (2005) report the same asymmetry for the
20
21 NYSE using the TORQ database. They argue this finding indicates that sellers are more
22
23 impatient than buyers, and prefer to take advantage of the lower costs of liquidity. In
24
25 any case, the marginal probabilities show that SPR and the NS^l have a major impact,
26
27 with NO^l being less relevant.
28
29
30
31

32
33 Regarding the AQ set, Table III shows these variables are less important than the best
34
35 quotes in characterizing incoming traders' patience. For secondary depth (NO^{25} and
36
37 NS^{25}), our findings support the hypothesis that the book beyond the best quotes may
38
39 reinforce the crowding out effect predicted by Parlour (1998). In addition, the marginal
40
41 probabilities confirm that, in this case, the finding is consistent among sellers and
42
43 buyers. Therefore, in our sample, the crowding out effect for sellers is fully satisfied
44
45 when we consider the secondary book depth. We also obtain a weak negative effect of
46
47 LS^{l2} and a strong positive effect of LO^{l2} on order aggressiveness. As previously
48
49 discussed, a small LS^{l2} may signal a crowded book on the same side of the market as the
50
51 incoming trader. This variable therefore reinforces the findings obtained with NS^l and
52
53 NS^{25} . Moreover, a rational buyer (seller) will interpret a higher dispersion on the ask
54
55 (bid) side of the book as a longer expected time-to-execution for a new limit order to
56
57 buy (sell), inducing the incoming buyer (seller) to be more aggressive. Finally, both
58
59
60

1
2
3 LS^{25} and LO^{25} tend to decrease the frequency of aggressive traders. In terms of statistical
4
5
6 significance, however, this last relationship is weak.
7

8
9 Table IV shows, for the first stage of the SOP model, the relative in-sample and out-
10
11 of-sample performance of each of the three model specifications (BM, BQM and
12
13 CBM). We report four alternative goodness-of-fit measures that correspond to the in-
14
15 sample (adjusted) pseudo- R^2 s of McFadden (1973, p.121), Maddala (1983, p.39) with
16
17 the Cragg and Uhler (1970) correction, Aldrich and Nelson (1984) with the Veall and
18
19 Zimmermann (1992) correction, and McKelvey and Zavoina (1975).¹⁷ For each
20
21 measure, we provide its median across the 36 stocks for the BM, its percentage increase
22
23 for the BQM with respect to the BM, and for the CBM with respect to the BQM.
24
25
26
27

28 [Table IV]
29
30

31 The results across measures are similar. Consider the McKelvey and Zavoina's
32
33 pseudo- R^2 . When the BQ variables are added to the simplest autoregressive model, we
34
35 get an in-sample fit improvement of 51.68% for sellers and 82.75% for buyers. In
36
37 addition, when the whole LOB is taken into account the fit improves an additional
38
39 52.40% for sellers and 42.57% for buyers. This increasing pattern indicates that the
40
41 state of the book determines, at least partially, the nature of the incoming trader.
42
43 Moreover, traders examine not only the best quotes, but also the less aggressive quotes.
44
45 Our results seem to attribute a more noteworthy role to the book beyond the best quotes
46
47 than it should be expected given Cao *et al*'s (2003) findings. Similarly, from the out-of-
48
49
50
51
52

53 ¹⁷ No one of these measures is universally accepted or employed. The values between zero and one have
54
55 no natural interpretation, though it has been suggested that the pseudo- R^2 value increases as the fit of the
56
57 model improves. In a comparative analysis performed by Veall and Zimmermann (1996), these authors
58
59 conclude that, for the particular case of the ordered probit model, the pseudo- R^2 due to McKelvey-
60
Zavoina outperforms the other measures and has a strong numerical relationship to the OLS- R^2 in the
latent variable. The Veall-Zimmermann and the Cragg-Uhler's measures also perform reasonably well.
We include the McFadden's pseudo- R^2 because it is the most common in statistical packages. For a
review of all these goodness-of-fit measures see Veall and Zimmermann (1996). As in standard
regression analysis, we use adjusted versions of these measures to take into account the inclusion of
additional explanatory variables.

1
2
3 sample McKelvey-Zavoina (adjusted) pseudo- R^2 , we observe that the predictive
4
5 capacity of the BQM outperforms that of the BM by a median 108.96% for sellers and
6
7 141.22% for buyers. When the complete book is considered there is an additional
8
9 improvement of 31.39% for sellers and 40.11% for buyers.
10
11

12
13 As an alternative to the previous point measures of goodness-of-fit, we perform the
14
15 following forecasting exercise. Using the in-sample estimated coefficients, we compute
16
17 the one-step-ahead forecasting probability for each of the 3 categories of aggressiveness
18
19 and for each out-of-sample observation. Then, we compute how often the SOP model
20
21 beats the unconditional probability given by the in-sample relative frequency of each
22
23 category. The results are summarized at the bottom panel of Table IV. We find that the
24
25 CBM outperforms the BQM and the BM. The CBM beats the unconditional frequency
26
27 more often than the other two models for 47.22% for both buyers and sellers; only for
28
29 22.22% (16.67%) of the stocks the BQM outperforms the other two models for sellers
30
31 (buyers). We also provide a direct comparison between models. The CBM does better
32
33 than the BQM for 94.44% (97.22%) of the stocks in sellers (buyers) model, in the sense
34
35 that the CBM allocates a higher probability to the actual category of aggressiveness than
36
37 the BQM. Both book models usually improve on the BM.
38
39
40
41
42
43

44
45 Table V summarizes the estimation of the second stage of the CBM. We report the
46
47 coefficients for patient (Panel A) and impatient (Panel B) traders, distinguishing
48
49 between buyer and seller-initiated orders, in separated panels. We also provide the
50
51 marginal probabilities in each case.
52
53

54 [Table V]
55
56

57
58 Table V provides evidence that the bid-ask spread is very important in determining
59
60 both patient and impatient order choices. Marginal probabilities in Panel A show that

1
2
3 patient traders submit more aggressive limit orders (C4) as the bid-ask spread increases.
4
5 This finding is consistent with the theoretical predictions of Foucault et al. (2005), with
6
7 previous empirical evidence (e.g., Biais et al., 1995, and Ranaldo, 2004), and with our
8
9 hypothesis in section 3. Regarding impatient traders, marginal probabilities in Panel B
10
11 report a strong decrease in the likelihood of submitting C5 orders, which favors an
12
13 increase in the proportion of C6 and C7 orders. This finding provides support to our
14
15 hypothesis in section 3 that an increase in the inside spread has more dramatic
16
17 consequences on the least aggressive type of traders among all the impatient ones.
18
19
20
21

22
23 Regarding quoted depth at the best quotes, Table V shows that same-side (opposite-
24
25 side) depth is more relevant for patient traders' (impatient traders') order choices.
26
27 Marginal probabilities in Panel A show that patient traders facing a thick best quote (a
28
29 large NS^l) are more likely to "jump the queue", which is consistent with Handa et al.'s
30
31 (2003) arguments, among others. As hypothesized, the submission of inside-the-quotes
32
33 (C4) limit orders increases as NS^l augments, in detriment of less aggressive types (C3
34
35 and C2) of limit orders. Contrarily, NO^l is only statistically significant for 3 (2) stocks
36
37 for patient buyers (sellers).
38
39
40
41

42
43 Even though a deep best bid (ask) increases the rate of arrival of impatient buyers
44
45 (sellers) (see Table III), the marginal probabilities in Panel B indicate that these
46
47 additional traders use to submit C5 (non-aggressive) market orders. This last finding is
48
49 week, anyway: NS^l is only statistically significant for 14 (10) stocks for impatient
50
51 sellers (buyers). Quite the opposite, the thickness of the best opposite quote has a strong
52
53 effect on order choice by impatient traders. As hypothesized in section 3, the higher the
54
55 NO^l , the higher (lower) the proportion of impatient traders choosing C5 (C6 and C7)
56
57 market orders. This behavior is observed for both buyers and sellers, despite the
58
59 asymmetric pattern evidenced in the first step of the SOP model. Thus, impatient traders
60

1
2
3 take advantage of liquidity surpluses and adjust the size of their orders to the available
4
5 opposite depth, probably seeking to equalize the trading costs across trade sizes, as in
6
7 Mendelson and Tunca (2000).
8
9

10
11 Albeit the BQ set is much more relevant, the AQ set also determines order choice at
12
13 the second stage of the SOP model. Panel A in Table V shows that patient buyers'
14
15 (sellers') order choices depend more on the book beyond the best bid (ask) than on the
16
17 book beyond the best ask (bid). Among all the AQ variables, LS^{l2} is found to be the
18
19 most relevant. In general, the length of their own side of the book decreases the
20
21 aggressiveness of incoming limit order traders. As with NS^l , a narrow LS^{l2} may signal a
22
23 crowded LOB (or a low picking-off risk), which encourages patient traders to place
24
25 their orders in front of the queue so as to gain time precedence. We also find NS^{25} to be
26
27 negatively related to the aggressiveness of patient traders, which is totally consistent
28
29 with the simulation results in Goettler et al. (2005). A crowded book beyond the best
30
31 ask (bid) quote signals the best quote is too low (high) leading incoming patient traders
32
33 on the same side to be more conservative. As far as we know, this is the first paper
34
35 providing empirical support to Goettler et al.'s (2005) prediction. The remaining same-
36
37 side variables are less critical. The incidence of the opposite side of the book on the
38
39 aggressiveness of the incoming limit order trader, as given by NO^l , LO^{l2} and LO^{25} , is
40
41 weak.
42
43
44
45
46
47
48

49
50 With regard to impatient traders, marginal probabilities in Panel B of Table V point
51
52 towards LO^{l2} as the most remarkable determinant of order choice in the AQ set. As
53
54 discussed in section 3, as LO^{l2} increases it becomes more attractive to submit non-
55
56 aggressive (C5) market orders, and less attractive to submit aggressive (C7) market
57
58 orders. We also find that NO^{25} is positively related to the aggressiveness of the
59
60 incoming impatient traders on the opposite side of the market, which is consistent with

1
2
3 Goettler et al.'s (2005) simulation results. Hence, C7 buy (sell) orders are more frequent
4
5 when the LOB away from the best ask (bid) quote signals that the best ask (bid) is too
6
7 low (high). Consistently, we also report a negative effect of NS^{25} (also LS^{12} and LS^{25}) on
8
9 the aggressiveness of the incoming impatient trader.
10
11

12
13 Table VI reports the relative in-sample and out-of-sample performance of the second
14
15 stage of the three SOP models considered.
16
17

18 [Table VI]
19
20

21 The in-sample analysis shows that limit order traders' choices crucially depend on
22
23 the book information. There is a median improvement 134.48% for sellers and 177.09%
24
25 for buyers when the best quotes are taken into account. More important, the CBM
26
27 improves on the BQM by a median of 263.19% for sellers and 210.51% for buyers,
28
29 implying that patient traders take their decisions by examining not only the best quotes
30
31 but also the book beyond the best quotes. The out-of-sample (adjusted) pseudo- R^2 leads
32
33 to the same conclusion. In addition, the CBM obtains the best scores against the relative
34
35 frequency rule. It always outperforms the BQM, allocating higher probabilities to the
36
37 actual event than the BQM and the BM.
38
39
40
41
42

43 The results for the impatient traders are remarkably different. Table VI shows that
44
45 impatient traders' order choices depend, to a great extent, on the best quotes. There is a
46
47 median in-sample fit improvement of 1100% for sellers and 991% for buyers when the
48
49 BQ set is added to the simplest model. However, the CBM improves on the BQM only
50
51 by a median 23.75% for sellers and 34.03% for buyers. This finding suggests that the
52
53 most impatient traders base their strategic decisions primarily on the best quotes. The
54
55 out-of-sample predictive performances support this conclusion: the pseudo- R^2 for sellers
56
57 (buyers), for example, increases a negligible 7.82% (21.88%) from the BQM to the
58
59
60

1
2
3 CBM. Moreover, book-based models rarely do better than the relative frequency rule
4
5 and the BQM probabilities outperform those of the CBM as many times as the CBM
6
7 outperforms the BQM.¹⁸
8
9

10
11 As a final empirical analysis, we compare the performance of the two-stage SOP
12
13 model with the one-stage OP model. We repeat the forecasting exercise at the bottom of
14
15 Tables IV and VI but comparing the CBM-SOP model with the CBM-OP model. This
16
17 time, we perform the analysis both in-sample and out-of sample. Our findings are
18
19 summarized in Table VII. We report the number of stocks and the percentage of
20
21 observations for which the SOP model outperforms the OP model, in the sense that the
22
23 former model attributes a higher probability to the actual event than the latter model.
24
25 We find that the SOP model beats the OP model in-sample for 100% of the stocks and
26
27 about 60% of the observations, both for buyer-initiated and seller-initiated orders. The
28
29 out-of-sample analysis provides similar figures. When we control for the type of order,
30
31 we find that the SOP model beats the OP model in almost all categories, both in-sample
32
33 and out-of-sample, with the exception being cancellations in-sample. The improvement
34
35 is particularly remarkable for the most aggressive categories.
36
37
38
39
40
41

42 [Table VII]
43
44

45 **6. Summary and conclusions**

46
47

48 We study the role the limit order book (LOB) plays in order choice by both patient
49
50 and impatient traders. Our contribution is twofold. This is the first paper in modeling
51
52 order aggressiveness using a two-stage sequential ordered probit (SOP) model. The
53
54
55

56
57 ¹⁸ We have also performed likelihood-ratio tests to compare the three model specifications. For all the
58
59 stocks and for all the stages of the SOP model, the BQM model is found to beat the BM model, and the
60
CBM to beat the BQM at the 1% level. Therefore, we corroborate that the book beyond the best quotes matters. Moreover, the chi-square values suggest that secondary levels of the book are more useful to explain patient traders' decisions than to impatient traders' decisions. These findings are available upon request from the authors.

1
2
3 SOP model allows studying patient traders' order choices and impatient traders' order
4
5 choices separately. This is valuable because we are able to separate the decision of
6
7 whether to withdraw, provide or consume liquidity, from the decision of choosing a
8
9 particular type of order. We show the superior forecasting performance, both in-sample
10
11 and out-of-sample, of the SOP model versus the traditional ordered probit model of
12
13 order aggressiveness.
14
15

16
17
18 We also contribute to the literature by considering the whole LOB, instead of just the
19
20 best quotes, in explaining order choice. We distinguish between two pieces of book
21
22 information, the best quotes and the book beyond the best quotes. We evaluate the
23
24 relative improvement on the in-sample and out-of-sample goodness-of-fit performance
25
26 of the SOP model when different pieces of book information are sequentially added as
27
28 regressors.
29
30

31
32
33 We find that the whole book contributes in determining whether a trader becomes a
34
35 liquidity provider or a liquidity consumer, though the inside spread and the quoted
36
37 depth at the best quote on the same side as the incoming trader are particularly
38
39 powerful. Patient traders crucially rely on the book beyond the best quotes to place their
40
41 limit orders. In contrast, impatient traders primarily depend on the best quotes. Our
42
43 findings also qualify previous empirical studies:
44
45

- 46
47
48 (a) The frequency of patient traders increases with the bid-ask spread, and decreases
49
50 with the thickness of the depth at the best quote on the same side of the market. The
51
52 thickness of the best quote on the opposite side of the market has an asymmetric
53
54 impact. It increases the frequency of patient buyers while it increases the frequency
55
56 of impatient sellers. The book beyond the best quotes reinforces Parlour's (1998)
57
58 "crowding out" effect.
59
60

- 1
2
3 (b) Consistently with Foucault et al. (1995), we find that the higher the bid-ask spread,
4 the higher the frequency of inside-the-quotes limit orders. A wider spread also
5 decreases the likelihood of non-aggressive market orders, increasing the frequency
6 of the most aggressive type of market orders.
7
8
9
10
11
12
13 (c) Patient traders base their order choices primarily on the state of the LOB on the
14 same side of the market, while impatient traders mostly look at the state of the LOB
15 on the opposite side. Patient traders submit more aggressive limit orders as the depth
16 at the best quote increases, while their choice is almost unaffected by the opposite
17 side depth at the best quote. Impatient traders submit less aggressive market orders
18 as the thickness of the opposite quote increases. They take advantage of liquidity
19 surpluses while adjusting the size of their orders to the available depth, as in
20 Mendelson and Tunca (2000).
21
22
23
24
25
26
27
28
29
30
31
32
33 (d) Consistently with Goettler et al. (2008), we show that a thicker LOB beyond the
34 best ask (bid) signals that the best quote is too low (high). Thus, the deeper the book
35 above (below) the best ask (bid), the more aggressive both patient and impatient
36 buyers (sellers) become.
37
38
39
40
41
42
43 (e) A lengthy book on the opposite (same) side of the market, decreases the
44 aggressiveness of incoming impatient (patient) traders.
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

References

- 1
2
3
4
5
6 Ahn, Hee-Joon, Kee-Hong Bae, and Kalok Chan, 2001, Limit orders, depth and volatility: evidence from
7 the stock exchange of Hong Kong, *Journal of Finance*, 56, 2, 769-790.
- 8 Aldrich, J.H., and F.D. Nelson, 1984, *Linear Probability, Logit, and Probit Models*, Sage University
9 Press.
- 10 Al-Suhaibani, Mohammad, and Lawrence Kryzanowski, 2000, An explanatory analysis of the order book,
11 order flow and execution on the Saudi Stock Market, *Journal of Banking and Finance*, 24, 1323-
12 1357.
- 13 Anand, Amber, Sutago Chakravarty, and Terrence Martell, 2005, Empirical evidence on the evolution of
14 liquidity: Choice of market versus limit orders by informed and uninformed traders, *Journal of*
15 *Financial Markets*, 8, 289-309.
- 16 Beber, Alessandro, and Cecilia Caglio, 2005, Order submission strategies and information: Empirical
17 evidence from the NYSE, *FAME Research Paper Series*, 146, International Center for Financial Asset
18 Management and Engineering.
- 19 Biais, Bruno, Pierre Hillion, and Chester Spatt, 1995, An empirical analysis of the limit order book and
20 the order flow in the Paris Bourse, *The Journal of Finance*, 50, 1655-1689.
- 21 Bloomfield, Robert, Maureen O'Hara and Gideon Saar, 2005, The "make or take" decision in an
22 electronic market: Evidence on the evolution of liquidity, *Journal of Financial Economics*, 75, 165-
23 199.
- 24 Borooah, Vani K., 2001, Logit and Probit: Ordered and Multinomial Models, *Quantitative Applications in*
25 *the Social Sciences Series*, Sage University Paper.
- 26 Cao, Charles, Oliver Hansch, and Xiaoxin Wang, 2003, The informational content of an open limit order
27 book, *Working Paper*, Pennsylvania State University.
- 28 Chan, Yue-cheong, 2005, Price movement effects on the state of the electronic limit-order book, *The*
29 *Financial Review*, 40, 195-221.
- 30 Chakrabarty, Bidisha, Zhaohui Han, and Konstantin Tyurin, 2006, A competing risk analysis of
31 executions and cancellations in a limit order market, *CAEPR Working Paper #2006-015*, Center for
32 Applied Economics and Policy Research, Indiana University.
- 33 Cohen, Kalman J., Steven F. Maier, Robert A. Schwartz, and David K. Whitcomb, 1981, Transaction
34 costs, order placement strategy and the existence of the bid-ask spread, *Journal of Political Economy*,
35 89, 287-305.
- 36 Copeland, Thomas E., and Dan Galai, 1983, Information effects on the bid-ask spread, *Journal of*
37 *Finance*, 38, 1457-1469.
- 38 Coppejans, Mark, Ian Domowitz, and Ananth Madhavan, 2004, Resiliency in an Automated Auction,
39 *Working Paper*, Barclays Global Investors.
- 40 De Winnie Rudy, and Catherine D'hondt, 2007, Hide-and-seek in the market: Placing and detecting
41 hidden orders, *Review of Finance*, 663-692.
- 42 Ellul, Andrew, Craig W. Holden, Pankaj Jain, and Robert Jennings, 2007, Order dynamics: Recent
43 evidence from the NYSE, *Journal of Empirical Finance*, 14, 636-661.
- 44 Foucault, Thierry, 1999, Order flow composition and trading costs in a dynamic limit order market,
45 *Journal of Financial Markets*, 99-134.
- 46 Foucault, Thierry, Ohad Kadan, and Eugene Kandel, 2005, Limit order book as a market for liquidity,
47 *Review of Financial Studies*, 18, 1171-1217.
- 48 Foucault, Thierry, Sophie Moinas, and Erik Theissen, 2007, Does anonymity matter in electronic limit
49 order markets? *Review of Financial Studies*, 20, 1707-1747.
- 50 Glosten, Larry, 1994, Is the electronic limit order book inevitable? *Journal of Finance*, 49, 1127-1161.
- 51 Goettler, Ronald L., Christine A. Parlour, and Uday Rajan, 2005, Equilibrium in a dynamic limit order
52 market, *Journal of Finance*, 5, 2149-2192.
- 53 Green, W.H., 2003, *Econometric Analysis*, 5th edition, Prentice Hall.
- 54 Griffiths, Mark D., Brian F. Smith, D. Alasdair S. Turnbull, and Robert W. White, 2000, The costs and
55 determinants of order aggressiveness, *Journal of Financial Economics*, 56, 65-88.
- 56 Hall, Anthony D., and Nikolaus Hautsch, 2007, Modelling the buy and sell intensity in a limit order book
57 market, *Journal of Financial Markets*, 10, 249-286.
- 58 Handa, Puneet, and Robert A. Schwartz, 1996, Limit order trading, *The Journal of Finance* 51, 1835-
59 1861.
- 60 Handa, Puneet, Robert A. Schwartz, and Ashish Tiwari, 2003, Quote setting and price formation in an
order driven market, *Journal of Financial Markets*, 6, 461-489.

- 1
2
3 Harris, Lawrence, 1998, Optimal Dynamic order submission strategies in some stylized trading problems.
4 *Financial Markets, Institutions and Instruments*, 7, 26-74.
- 5 Harris, Lawrence E., and Joel Hasbrouck, 1996, Market vs. limit orders: The SuperDOT evidence on
6 order submission strategy, *Journal of Financial and Quantitative Analysis*, 31, 213-231.
- 7 Harris, Lawrence E., and Venkatesh Panchapagesan, 2005, The information-content of the limit order
8 book: evidence from NYSE specialist actions, *Journal of Financial Markets*, 8, 25-67.
- 9 Hasbrouck, Joel, and Gideon Saar, 2007, Technology and liquidity provision: The blurring of traditional
10 definitions, *Working Paper*, New York University.
- 11 Irvine, Paul, Benston, George, and Eugene Kandel, 2000, Liquidity beyond the inside spread: measuring
12 and using information in the limit order book, *Working Paper*, Emory University.
- 13 Kaniel, Ron, and Hong Liu, 2006, So what orders do informed traders use? *Journal of Business*, 79,
14 1867-1913.
- 15 Kavajecz, Kenneth A., and Elisabeth R. Odders-White, 2004, Technical analysis and liquidity provision,
16 *The Review of Financial Studies*, 17, 1043-1071.
- 17 Lo, Andrew W., A. Craig MacKinlay, and June Zhang, 2002, Econometric models of limit-order
18 executions, *Journal of Financial Economics*, 65, 31-71.
- 19 Maddala, G.S., 1983, *Limited-dependent and Qualitative Variables in Econometrics*, New York:
20 Cambridge University Press.
- 21 Moinas, Sophie, 2004, Hidden orders and liquidity on a limit order market, *Working Paper*.
- 22 Mendelson, Haim, and Tunay I. Tunca, 2000, Transaction costs, information asymmetry, and dynamic
23 market design, *Working Paper*, Stanford University.
- 24 McFadden, D., 1973, Conditional logit analysis of qualitative choice behavior, in Zarembka, P, (Ed.)
25 *Frontiers in Econometrics*, pp. 105-142, Academic Press, New York.
- 26 McKelvey, R., and W. Zavoina, 1975, A statistical model for the analysis of ordinal level dependent
27 variables, *Journal of Mathematical Sociology*, 4, 103-120.
- 28 O'Hara, Maureen, 2007, Optimal microstructures, *European Financial Management Journal*, 852-832.
- 29 Pardo, Angel, and Roberto Pascual, 2007, On the hidden side of liquidity, *Working Paper*, SSRN Library.
- 30 Parlour, Christine A., 1998, Price dynamics and limit order markets, *Review of Financial Studies*, 789-
31 816.
- 32 Pascual, Roberto, and David Veredas, 2006, Does the limit order book matters in explaining long-run
33 volatility?, *SSRN Working Paper*.
- 34 Ranaldo, Angelo, 2004, Order aggressiveness in limit order book markets, *Journal of Financial Markets*,
35 7, 53-74.
- 36 Rosu, Loanid, 2005, A dynamic model of the limit order book, *SSRN Working Paper*.
- 37 Seppi, Duane, 1997, Liquidity provision with limit orders and a strategic specialist, *Review of Financial*
38 *Studies*, 10, 103-150.
- 39 Veall, Michael R., and Klaus F. Zimmermann, 1992, Pseudo-R²'s in the ordinal probit model, *Journal of*
40 *Mathematical Sociology*, 16, 332-342.
- 41 Veall, Michael R., and Klaus F. Zimmermann, 1996, Pseudo-R² measures for some common limited
42 dependent variable models, *Journal of Economic Surveys*, 10, 3, 241-259.
- 43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

TABLE I
Sample Statistics: Book and Activity

This table provides average daily descriptive statistics on the 36 stocks in our sample, from July to December 2000. "Quote midpoint" is the average between the best offer and bid quotes. "Spread" is the distance, in number of ticks, between the best ask and bid quotes. "Ask (Bid) depth" is the accumulated number of shares offered at the five best ask (bid) quotes on the LOB. "Ask (Bid) orders" is the accumulated number of limit orders supporting the best ask (bid) quotes on the book. Ask₁-Ask₅ (Bid₁-Bid₅) is the distance, in number of ticks, between the first and the fifth ask (bid) quotes on the book. "Daily Vol./1000 (Trades)" is the daily average share volume standardized by 1000(number of trades). The tick is 0.01 euros for all stocks and during all the sample period.

	Quote Midpoint	Spread (# ticks)	Ask Depth (Book)	Bid Depth (Book)	Ask Orders (Book)	Bid Orders (Book)	Ask ₁ -Ask ₅ (# ticks)	Bid ₁ -Bid ₅ (# ticks)	Daily Vol./1000	Daily Trades
ACR	9.30	3.17	10571.1	14832.1	18.36	11.84	5.83	5.78	277.85	370.48
ACS	27.11	13.70	4028.2	4392.9	7.01	7.69	19.22	18.79	133.83	249.15
ACX	32.09	13.07	3701.3	3926.0	8.03	7.40	18.91	19.57	175.34	328.65
ALB	27.55	19.03	3083.9	3232.3	6.31	6.44	23.59	21.30	136.75	188.77
ALT	16.31	4.05	12583.4	11611.8	8.60	12.94	8.05	9.07	1131.46	773.14
AGS	14.19	6.63	5316.6	5719.1	7.47	7.75	11.24	11.29	171.98	232.81
AMS	10.23	2.84	14416.6	16484.7	17.03	12.66	6.33	6.14	1177.81	924.65
ANA	38.48	14.68	2955.1	2883.1	7.20	7.01	22.24	21.73	125.52	288.21
AUM	16.63	8.90	6913.5	6626.0	7.69	9.65	12.75	18.70	102.09	76.23
BBV	16.07	1.71	84451.5	53025.4	13.90	39.14	4.84	4.96	6974.52	2023.90
BKT	44.19	11.73	3694.6	3734.3	9.17	7.99	18.67	16.34	224.68	559.02
CAN	20.97	12.06	8585.1	8583.9	6.45	8.07	15.58	18.93	146.01	148.24
CTG	18.94	7.29	6700.7	6860.1	8.53	8.99	10.93	10.97	372.38	369.15
DRC	9.94	3.60	18047.6	11783.3	9.62	15.56	6.69	7.81	624.81	440.19
ELE	20.80	2.45	24505.2	22732.6	10.07	12.75	6.34	6.34	3099.08	1496.06
FCC	19.56	8.63	5168.4	5081.2	8.01	8.35	13.52	13.90	183.95	322.99
FER	13.92	4.96	5974.4	6629.5	9.79	10.07	8.90	8.74	191.16	396.99
GPP	4.40	1.71	42577.5	40534.4	20.56	20.38	4.45	4.50	1107.38	620.10
IBE	13.80	2.57	31215.0	27039.0	11.22	15.85	5.74	6.20	2112.56	758.00
IDR	17.81	5.87	5819.6	5662.0	9.01	9.49	10.07	9.64	290.02	530.90
MAP	18.27	12.12	6629.1	5302.8	7.10	10.03	16.58	21.56	130.70	116.61
NHH	13.09	6.26	11189.7	9085.4	7.43	9.29	10.02	11.47	357.15	232.24
POP	34.72	8.57	8096.0	5174.1	7.43	12.55	13.66	14.05	395.66	463.60
PRS	23.68	7.95	5207.0	5825.4	8.76	11.72	12.52	14.06	382.35	554.59
REE	10.66	4.49	7034.9	9465.6	10.03	9.66	7.75	7.69	160.62	279.86
REP	20.32	2.41	25296.6	24654.2	13.76	12.37	6.14	5.99	3528.20	1655.46
SCH	11.45	1.34	178610.4	115052.2	28.81	97.43	4.33	4.31	9719.32	2962.41
SGC	33.37	12.24	3056.1	3041.5	8.58	7.21	18.59	17.10	202.47	545.57
SOL	10.92	4.70	7633.1	8448.2	9.47	8.42	8.27	7.89	280.60	292.52
TEF	21.90	1.56	65666.8	56251.7	23.83	19.05	4.90	4.67	19714.10	6608.87
TPZ	4.92	1.60	50389.1	68614.1	40.60	20.30	4.45	4.38	1433.03	847.68
TRR	34.53	4.75	8983.2	8675.3	15.11	11.48	9.59	8.37	2935.99	4596.82
TPI	8.83	2.41	14769.5	17494.5	17.24	12.05	5.61	5.39	1133.28	1008.90
UNF	20.47	4.40	14094.7	19545.4	8.21	9.20	8.89	7.72	768.06	413.85
VAL	6.71	3.01	14156.4	12739.4	10.09	11.35	5.95	6.49	334.15	253.67
ZEL	34.95	4.25	6826.7	5900.9	10.71	11.15	7.78	7.52	865.26	1776.36
Average	19.47	6.41	20220.79	17608.74	11.98	14.26	10.53	10.81	1697.23	936.30
Std. Dev.	(10.07)	(4.53)	(32720.8)	(23084.8)	(7.13)	(15.41)	(5.46)	(5.69)	(3677.81)	(1324.33)

TABLE II
Statistics: Aggressiveness

This table provides general statistics on order aggressiveness for the 36 SSE-listed stocks in our sample, from July to December 2000. We classify orders in terms of 7 categories of aggressiveness (C1 to C7). C1 are cancellations. C2 are limit orders to sell (buy) above (below) the best ask (bid) quotes. C3 are limit orders to sell (buy) hitting the best ask (bid) quote. C4 are limit orders within the best quotes. C5 are market orders for a lower size than the depth available at the best quote on the opposite side of the book. C6 are totally or partially executed orders that consume (only) the best quote on the opposite side of the book. C7 are totally or partially executed orders that consume more than one level of quotes on the opposite side of the book. In Panel A we report the percentage of all seven categories of orders over all the orders submitted for the 36 stocks in the sample. We provide the distribution of all orders (buyer plus seller-initiated) and also the distribution of buyer-initiated and seller-initiated orders separately. In Panel B, we report similar statistics for each category of buyer-initiated and seller-initiated orders but this time over the total number of buyer or seller-initiated orders submitted, respectively.

	Limit Orders				Small M.O. (C5)	Market to Limit (C6)	Large M.O. (C7)	Obs.
	Cancellat. (C1)	Ab./Bel. (C2)	At (C3)	Within (C4)				
A: Type/All orders								
All	14.41	13.45	11.05	10.89	40.39	6.07	3.75	7971432
Buyer initiated	7.49	6.93	5.69	5.73	21.37	3.04	1.81	4150054
Seller initiated	6.92	6.52	5.36	5.16	19.02	3.02	1.94	3821378
B: Type/Signed orders								
Buyer initiated	14.39	13.31	10.93	11.01	41.05	5.85	3.47	4150054
Seller initiated	14.43	13.59	11.19	10.76	39.67	6.31	3.77	3821378

TABLE III
First Stage of the SOP Model: Estimates and marginal probabilities

This table summarizes the estimation of the first stage of a sequential ordered probit (SOP) model for order aggressiveness with two stages. In this first stage, the dependent variable has three levels of aggressiveness: cancellations (withdrawal of liquidity), limit orders (provision of liquidity), and market orders (consumption of liquidity). In the second stage, the trader that has chosen to submit a limit order has to decide whether to place the limit order away from the best quotes, at the best quotes, or within the best quotes; the trader that chooses to submit a market order has to decide the size of his/her order: less volume than available at the best quote on the opposite side of the book, all that is available at the best quote on the opposite side of the book, or more volume than available at the best quote on the opposite side of the book. The estimation of the second stage is summarized in Table V. We report the median of the estimated coefficients for the 36 stocks in the sample, the percentage of statistically significant coefficients, and the percentage of statistically significant and positive coefficients. We provide separated results for buyer initiated orders and seller initiated orders. Additionally, we provide cross-sectional average marginal probabilities evaluated at the mean of the regressors. The marginal probabilities (x1000) measure the expected increase in the probability of each category after a marginal increase in a given regressor. The regressors, defined with respect to an incoming order, are: (1) Computed using the best ask and bid quotes: the bid-ask spread (SPR), the number of orders on the same side of the market (NS^1), and the number of orders on the opposite side of the market (NO^1). (2) Computed using from the 2nd to the 5th best quotes: the accumulated number of orders on the same side of the market NS^{25} , the accumulated number of orders on the opposite side of the market NO^{25} . (3) The distance between the best and the second best (the second best and fifth best) quotes on the same side of the market LS^{12} (LS^{25}), the distance between the best and the second best (the second best and fifth best) quotes on the opposite side of the market LO^{12} (LO^{25}). All models include one lag of the dependent variable. All the statistics are calculated across the 36 stocks in the sample.

1 st Stage - Withdrawing, providing or consuming liquidity										
Sellers	Agr(-1)	SPR	NS^1	NO^1	NS^{25}	NO^{25}	LS^{12}	LO^{12}	LS^{25}	LO^{25}
Median Coef.	0.2072	-0.0167	0.0279	-0.0332	0.0040	-0.0035	-0.0074	0.0215	0.0026	-0.0057
Significant (%)	100	94.44	97.22	91.67	72.22	52.78	63.89	94.44	50.00	61.11
Sig. & Positive (%)	100	0.00	97.22	2.78	69.44	8.33	25.00	94.44	27.78	16.67
Buyers										
Median Coef.	0.1859	-0.0203	0.0495	-0.0220	0.0056	-0.0012	-0.0155	0.0262	-0.0066	-0.0059
Significant (%)	100	100.00	94.44	88.89	88.89	58.33	63.89	94.44	69.44	50.00
Sig. & Positive (%)	100	0.00	91.67	0.00	83.33	11.11	30.56	94.44	27.78	13.89
Marginal probabilities (cross-sectional means)										
Sellers	Agr(-1)	SPR	NS^1	NO^1	NS^{25}	NO^{25}	LS^{12}	LO^{12}	LS^{25}	LO^{25}
Withdrawing	-38.607	8.226	-8.362	-1.045	-1.693	0.251	2.180	-5.399	0.757	0.234
Providing liquidity	-26.688	6.294	-3.937	-0.329	-0.765	0.158	2.604	-3.725	1.197	0.494
Consuming liquidity	65.296	-14.520	12.299	1.374	2.458	-0.409	-4.783	9.123	-1.954	-0.728
Buyers										
Withdrawing	-34.735	9.760	-11.952	-0.034	-2.764	0.185	2.302	-6.582	1.173	0.144
Providing liquidity	-18.223	6.218	-3.181	0.353	-0.666	0.161	2.757	-4.254	1.224	0.118
Consuming liquidity	52.958	-15.978	15.134	-0.319	3.430	-0.346	-5.058	10.836	-2.396	-0.263

TABLE IV
First Stage of the SOP Model: Performance

This table summarizes the results of an in-sample goodness-of-fit analysis and an out-of-sample predictive-ability analysis of three different specifications of the first stage of a sequential ordered probit (SOP) model for order aggressiveness. In its first stage, the dependent variable has three levels of aggressiveness: cancellations (withdrawal of liquidity), limit orders (provision of liquidity), and market orders (consumption of liquidity). In the second stage, the trader that has chosen to submit a limit order has to decide whether to place the limit order away from the best quotes, at the best quotes, or within the best quotes; the trader that chooses to submit a market order has to decide the size of his/her order: less volume than available at the best quote on the opposite side of the book, all that is available at the best quote on the opposite side of the book, or more volume than available at the best quote on the opposite side of the book. The three specifications considered are: the Baseline Model (BM), which only includes the first lag of the dependent variable in the set of explanatory variables; the Best Quotes Model (BQM), which adds variables computed from the best quotes of the book to the BM model; the Complete Book Model (CBM), which adds variables computed using from the 2nd to the 5th best quotes of the book to the BQM. The regressors computed using the best quotes are: the bid-ask spread (*SPR*), the number of orders on the same side of the market (*NS^s*), the number of orders on the opposite side of the market (*NO^s*). The regressors computed using the additional 4 quotes available of the book are: the accumulated number of orders on the same side of the market *NS²⁵*, the accumulated number of orders on the opposite side of the market *NO²⁵*, the distance between the best and the second best (the second best and fifth best) quotes on the same side of the market *LS²* (*LS²⁵*), the distance between the best and the second best (the second best and fifth best) quotes on the opposite side of the market *LO²* (*LO²⁵*). The in-sample analysis uses data from July to November 2000 and the out-of-sample analysis uses data from December 2000. The table reports the in-sample adjusted pseudo-R²s of McFadden (1973), Maddala (1983) with the Cragg-Uhler (1970) correction, Aldrich-Nelson (1984) with the Veall-Zimmermann (1992) correction, and McKelvey-Zavoina (1975). The out-of-sample adjusted McKelvey-Zavoina pseudo-R² is also provided. Finally, the table provides: (1) the percentage of stocks for which a given model is the best against using relative frequencies to predict the aggressiveness of the incoming order; (2) the percentage of stocks for which the CBM and the BQM outperform the BM on the basis that they usually allocate higher probabilities to the actual event than BM, and (3) the percentage of stocks for which the CBM outperforms the BQM on the basis of the same prediction rule.

Sequential Ordered Probit Model						
1st Stage: Withdrawing, consuming or providing liquidity						
In-sample Adjusted Pseudo-R ² s	Sellers			Buyers		
BM Pseudo-R ²						
McF	0.0097			0.0079		
MadCU	0.0220			0.0181		
AN	0.0283			0.0233		
MZ	0.0241			0.0198		
BQM (% increase over BM)						
McF	52.55			83.49		
MadCU	51.87			82.55		
AN	50.94			81.74		
MZ	51.68			82.75		
CBM (% increase over BQM)						
McF	49.80			41.91		
MadCU	51.48			43.06		
AN	51.17			42.60		
MZ	52.40			42.57		
Out-sample Adjusted Pseudo-R ² (MZ)						
	Sellers			Buyers		
BM Pseudo-R ²	0.0194			0.0156		
BQM (% increase over BM)	108.96			141.22		
CBM (% increase over BQM)	31.39			40.11		
Out-sample comparison						
	CBM	BQM	BM	CBM	BQM	BM
Best model against the relative frequency	47.22	22.22	30.56	47.22	16.67	36.11
Better performance than the BM	100	97.22		97.22	100	
Better performance than the BQM	94.44			97.22		

TABLE V
Second Stage of the SOP Model: Estimates and marginal probabilities

This table summarizes the estimation of the second stage of a sequential ordered probit (SOP) model for order aggressiveness with two stages. In the first stage, the dependent variable has three levels of aggressiveness: withdrawal of liquidity, provision of liquidity, and consumption of liquidity. The estimation of the first stage is reported in Table III. In the second stage, the liquidity provider (Panel A) has to decide whether to place the limit order away from the best quotes, at the best quotes, or within the best quotes; the liquidity consumer (Panel B) has to decide the size of his/her market order: less volume than available at the best quote on the opposite side of the book, all that is available at the best quote on the opposite side of the book, or more volume than available at the best quote on the opposite side of the book. A positive estimated coefficient means that the associated explanatory variable is positively related to order aggressiveness. We report the median estimated coefficient for the 36 stocks in the sample, the percentage of statistically significant coefficients and the percentage of statistically significant and positive coefficients. Additionally, we provide cross-sectional average marginal probabilities evaluated at the mean of the regressors. The marginal probabilities measure the expected increase in the probability of each category after a marginal increase in a given regressor. We provide separated results for buyer initiated orders and seller initiated orders. The regressors, defined with respect to an incoming order, are: (1) Computed using the best ask and bid quotes: the bid-ask spread (SPR), the number of orders stored on the same side of the book (NS^1), the number of orders stored on the opposite side of the book (NO^1), (2) Computed using from the 2nd to the 5th best quotes: the accumulated number of orders on the same side of the market (NS^{25}), the accumulated number of orders on the opposite side of the market (NO^{25}), (3) The distance between the best and the second best (the second best and fifth best) quotes on the same side of the market LS^{12} (LS^{25}), the distance between the best and the second best (the second best and fifth best) quotes on the opposite side of the market LO^{12} (LO^{25}). All models include one lag of the dependent variable.

Panel A: 2nd Stage - Limit order traders (Patient Traders)										
Sellers	Agr(-1)	SPR	NS^1	NO^1	NS^{25}	NO^{25}	LS^{12}	LO^{12}	LS^{25}	LO^{25}
Median Coef.	0.1017	0.0201	0.0733	0.0032	-0.0063	0.0059	-0.0838	-0.0074	-0.0081	0.0099
Significant (%)	94.44	88.89	88.89	8.33	91.67	22.22	94.44	16.67	66.67	19.44
Sig. & Positive (%)	94.44	88.89	88.89	8.33	0.00	22.22	0.00	2.78	2.78	13.89
Buyers										
Median Coef.	0.0966	0.0167	0.0934	-0.0015	-0.0076	0.0030	-0.0888	-0.0149	-0.0062	0.0105
Significant (%)	97.22	91.67	97.22	5.56	94.44	33.33	97.22	30.56	52.78	30.56
Sig. & Positive (%)	97.22	91.67	97.22	2.78	0.00	27.78	0.00	5.56	0.00	16.67
Marginal probabilities (cross-sectional means) (x1000)										
Sellers	Agr(-1)	SPR	NS^1	NO^1	NS^{25}	NO^{25}	LS^{12}	LO^{12}	LS^{25}	LO^{25}
Above the best quote	-33.986	-11.262	-21.943	0.315	3.252	-0.529	40.913	0.752	3.219	-1.496
At the best quote	-0.688	2.111	-1.741	0.348	0.396	-0.036	-5.711	0.358	-0.060	0.451
Inside the spread	34.674	9.150	23.684	-0.663	-3.648	0.565	-35.203	-1.110	-3.159	1.045
Buyers										
Below the best quote	-32.701	-9.672	-28.694	-0.283	3.346	-0.833	43.119	2.673	1.681	-0.562
At the best quote	-3.144	1.429	-4.307	-0.004	0.673	-0.149	-5.068	-0.159	0.091	0.274
Inside the spread	35.845	8.242	33.001	0.287	-4.020	0.982	-38.051	-2.513	-1.773	0.289
Panel B: 2nd Stage - Market order traders (Impatient Traders)										
Sellers	Agr(-1)	SPR	NS^1	NO^1	NS^{25}	NO^{25}	LS^{12}	LO^{12}	LS^{25}	LO^{25}
Median Coef.	0.1144	0.0424	-0.0099	-0.2878	-0.0020	0.0078	0.0123	-0.0212	0.0178	-0.0106
Significant (%)	97.22	66.67	27.78	100	58.33	63.89	36.11	63.89	66.67	55.56
Sig. & Positive (%)	97.22	66.67	2.78	0	8.33	61.11	33.33	0	63.89	5.56
Buyers										
Median Coef.	0.1079	0.0304	-0.0306	-0.2401	-0.0058	0.0079	0.0220	-0.0269	0.0145	-0.0076
Significant (%)	94.44	86.11	38.89	97.22	77.78	66.67	58.33	80.56	61.11	27.78
Sig. & Positive (%)	91.67	86.11	5.56	0	5.56	63.89	52.78	0	58.33	2.78
Marginal probabilities (cross-sectional means) (x1000)										
Sellers	Agr(-1)	SPR	NS^1	NO^1	NS^{25}	NO^{25}	LS^{12}	LO^{12}	LS^{25}	LO^{25}
Non aggressive	-30.580	-14.058	2.305	78.642	0.721	-1.773	-2.709	5.438	-3.822	1.853
Market-to-limit	14.876	7.802	-1.048	-38.792	-0.341	0.822	1.472	-2.733	2.111	-0.961
Agressive	15.704	6.256	-1.257	-39.849	-0.380	0.951	1.237	-2.704	1.711	-0.892
Buyers										
Non aggressive	-30.501	-16.578	5.159	71.845	2.758	-1.857	-4.369	9.654	-3.537	1.368
Market-to-limit	15.790	9.805	-2.535	-36.542	-1.426	0.909	2.502	-5.491	2.041	-0.801
Agressive	14.711	6.772	-2.624	-35.303	-1.331	0.948	1.866	-4.164	1.495	-0.568

TABLE VI
Second Stage of the SOP Model: Performance

This table summarizes the results of an in-sample goodness-of-fit analysis and an out-of-sample predictive-ability analysis for three different specifications of a sequential ordered probit (SOP) model of order aggressiveness. In its first stage, the dependent variable has three levels of aggressiveness: cancellations (withdrawal of liquidity), limit orders (provision of liquidity), and market orders (consumption of liquidity). In the second stage, the trader that has chosen to submit a limit order has to decide whether to place the limit order away from the best quotes, at the best quotes, or within the best quotes; the trader that chooses to submit a market order has to decide the size of his/her order: less volume than available at the best quote on the opposite side of the book, all that is available at the best quote on the opposite side of the book, or more volume than available at the best quote on the opposite side of the book. The three specifications are: the Baseline Model (BM), which only includes the first lag of the dependent variable in the set of explanatory variables; the Best Quotes Model (BQM), which adds variables computed from the best quotes of the book to the BM model; the Complete Book Model (CBM), which adds variables computed using from the 2nd to the 5th best quotes of the book to the BQM. The regressors computed using the best quotes are: the bid-ask spread (SPR), the number of orders on the same side of the market (NS^l), and the number of orders on the opposite side of the market (NO^l). The regressors computed using the additional 4 quotes available of the book are: the accumulated number of orders on the same side of the market NS^{25} , the accumulated number of orders on the opposite side of the market NO^{25} , the distance between the best and the second best (the second best and fifth best) quotes on the same side of the market LS^{12} (LS^{25}), the distance between the best and the second best (the second best and fifth best) quotes on the opposite side of the market LO^{12} (LO^{25}). The in-sample analysis uses data from July to November 2000 and the out-of-sample analysis uses data from December 2000. The table reports the in-sample adjusted pseudo- R^2 's of McFadden (1973), Maddala (1983) with the Cragg-Uhler (1970) correction, Aldrich-Nelson (1984) with the Veall-Zimmermann (1992) correction, and McKelvey-Zavoina (1975). The out-of-sample adjusted McKelvey-Zavoina pseudo- R^2 is also provided. Finally, the table provides: (1) the percentage of stocks for which a given model is the best against using relative frequencies to predict the aggressiveness of the incoming order (2) the percentage of stocks for which the CBM and the BQM outperform the BM on the basis that they usually allocate higher probabilities to the actual event than BM, and (3) the percentage of stocks for which the CBM outperforms the BQM on the basis of the same prediction rule. Table IV provides the results for the first stage of the SOP model.

2nd Stage Sequential Ordered Probit Model

In-sample Adjusted Pseudo- R^2 's	Passive Traders (Liquidity Providers)			Active Traders (Liquidity Consumers)								
	Sellers		Buyers	Sellers		Buyers						
BM Pseudo- R^2												
McF	0.0037		0.0031	0.0019		0.0022						
MadCU	0.0092		0.0078	0.0033		0.0041						
AN	0.0119		0.0100	0.0043		0.0054						
MZ	0.0107		0.0087	0.0043		0.0053						
BQM (% increase over BM)												
McF	131.04		177.32	1095.97		1005.21						
MadCU	134.62		176.86	1104.73		976.90						
AN	134.34		175.29	1088.45		956.08						
MZ	143.11		197.91	3739.82		3128.25						
CBM (% increase over BQM)												
McF	272.94		221.91	25.78		34.43						
MadCU	259.05		213.73	24.14		34.78						
AN	250.48		207.29	23.36		33.62						
MZ	267.34		202.79	1.85		13.54						
Out-sample Adjusted Pseudo- R^2 (MZ)	Sellers-M2		Buyers-M2		Sellers-M2		Buyers-M2					
BM Pseudo- R^2	0.0107		0.0086		0.0054		0.0772					
BQM (% increase over BM)	126.73		152.47		2097.78		1258.31					
CBM (% increase over BQM)	334.57		325.53		7.82		21.88					
Out-sample comparison	CBM	BQM	BM	CBM	BQM	BM	CBM	BQM	BM	CBM	BQM	BM
Best model against the relative frequency	91.67	8.33	0.00	88.89	11.11	0.00	38.89	22.22	38.89	27.78	19.44	52.78
Better performance than the BM	100	97.22		100	100		44.44	22.22		47.22	27.78	
Better performance than the BQM	100			100			38.89			50.00		

TABLE VII
SOP Model vs. OP Model: Performance

This table compares the performance of a two-stage sequential ordered probit (SOP) model for order aggressiveness with a one-stage ordered probit model (OP) model. In the first stage of the SOP model, the dependent variable has three levels of aggressiveness: cancellations (withdrawal of liquidity), limit orders (provision of liquidity), and market orders (consumption of liquidity). In the second stage, the trader that has chosen to submit a limit order has to decide whether to place the limit order away from the best quotes, at the best quotes, or within the best quotes; the trader that chooses to submit a market order has to decide the size of his/her order: less volume than available at the best quote on the opposite side of the book, all that is available at the best quote on the opposite side of the book, or more volume than available at the best quote on the opposite side of the book. The OP model considers seven categories of order aggressiveness: cancellations (C1), limit orders away from the best quotes (C2), limit orders at the quotes (C3), limit orders inside the quotes (C4), non-aggressive market orders (C5), market-to-limit orders (C6), and aggressive market orders (C7). Both models include the same set of explanatory variables: the first lag of the dependent variable, the bid-ask spread (SPR), the number of orders on the same side of the market (NS^j), the number of orders on the opposite side of the market (NO^j), the accumulated number of orders on the same side of the market NS^{25} , the accumulated number of orders on the opposite side of the market NO^{25} , the distance between the best and the second best (the second best and fifth best) quotes on the same side of the market LS^{12} (LS^{25}), the distance between the best and the second best (the second best and fifth best) quotes on the opposite side of the market LO^{12} (LO^{25}). The in-sample analysis uses data from July to November 2000 and the out-of-sample analysis uses data from December 2000. The table reports the number of stocks (up to 36) and the percentage of observations for which the SOP model outperforms the OP model in the sense that the SOP model it allocates a higher probability to the actual event than the OP.

Buys				Sells				
In sample		SOP		In sample		SOP		
		vs. OP	Obs.			vs. OP	Obs.	
	Stocks	36	3506233		Stocks	36	3242140	
	% Obs.	60.28%			% Obs.	59.63%		
Out of sample				Out of sample				
	Stocks	35	552053		Stocks	36	513959	
	% Obs.	58.72%			% Obs.	58.60%		
Buys - Aggressiveness				Sells - Aggressiveness				
In sample				In sample				
	C1	Stocks	22	515523	C1	Stocks	19	474179
		% Obs.	48.58%			% Obs.	47.52%	
	C2	Stocks	32	474259	C2	Stocks	33	446853
		% Obs.	52.70%			% Obs.	53.92%	
	C3	Stocks	36	389114	C3	Stocks	36	365332
		% Obs.	62.63%			% Obs.	62.66%	
	C4	Stocks	36	388340	C4	Stocks	36	344202
		% Obs.	60.73%			% Obs.	60.27%	
	C5	Stocks	36	1413969	C5	Stocks	36	1278047
		% Obs.	63.88%			% Obs.	62.79%	
	C6	Stocks	32	206444	C6	Stocks	35	204115
		% Obs.	64.57%			% Obs.	61.24%	
	C7	Stocks	36	118584	C7	Stocks	36	129412
		% Obs.	81.69%			% Obs.	79.60%	
Out of sample				Out of sample				
	C1	Stocks	29	92207	C1	Stocks	28	87503
		% Obs.	52.41%			% Obs.	52.14%	
	C2	Stocks	32	78144	C2	Stocks	32	72537
		% Obs.	54.70%			% Obs.	53.84%	
	C3	Stocks	35	64379	C3	Stocks	34	62124
		% Obs.	63.57%			% Obs.	70.00%	
	C4	Stocks	35	68622	C4	Stocks	35	66836
		% Obs.	60.43%			% Obs.	61.23%	
	C5	Stocks	23	193036	C5	Stocks	22	168404
		% Obs.	57.32%			% Obs.	55.93%	
	C6	Stocks	36	36124	C6	Stocks	32	36902
		% Obs.	67.51%			% Obs.	62.31%	
	C7	Stocks	35	19541	C7	Stocks	34	19653
		% Obs.	80.19%			% Obs.	75.84%	