

# Demand for Immediacy: Time Is Money<sup>1</sup>

Isabel Tkatch<sup>2</sup>  
School of Business Administration  
Hebrew University,  
Jerusalem, Israel 91905  
msisabel@mscc.huji.ac.il  
and

Eugene Kandel  
School of Business Administration  
and Department of Economics  
Hebrew University,  
Jerusalem, Israel 91905  
mskandel@mscc.huji.ac.il

This Draft: November 21, 2004

<sup>1</sup>We would like to thank Thierry Foucault, Ilan Guttman, Ohad Kadan, Daniel Passerman, Orly Sade, Yosi Yahav, as well as seminar participants at Hebrew University for helpful comments and suggestions. Financial support from the Israeli Science Foundation and Falk Institute for Economic Research is gratefully acknowledged. We would like to thank the Tel Aviv Stock Exchange for providing us with the data set, Miki Froimovich, for facilitating the process, and Roby Goldenberg and Dror Shalit for many useful discussions.

<sup>2</sup>Corresponding author.

## **Abstract**

### **Demand for Immediacy: Time IS Money**

Do traders in securities markets value immediacy of execution? Theoretical literature on dealer markets did not have to deal with this question, thus until recently this question was mostly ignored. Recent interest in the time-to-execution as measure of market quality in order driven markets coincided produced empirical research on this topic. Lack of theoretical models made it difficult to test the importance of immediacy. We build on recent theoretical developments to construct an estimation procedure that tests the effect of immediacy directly. Using data from the Tel Aviv Stock Exchange we show that the expected time-to-execution is an important determinant of order submission strategies. Moreover, the qualitatively similar results that we get for stocks and government bonds, that are traded in the same platform, suggest that immediacy considerations have a significant effect on the price process regardless of information asymmetry. Finally, our results corroborate predictions of theoretical models of order driven markets.

# 1 Introduction

Do traders in securities markets value immediacy of execution? Demsetz (1968) stated (p.41): “*Waiting costs are relatively important for trading in organized markets, and would seem to dominate the determination of spreads.*”. While Demsetz was referring to order-driven markets, the subsequent theoretical literature in Market Microstructure focused mostly on dealer markets. The main considerations for a dealer are the inventory risk (e.g. Stoll 1978), and the adverse-selection costs in an environment of rapidly changing fundamental values (e.g. Copeland and Galai 1983, Glosten and Milgrom 1985, Kyle 1985, Easley and O’Hara 1992). The waiting costs cannot and do not play a role in this setting since dealers always provide immediate execution.

Over the years, order-driven markets in various forms came to dominate equity trading around the world. This prompted a vast body of empirical research that analyzed data from order-driven markets using ideas derived from the theory based on dealer market environment, and interpreting the results accordingly. The basic premise was that competition among limit orders should generate approximately the same prices as those set by dealers. Consequently, as the theory did not address the issue of demand for immediacy, it did not play a role in most of these studies.

Dissatisfaction with the applicability of dealer-markets theory to order-driven markets prompted several authors to model order-driven markets explicitly. Glosten (1994), Chakravarty and Holden (1995), Rock (1996), Seppi (1997), Biais, Martimort and Rochet (2000) and Parlour and Seppi (2001), among others, focus on the optimal bidding strategies for limit order

traders in static models, in which the arrival of market and limit orders is determined exogenously. Yet, we know that in order-driven markets traders may choose endogenously whether to submit market orders that demand immediacy or limit orders that supply it. Parlour (1998), Foucault (1999), and Goettler, Parlour and Rajan (2004) study dynamic models with an endogenous choice between limit and market orders, but in neither model limit order traders bear waiting costs, i.e. care explicitly about the immediacy of execution.

Recently, time-to-execution received much attention from regulators; SEC now considers time-to-execution one of the parameters of market quality and requires brokerage firms to publish statistics on execution times. In recent years, several empirical papers (see below) studied time-to-execution in high-frequency trade data. These papers are not based on theoretical models of equilibrium in order-driven markets, thus cannot really test the importance of the demand for immediacy in these markets.

Foucault, Kadan, and Kandel (2004) (hereafter FKK) develop a dynamic model of a pure limit order book with fully strategic, liquidity motivated traders, who could be either patient or impatient (differing values of immediacy). They deliberately depart from the literature by assuming that the fundamental value of the asset is constant over time, thus removing the conventional drivers of book dynamics: the inventory cost and the adverse-selection cost. Instead, they make the demand for immediacy their traders' main concern, as in Demsetz (1968). The trade-off between the price-aggressiveness of an order and its time-to-execution is the key feature of the model. They find that order price-aggressiveness and time-to-execution are determined jointly in equilibrium by four parameters: the degree of im-

patience (costs of waiting), the proportion of the impatient traders in the population, the arrival rate of traders, and the tick size. Rosu (2004) develops a model with quite different assumptions, yet with the same basic idea of testing the effects of the demand for immediacy. He gets qualitatively similar results.

The restrictive assumptions of the FKK and Rosu (2004) models prevents them from being taken directly to the data. Yet, the intuition for any order-driven market is very clear: if demand for immediacy is indeed important to traders, longer expected time-to-execution must increase the price-aggressiveness of submitted orders. We tests this proposition.

We must use a set of simultaneous equations to address the endogeneity of the two variables of interest: the execution-time and the price-aggressiveness of the order. We must also account for selection bias, since we only observe the time-to-execution of order submission strategies that were chosen, but not of the counterfactuals. Finally, given the frequent cancellations of orders, we must account for censored observations. We adopt a three stage estimation methodology: the first stage is an ordered-probit model of the order submission strategy with three levels of price-aggressiveness: market; price-improving limit and non-improving limit. This stage allows us to derive regressors that correct for selection bias in the second stage estimation of time-to-execution. In the second stage we follow Lo, MacKinlay and Zhang (2002) (hereafter LMZ), who address the estimation problem of a positive and censored dependent variable using a lifetime regression model. In the third stage we use the estimated expected execution times, calculated on the basis of the second stage results, as additional explanatory variables in a structural ordered-probit model of the order submission strategy. While

there is a multicollinearity problem in certain specifications, in general we are able to identify the effects of the variables of interests.

We use the Tel Aviv Stock Exchange (TASE) data to estimate the model empirically. TASE is a pure limit order market without intermediaries or hidden orders, which fits the assumptions of the theoretical models. In TASE, stocks and bonds are traded on the same platform with minor differences in hours and trading rules. This feature enables us to carry out the same analyses for stocks and bonds, and compare the results. TASE made all its data files (orders, cancellations, trades with cross-references to orders) available to us, so we are able to avoid problems that arise from using small-size or aggregate data sets. A potential concern is that TASE is a far less liquid market than the US or some European markets. We test whether this fact systematically affects the dynamics of order submission by replicating LMZ (2002) analyses for our stocks data, and obtain very similar results<sup>1</sup>. This suggests that lessons learned on TASE extend beyond this market.

The model ignores information asymmetry, which dominate the microstructure literature. We fully agree with O'Hara (2003), that adverse selection is an important factor in price dynamics. However, we conjecture that in order driven markets demand for immediacy plays an important role as well. Our comparison between stocks and bonds that are traded on the same platform, allows us to test the importance of immediacy in different regimes of information asymmetry. Government bonds, after all, are much less prone to asymmetric information than stocks.

Our results corroborate the main prediction of Demsetz (1968): *time is*

---

<sup>1</sup>A detailed discussion of the LMZ (2002) model and a comparison of TASE results to those obtained for the ITG dataset are presented in **Appendix 2**.

*money* for traders on TASE, i.e. traders are willing to pay money (improve on the price) to get a reduction in the expected time-to-execution. Moreover, the qualitatively similar results we get for stocks and government bond strengthen our claim that liquidity based considerations have a significant effect on the price process regardless of the degree of information asymmetry. Finally, our results corroborate other empirical predictions of the FKK (2004) model.

Section 2 surveys the literature on time-to-execution and order price-aggressiveness. Section 3 sets up the econometric model and discusses its intuition. Section 4 describes the data and sample, and defines the variables of interest. Section 5 presents the results and section 6 concludes. A detailed discussion of the econometric model and of robustness issues is delegated to the appendix.

## **2 Literature Survey**

Time-to-execution has been investigated recently in several empirical studies. Engle and Russell (1998) and Engle (2000) develop and estimate autoregressive conditional duration models. Since order data is usually unavailable in the US trading systems, these models concentrate on the duration between trades and on the characteristics of the transaction process such as the transaction stock price and volatility. Handa and Schwartz (1996), among others, try to extract the execution time of an order from transaction time and price data assuming that the order is executed as soon as a suitable transaction price is observed, yet Lo, MacKinlay and Zhang (2002) show that this type of procedure does not produce a good proxy for the true time-to-execution. Hasbrouck and Saar (2002) use a dataset provided by

Island (ECN) where order submission time is available. They test whether asset price volatility affects the order price-aggressiveness and its time-to-execution, but do it in separate and unrelated models. Lo, MacKinlay and Zhang (2002) use a unique dataset from ITG, an institutional brokerage firm, where one can directly observe order submission information. They estimate an econometric model of time-to-execution, claiming that it is an important dimension of market quality in limit order markets (SEC 1997). They use proxies for the aggressiveness of limit orders as explanatory variables for time-to-execution, as if they were exogenously determined. This approach is not consistent with an equilibrium model, where both time-to-execution and price-aggressiveness of orders are determined endogenously.

Hollifield, Miller, Sandas and Slive (2003) (hereafter HMSS) also use a duration model to estimate time-to-execution and derive the probability of execution of various orders. Their main interest is in the second stage results where they estimate the distribution of private values and the arrival rate of traders. Even though HMSS (2003) do not state it explicitly, one could write their model as a set of simultaneous equations. While they set their exercise in a rich theoretical model in which traders' choices depend on their private values, they do not estimate the value of immediacy. Moreover, their empirical estimation requires many structural assumptions.

Order submission strategies (or price-aggressiveness) are analyzed in many empirical papers, usually in the context of the explicit transaction costs associated with these strategies rather than their expected execution time. Harris and Hasbrouck (1996) evaluate performance measures for market and limit orders, and compare those for fixed order size, side and spread groups. They do not consider time-to-execution. Griffiths, Smith, Turnbull

and White (2000) use order submission strategy and its interaction with other “decisions” and exogenous variables, as explanatory variables in a regression model of price impact. This approach ignores the fact that traders take the expected price impact into account when choosing an order. Ellul, Holden, Jain and Jennings (2003) model the order submission strategy as the dependent variable in a multinomial logit model. They use stock characteristics and market conditions as explanatory variables, but ignore the subsequent cost aspect of the order strategy.

Given the results of FKK and Rosu (2004) models, it is clear that if the demand for immediacy is important, one cannot separate the order submission strategy from its time-to-execution. On the other hand, given that the two variables are endogenously determined in equilibrium, one variable cannot be used as an explanatory variable for the other. Moreover, the execution-time is observed only for choices that were actually made, which requires us to deal with a selection bias in this variable, and finally, order cancellations force us to deal with censored data. In the following section we present an econometric model that addresses these issues.

### **3 A Structural Model of Expected Execution Times and Order Submission Strategy**

We build on a switching regression model with endogenous switching described in Maddala (1983) and modify it to fit censored lifetime variables, as required for the analysis of our data. In this section we sketch the econometric model, while the detailed derivation is delegated to **Appendix 1**. The endogenous variables in the model are: order submission strategy (or price-aggressiveness), which is transformed into a discrete variable with three

levels (market order, price-improving limit and non-improving limit order), and the associated times-to-execution variables. Note, that since the expected time-to-execution of a market orders is zero by definition we only have two lifetime variables for price-improving and non-improving limit orders.<sup>2</sup>

Consider a sequence of  $i = 1, \dots, n$  limit orders, and denote by  $C_i$  a latent variable representing the price-aggressiveness level of order  $i$ . We assume that the price-aggressiveness depends on the expected time-to-execution and on a set of exogenous explanatory variables denoted by  $Z$ . We denote by  $Y_1$  the log expected time-to-execution of price-improving limit orders, and by  $Y_2$  the log expected time-to-execution of limit orders that do not improve on the quoted price. We assume that both lifetime variables depend on the same set of exogenous explanatory variables, denoted by  $X$ , which may overlap the set  $Z$ .

Let

$$Y_{1i} = \beta_1' X_i + u_{1i},$$

$$Y_{2i} = \beta_2' X_i + u_{2i},$$

and

$$C_i = \gamma' Z_i + \delta_1 Y_{1i} + \delta_2 Y_{2i} - u_i,$$

where

$$\begin{bmatrix} u_{1i} \\ u_{2i} \\ u_i \end{bmatrix} \sim N(0, \Sigma), \quad \Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{1u} \\ & \sigma_2^2 & \sigma_{2u} \\ & & \sigma_u^2 \end{bmatrix}.$$

---

<sup>2</sup>While we use the term “time-to-execution” throughout this paper, a more accurate title name for the endogenous lifetime variables is “time-to-first-fill”, which is the time till the first share of the order is executed. As we show below, this allows us to assume that the order size has no direct effect on the time to execution.

The system of equation above is defined to capture the trade-off between execution time and price. We have one equation for each expected log execution-time,  $Y_1$  and  $Y_2$ . Each variable is observed only in the case the associated strategy is chosen. Given the price priority rules, it is obvious that the execution-time of an aggressive limit order is shorter than that of a non-aggressive one, which implies that  $\widehat{\beta}'_1 X \leq \widehat{\beta}'_2 X$  for every order  $i$ . Furthermore, following FKK, we expect traders' price-aggressiveness level to increase in the expected time-to-execution of a less aggressive order strategy. Thus, for example, the probability of a market order should be higher when the trader expects relatively high execution-times for limit orders. This assertion should result in positive signs of the coefficients of the expected lifetime variables,  $\widehat{\delta}_1, \widehat{\delta}_2 > 0$ .

Neither of the dependent variables defined above,  $Y_1$ ,  $Y_2$ , and  $C$ , are directly observable. Instead, we observe the following system:

$$I_i = \begin{cases} 0 & \text{if } 0 < C_i < +\infty \text{ (market order - most aggressive)} \\ 1 & \text{if } -\mu_2 < C_i \leq 0 \text{ (price-improving limit order)} \\ 2 & \text{if } -\infty < C_i \leq -\mu_2 \text{ (non-improving limit order)} \end{cases},$$

and

$$Y_i = \begin{cases} 0 & \text{if } I_i = 0 \\ Y_{1i} & \text{if } I_i = 1 \\ Y_{2i} & \text{if } I_i = 2 \end{cases}.$$

Moreover, since a significant percentage of the orders are cancelled before execution, we do not observe  $Y_i$ , but rather a mixture of truncated and non-truncated lifetime values. Following the study of LMZ (2002), we assume that all cancellations, which take place before any fraction of the order is executed, represent a naive censoring of the lifetime values. Finally, since only a discrete version of the price-aggressiveness level is observable, the

parameters  $\gamma$ ,  $\delta_1$ , and  $\delta_2$  are estimable only up to a proportionality factor. We denote those variables by  $\gamma^*$ ,  $\delta_1^*$ , and  $\delta_2^*$  respectively, and the appropriate criterion variable is denoted by  $C^*$ .

The estimable form of our model is as follows:

$$I_i = \begin{cases} 0 & \text{if } -\infty < u_i^* < \gamma^* Z_i^* + \delta_1^* Y_{1i} + \delta_2^* Y_{2i} \\ 1 & \text{if } \gamma^* Z_i^* + \delta_1^* Y_{1i} + \delta_2^* Y_{2i} \leq u_i^* < \gamma^* Z_i^* + \delta_1^* Y_{1i} + \delta_2^* Y_{2i} + \mu_2^* \\ 2 & \text{if } \gamma^* Z_i^* + \delta_1^* Y_{1i} + \delta_2^* Y_{2i} + \mu_2^* \leq u_i^* < +\infty \end{cases},$$

$$Y_i = \begin{cases} 0 & \text{if } I_i = 0 \\ \beta_1' X_i + u_{1i} & \text{if } I_i = 1 \text{ and } S_i = 0 \\ \beta_2' X_i + u_{2i} & \text{if } I_i = 2 \text{ and } S_i = 0 \end{cases}$$

and

$$S_i = \begin{cases} 1 & \text{if observation } i \text{ is censored} \\ 0 & \text{if observation } i \text{ is not censored} \end{cases},$$

where<sup>3</sup>

$$\begin{bmatrix} u_{1i} \\ u_{2i} \\ u_i^* \end{bmatrix} \sim N(0, \Sigma^*), \quad \Sigma^* = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{1u^*} \\ & \sigma_2^2 & \sigma_{2u^*} \\ & & 1 \end{bmatrix},$$

$$C_i^* = \frac{C_i}{\sigma^*}, \quad \gamma^* = \frac{\gamma}{\sigma^*}, \quad \delta_j^* = \frac{\delta_j}{\sigma^*} \quad (j = 1, 2), \quad u_i^* = \frac{[u_i - \delta_1 u_{1i} - \delta_2 u_{2i}]}{\sigma^*},$$

and

$$(\sigma^*)^2 = \text{Var}[u_i - \delta_1 u_{1i} - \delta_2 u_{2i}].$$

We use a three stage estimation procedure to get consistent estimates for the unknown parameters of our model, those of the criterion function ( $C^*$ ) that determines the order submission strategy, and those in each conditional lifetime equation ( $Y_1$  and  $Y_2$ ). Since the order submission strategy is a discrete ordered variable, we use an ordered probit model to estimate the

---

<sup>3</sup>Note that if  $S_i = 1$  we can only say that  $Y_i > \beta_j' X_i + u_{1j}$ ,  $j = 1, 2$ .

parameters of the criterion function. As for the lifetime variables, we follow LMZ (2002) in using a lifetime regression model.

In the first stage we estimate the parameters of the reduced-form of the criterion function:

$$C_i^* = \gamma^{*'} Z_i + (\delta_1^* \beta_1 + \delta_2^* \beta_2)' X_i - u_i^*.$$

Each lifetime variable is only observable for a sub-sample, thus we use the estimates obtained in the first stage to correct the for selection bias. We calculate the appropriate weights,  $W_{1i}^*$  and  $W_{2i}^*$ , (described in detail in the appendix) and use them as additional explanatory variables in the following two lifetime equations:

$$Y_i = \begin{cases} \beta_1' X_i + \sigma_{1u^*} W_{1i}^* + \varepsilon_{1i} & \text{if } I_i = 1 \\ \beta_2' X_i + \sigma_{2u^*} W_{2i}^* + \varepsilon_{2i} & \text{if } I_i = 2 \end{cases},$$

where we assume that cancellations represent naive censoring of the lifetime variables, which has lognormal distribution. Finally, in the third stage, we use the estimates of the expected-lifetime for each strategy,  $\widehat{\beta}_1' X$  and  $\widehat{\beta}_2' X$ , which are calculated on the basis of the second stage results. We estimate the parameters of the following structural-form of the criterion function as an ordered probit model:

$$C_i^* = \gamma^{*'} Z_i + \delta_1^* \widehat{\beta}_1' X_i + \delta_2^* \widehat{\beta}_2' X_i - u_i^*.$$

Note that the set of exogenous variables affecting the order price-aggressiveness,  $Z$ , and the set of variables affecting the time-to-execution,  $X$ , must not overlap completely for us to be able to identify the relevant parameters.

We use the switching lifetime regression model with endogenous switching described above to answer the following questions:

1. Whether the expected saving in time-to-execution increases order aggressiveness,  $\delta_1^*, \delta_2^* > 0$ ?
2. Whether the result in (1) depends on the degree of information asymmetry, i.e. whether the results are similar for the stocks and the government bonds, which are traded on the same platform.
3. What are the effects of other parameters of the FKK (2004) model that we incorporate in the vectors of the exogenous explanatory variables for the expected time-to-execution ( $X$ ) and for the order submission strategy ( $Z$ ). In particular, we are interested in the effects of competition between the suppliers of liquidity and the arrival rate of traders on the expected time-to-execution, and in the effects of spread, tick size and time on the price-aggressiveness level of the submitted order.

## 4 Data, Sample and the Choice of Controls

### 4.1 TASE Trading System

Trading on the Tel Aviv Stock Exchange (TASE) consists of three phases: The trading of shares starts with a call auction at 9:45 AM; a continuous phase between 9:45 and 16:45 and a closing phase between 16:45 and 17:00. Each phase is characterized by a different trading mechanism. The day starts with an empty order book at 8:30, when traders start submitting either limit or market orders, and the best three prices on each side are presented continuously starting at 9:00. At 9:45 orders are crossed using the call auction mechanism, with time and price priority rules, and the continuous trading phase begins. The continuous trading phase on the TASE is a pure limit order market without intermediaries. Traders may post either

market or limit orders, and those are executed according to price and time priority rules. At 16:45 the closing phase begins; it is a simple crossing of traders' market orders that are executed, if possible, at the closing price. All unexecuted orders are cancelled at the end of the trading day, and the next day's book is empty till the first orders come in.

Some companies from the sample are cross-listed on European and US markets, usually on NASDAQ. Nevertheless, due to time differences, trading of those stocks is simultaneous on TASE and on US venues only for the last hour of the day.

Equities, bonds, index options and futures contracts in TASE are all traded on the same platform, with some minor differences in hours, minimum order size and tick size. The continuous trading phase for bonds starts at 10:45 and ends at 16:35, which allows traders of the different instruments to devote exclusive attention to each instrument at the time of the opening auction and of the closing procedure.

## 4.2 Data and Sample

We use order submission, cancellation and trading data, for 32 stocks and 45 government bonds traded on the TASE, for a period of three months: May 1 2000 to July 31 2000. The sample of stocks includes those with the largest trading volume<sup>4</sup> that had no tick size changes during the sample period. The list of stocks and some descriptive statistics are presented in **Table 1S**. The sample of bonds includes those that were traded during the entire sample period and were the most liquid<sup>5</sup>. All are government bonds which belong

---

<sup>4</sup>All the stocks in our sample are included in the TA100 index and most are in TA25. Both indexes are based on the trading volume as a measure of liquidity.

<sup>5</sup>Each bond had a 3 months total trading volume of at least 10,000,000 NIS, a median daily trading volume of at least 50,000 NIS and a median daily number of transactions of

to one of three categories: nominal bonds, inflation-linked, and dollar-linked bonds<sup>6</sup>. The list of the bonds and some descriptive statistics are presented in **Table 1B**. Note that all bonds have a tick size of 0.0001NIS, which is finer than the minimal tick size category for stocks<sup>7</sup>.

Each order in our dataset is time stamped and matched with the book status prior to it's submission and with the subsequent trades and/or cancellation. Note that even though we have data of all trades for each order, in this study we concentrate on the analysis of the duration till the first transaction or cancellation (time-to-first-hit). About 2.6% (14,773) of the total number of observations in the sample of stocks and 3.5% (3,124) of the total number of observations in the sample of bonds are erroneous, and thus they are excluded from our analyses. Most errors are due to missing observations in the trades file supplied by the exchange.

In the continuous trading phase traders may submit two types of orders, market and limit, but pure market orders are very rare (less than 0.5% of the orders in the sample of stocks and less than 0.2% in the sample of bonds). Instead, traders submit marketable limit orders (limit buy orders priced above the best ask or limit sell orders priced below the best bid), and

---

<sup>6</sup>Inflation linked bonds are very popular in Israel, despite the current low inflation. It is a remnant of high inflation times in the 1980's.

<sup>7</sup>The government bonds in our sample can be organized in the three categories mentioned above. 'Makam', 'Gilon' and 'Shahar' bonds pay nominal principal and interest. The 'Makam' is a discount bond of up to one year maturity. 'Gilon' is issued for a term of up to 10 years, and pays a floating rate coupon twice a year. 'Shahar' is issued for a term of 1 year or more, up to 20 years, and pays a coupon once a year (fixed rate). The principal and interest payments of the 'Kfir', 'Galil' and 'Sagi' bonds are linked to inflation. 'Kfir' is issued for a term of up to 25 years, and pays a floating rate coupon once a year. 'Galil' is issued for a term of up to 20 year, and pays a fixed rate coupon once a year. 'Sagi' is issued for 5 or 6 years, and pays a fixed rate coupon once a year. 'Gilboa' is a dollar-denominated bond of 5 year maturity that pays a floating rate coupon linked to LIBOR.

for the purposes of our analyses we treat these orders as if they were market orders. In **Table 2** we present the distribution of order type by order side, and while both samples are relatively balanced with respect to order side, more than 60% of the observations are limit orders.

All unexecuted orders on the TASE are cancelled automatically at the end of the trading day yet, some are actively cancelled by traders during the day. The distribution of cancellations by cancellation type is presented in **Table 3**. In the sample of stocks we observe about 192,235 cancellations (58% of the total number of submitted limit orders) while in the sample of bonds the percentage cancelled limit orders is lower (40,623 or 47%) and there are fewer active cancellations. The distribution of time to cancellation by cancellation type is presented in **Table 4**. On average, active cancellations occur much faster than end-of-the-day cancellations.

The distribution of time-to-first-fill, for various order submission strategies (levels of price-aggressiveness), is presented in **Table 5**. The percentage of price improving limit orders in the sample of stocks is relatively low (26% of the limit orders) and, as expected, they stay in the book (till execution or cancellation) for less than 1/3 of the time that the non-improving limit orders do. In the sample of bonds, we observe price-improving limit orders more frequently (46%) yet, on average, all limit orders in the sample of bonds spend more time in the book before they are filled or cancelled (124 versus 87 minutes). This difference may be the result of the lower rate of cancellation and, in particular, the lower percentage of active cancellations in the sample of bonds.

In summary, the size of the sample of government bonds is smaller than that of stocks (86,245 versus 548,957 observations), there are no special

differences in the distributions of order side and order type in both samples, the proportion of limit order cancellations in the sample of bonds is smaller than that in the sample of stocks (47% versus 58%), and most cancellations in the sample of bonds occur at the end of the trading day (58% versus 35%). On average, time-to-first-fill is longer for bonds, especially when we examine price-improving limit orders (96 versus 33 minutes).

### 4.3 Definitions of Variables

First, let us introduce some *notations*:

$p$  = order price (NIS)

$q$  = order size (# of shares)

$p_{a1}$  = best quoted price on the sell side (lowest ask)

$p_{b1}$  = best quoted price on the buy side (highest bid)

$q_{a1}$  = quantity (# of shares) at the  $p_{a1}$  (lowest ask) price level

$q_{b1}$  = quantity (# of shares) at the  $p_{b1}$  (highest bid) price level

Next, let us define the following *dependent variables*:

$$I = \begin{cases} 0 & \text{if market order or a limit order, that improves the price} \\ & \text{by more than 50\% of the spread size} \\ 1 & \text{if limit order that improves the price} \\ & \text{by 50\% of the spread size or less} \\ 2 & \text{if non-improving limit order} \end{cases}$$

is a discrete variable measuring levels of order aggressiveness, where market is considered to be the most aggressive and non-improving limit is considered the least aggressive order submission strategy. Obviously, there is more than one way to define the categories of this variable and we discuss this issue in detail in **Appendix 1**. We would like to stress that our goal was to get a reasonable definition. We would like to capture different levels of price

aggressiveness, without losing the ability to draw meaningful conclusion from the analysis and without losing too many observations in the process<sup>8</sup>.

We denote by  $Y_j$  the log time to first-fill for a limit order of type  $j = 1, 2$ . We measure the lifetime of the order, in minutes, from the moment it is received on the exchange till the first share is executed (in contrast to completion time, where the whole quantity is executed). A cancellation of the limit order, before any fraction is executed, is considered as naive censoring of the lifetime variable.

Finally, let us define the *explanatory variables* for buy-side orders:

*Measures of the competition between suppliers of liquidity*

*compete* =  $\log(p_{a1}q_{a1})$  is the log of the monetary size of the queue at the best bid (competing / same-side volume of limit orders in the book).

*opposed* =  $\log(p_{b1}q_{b1})$  is the log of the monetary size of the queue at the best ask (opposed-side volume of limit orders in the book).

Our prediction for *compete* and *opposed* are similar to those in Parlour (1998) queueing model. The "crowding out" effect of large same-side depth in the book discourages traders from posting non-aggressive orders. If the buy-side depth is high, buyers expected execution time of limit orders to be high. On the other hand, when opposite-side depth is high, the "crowding out" effect should influence opposite-side orders in the same manner. If the sell-side depth is high, buyers expect that subsequent sell-side orders will be

---

<sup>8</sup>We have tried 4 levels of price aggressiveness, where price-improving limit orders are further divided into two categories, 1 tick improvements and aggressive ones. The results for that analysis are similar, except for a multicollinearity problem which is discussed in more detail in **Appendix 1**.

aggressive, thus the high sell-side depth should decrease the expected execution time of a buy-side limit order. High values of the variable *compete* and low values of *opposed* should both lead to a more aggressive order submission strategy. The volume of past unexecuted limit orders is also a proxy for the proportion of patient traders in the population in the FKK (2004) theoretical model. Past unexecuted limit orders in the book, *compete*, compete for liquidity with any non-aggressive limit orders that will be submitted in the future. This interpretation also implies a positive effect of the variable *compete* on the level of order strategy aggressiveness.

*Measures of the order arrival rate and order-side imbalance*

*arrivalrate* is the log of the monetary volume of orders, submitted on the opposite side, in the last half an hour. This variable represents the opposite-side order arrival rate. *arrivalrate* is a proxy for the parameter  $\lambda$  of the FKK (2004) theoretical model, which denotes the arrival rate of orders. Traders expect a time-to-execution to be long when the arrival rate of orders on the opposite-side is low (small values of *arrivalrate*).

$$sameside = \begin{cases} 1 & \text{if previous order was a buy order} \\ 0 & \text{otherwise} \end{cases}$$

is a dummy variable that is equal one if the previous and current submitted orders are both buy-side orders.

*imbalance* is the difference between the volume of same-side and opposite-side orders, normalized by the total volume measured in the last half hour.

Both *sameside* and *imbalance* measure order imbalance. The variable *imbalance* takes values between -1 and +1, which represent the magnitude of order imbalance. For a buy side order, *imbalance* is positive when the total

volume of buy-side orders is larger than that of sell-side orders, and negative when the situation is reversed. The variable *sameside* is an instantaneous measure of the same property. If the previous and current orders are both buy-side it is probable that there are more buyers than sellers in the market at the time the order is submitted. Buy-side traders expect longer time-to-execution when there is larger volume on the buy-side than on the sell-side (positive value of both *imbalance* and *sameside*).

*Measures of the trade-off between time and price aggressiveness*

$spread = \log(p_{a1} - p_{b1})$  is the log of the monetary size of the inside spread before the order is submitted. We do not expect traders to post market orders or make aggressive price improvements when the spread is large.

$$tick = \begin{cases} 0.1 & \text{if } p < 5_{NIS} \\ 1 & \text{if } 5 \leq p < 50 \\ 10 & \text{if } 50 \leq p < 500 \\ 100 & \text{if } p \geq 500 \end{cases}$$

is the minimum price improvement (in Agorot, where 1NIS = 100 Agorot).

Note that both variables, *spread* and *tick*, should be included in the model. For a given tick size, higher monetary size of the bid-ask spread means that the size of the spread measured in ticks is relatively large. For a given monetary spread size, large tick size means that the minimal price improvement is costly. Given the setting of FKK (2004) theoretical model, we expect traders to post less aggressive orders when an aggressive order strategy is costly, i.e. when the spread and the tick size are large.

$\widehat{\beta}'_j X = \widehat{Y}_j - \widehat{\sigma}_{ju} \widehat{W}_j^*$ ,  $j = 1, 2$  are the estimates of the expected lifetimes of strategy  $j$  limit orders. The values of this variable are calculated for each

observation in our sample, based on the second stage estimation results.

As we already mentioned before, given the FKK (2004) theoretical model setting, we expect traders' aggressiveness level to rise when the expected execution time of a less aggressive order strategy is long, and when the expected execution time of an aggressive order strategy is short.

*Control variables for intraday and day-of-the-week activity patterns*

$$daytime = \begin{cases} 1 & \text{if order is posted at 9:45 - 10:45} \\ 2 & \text{if order is posted at 10:45 - 11:45} \\ \vdots & \\ 7 & \text{if order is posted at 15:45 - 16:45} \end{cases}$$

is a discrete variable that indicates the time of order submission during the continuous trading phase. The variable is transformed into six dummy variables that control for an intraday effect.

$$weekday = \begin{cases} 5 & \text{if last day of the trading week} \\ 1 & \text{if first day of the trading week} \\ 3 & \text{otherwise (midweek)} \end{cases}$$

is a discrete variable that indicates the day of the week. The variable is transformed into two dummy variables that measure the day of the week effect.

Many empirical studies document intraday and daily patterns for variables such as spreads volumes and volatility. We would like to control for the effects of those time patterns in our model, and also to examine whether they have an impact on execution-times and order-submission-strategies.

*Cross sectional control variables*

The following variables are designed to capture the variation across stocks, since we use all orders and all stocks in one pool.

*price* is log of the average opening stage price (NIS) of the stock.

*volume* is log of the average daily volume (NIS).

*volatility* is the average of daily price range normalized by the range midpoint. The range is the difference between the maximum and minimum midpoint prices at the continuous trading phase. This measure of volatility is designed that way to overcome the bias caused by both, the bid ask bounce and the tick size regime (Hau 2003).

*TA25* is a dummy variable that is equal one if the stock is included in the TA25 index (25 largest stocks on the TASE).

*duallist* is a dummy variable that is equal one if the stock is also traded on some other exchange (usually NASDAQ).

Out of the 32 stocks in our sample, that had the highest trading volume and no changes in the tick size during the sample period, 18 are included in the TA25 index and 11 are traded on some other exchange (6 stocks are included in both groups).

#### *Daily control variables*

The following variables are designed to capture the variation of market conditions across days, again, since we use all days in our sample in one pool.

*Zvolume* is the  $Z$  score of the daily volume. The daily volume is calculated for each stock and day, and then normalized using the sample period mean and standard deviation of the daily volume for that stock.

*Zvolatility* is the  $Z$  score of the daily price range described above. The price range is calculated for each stock and day, and then normalized using

the sample period mean and standard deviation of the daily price range for that stock.

*Other variables*

$$sell = \begin{cases} 1 & \text{if sell side order} \\ 0 & \text{if buy side order} \end{cases}$$

is a dummy variable which indicates the order side. In some empirical studies, sellers are documented to be more aggressive than buyers.

$ordersize = \log(pq)$  is the log of the monetary order size (NIS).

We should use the order size as a control variable, since most theoretical models deal with fixed size orders. The effect of this variable on aggressiveness could go either way. A large order may drive the trader to ensure execution by using an aggressive strategy. On the other hand, such a trader might not want to expose his demand for liquidity and make a costly price concession.

$$round = \begin{cases} 1 & \text{if } \text{mod}(\frac{p_{b1}/tick}{5}) = 0 \\ 0 & \text{if } \text{mod}(\frac{p_{b1}/tick}{5}) > 0 \end{cases}$$

is a dummy variable that indicates whether the quoted price is a round number (a multiple of 5) or not. Since round numbers are observed more frequently than others for the spread size, we assume that traders are also more likely to choose round prices for limit orders. Thus, if the quoted price is indeed a round number, we expect traders to use that price, which will result in a negative effect on the aggressiveness level of the submitted order. Alternatively, if our traders are perfectly rational, this variable shouldn't have a significant effect on their order strategy.

*Special definitions for the sample of bonds*

Most variables, their definitions and their predicted effects are the same for both samples, stocks and government bonds. Yet there are a couple of differences. All government bonds in our sample are traded exclusively on the TASE and are not included in any stock index, thus the dummy variables *TA25* and *duallist* do not appear in the bonds model. The variable *tick* is also missing from that model, since all bonds in our sample have the same tick size. On the other hand, the discrete variable characterizing the nature of the bond *bondtype*, is not relevant for the stocks model.

$$bondtype = \begin{cases} Nom & \text{if the interest payment is nominal} \\ Ind & \text{if the interest payment is linked to inflation} \\ Dollar & \text{if the interest payment is linked to the dollar} \end{cases}$$

is a discrete variable that will be transformed into two dummy variables.

The last difference between the two models arises from the fact that bonds start trading later than stocks. Thus, the variable *daytime*, which indicates the time of order submission during the continuous trading phase, has only six categories in the bonds model.

## 5 Results

The results of the three stage estimation are presented in **Tables 6-8**. Note that the model parameters are identifiable since the two sets of explanatory variables, those included in the execution-time equations and those included in the equation of price-aggressiveness, are not identical.

The variables *ordersize round spread* and *tick* are excluded from the execution-time equations, while *compete opposed arrivalrate sameside* and *imbalance* are excluded from the equation of price-aggressiveness. We assume that order size has a rather significant influence on the execution-time of the whole quantity (time-to-completion), but it doesn't have a direct ef-

fect on the execution-time of the first share (time-to-first-fill). Thus, even though *ordersize* is taken into account when the order is submitted, it has a direct influences only on the price-aggressiveness of that order. We make the same assumption, of indirect influence on time, for round prices, spread and tick size. The variable *compete* is a proxy for the FKK (2004) model parameter  $\theta$  (the proportion of patient traders in the population) and, just as it is in the theoretical model, we assume that  $\theta$  has an influence on the expected execution-times of the different order strategies only through it's effect on the choice of strategy (price-aggressiveness). The same is true for *arrivalrate* which proxies the theoretical model parameter  $\lambda$  (arrival rate of traders) and for the variables that measure order side imbalance, *sameside* and *imbalance*.

## 5.1 Stocks

The results of the reduced form ordered probit analysis are not reported since they are merely an intermediary stage, and they do not necessarily represent the true effects of the explanatory variables on price-aggressiveness. At this stage, the effects are mixed and include both, the influence on the aggressiveness and on the lifetime of the order.

The results of the appropriate lifetime regression models, for each one of the two sub-samples of stocks, are presented in **Table 6S**. Note that the results are presented for  $Y = \log(\text{time})$  rather than for time itself. The measures of competition between suppliers of liquidity (*compete* and *opposed*), competition for order flow within one side of the market (*sameside* and *imbalance*) and the arrival rate of traders (*arrivalrate*) have significant effects on the lifetimes of both types of limit orders. As predicted, intense

competition prolongs the expected execution-time while high arrival rates tend to shorten that time. The dominant effect on the expected execution-time of price improving limit orders is of the instantaneous measure of order imbalance *sameside*, while the size of the queue as a proxy of competition for liquidity supply *compete* is as relevant as *sameside* only for traders who submit non-improving limit orders. It is only natural that the queue on the opposite side *opposed* has relatively more influence on the expected execution time of orders at the beginning of the queue than on those who choose not to improve on the quoted price. Yet, in both cases, the variable *opposed* has a negative effect on time.

The effects of the control variables for order submission time (*daytime* and *weekday*) are not always significant, but they have some interesting patterns. All types of limit orders face shorter execution time if submitted at the end of the trading day, and for non-improving limit orders, the expected execution time decreases monotonically during the day. For improving limit orders, we observe a different pattern of the expected execution time, which is negatively correlated with the activity levels observed during the day. The expected execution time of those orders is longest at midday, when the activity levels in the markets are usually low, and it is shorter at the end of the trading day, before the market closes. While the intraday pattern of expected execution-time is consistent with other measures of market activity, and could be interpreted as the result of intraday changes in the supply and demand of liquidity, the daily pattern is practically in the opposite direction. In fact, all types of limit orders face the longest execution time on the last day of the week, and the effects of the day of the week variables tend to be weaker than those of the intraday variables.

The effects of the volume and volatility variables on the expected execution-time is usually negative, whether they are defined as cross sectional (*volume*, *volatility*) or daily variables (*Zvolume*, *Zvolatility*). The dominant effect is that of the cross sectional variation in trading volume and, in general, execution time is affected by volume more than by volatility. The price level (*price*) is also used as a cross sectional control variable, and has a positive effect on the expected execution time of all limit order types. It will take longer time to execute a non-improving limit order if it is not included in the TA25 index, not traded on any other exchange and if it is a sell side order. It takes longer to execute an improving limit order if it is included in the TA25 index, not traded on any other exchange and it is a buy side order.

Finally, the coefficients of the weight variables ( $\widehat{W}_j^*$ ,  $j = 1, 2$ ), which correct for selection bias, are all large negative and significant, which suggests that such correction was required. The fact that they are negative means there is a positive correlation between the expected execution-time and the level of price-aggressiveness.

In **Table 7** we present the estimates of limit order lifetime. First, we compare the realized and the estimated lifetime within each sub-sample. The realized lifetime is the minimum of time-to-first-fill (execution of the first unit) and time-to-cancellation. The estimated lifetime within the sub-sample is corrected for selection bias, and thus it is the one relevant for this within sub-sample comparison. The distribution of the estimated lifetime stochastically dominates that of the realized one for all types of limit orders. This is expected since about 50% of the lifetime observations are censored (cancelled before any share is executed). The difference between the real-

ized and estimated lifetimes gets larger for the less aggressive limit order strategy, which is consistent with the larger percentage of censored observations we observe for that sub-sample, and may also indicate that censoring is correlated with the order submission strategy.

The econometric model we use enables us to estimate the expected lifetimes, conditional on each specific strategy the trader may choose, for every observation in our sample<sup>9</sup>. We compare two lifetime estimates: the estimated lifetime conditional on choosing an order-strategy of price-improving limit versus the one estimated conditional on the order-strategy of non-improving limit<sup>10</sup>. The difference between the lifetime estimates, of the improving and non-improving limit order strategy is positive for all the observations in our sample. Moreover, in **Table 7** we see that the distribution of the lifetime estimates for an aggressive order-strategy stochastically dominates the estimated distribution for the less aggressive one. All that is consistent with the price priority rule of the exchange.

The results of the third stage (structural probit) estimation are presented in **Table 8S**. The variables representing the trade-off between expected execution-time ( $\widehat{\beta}_j' X_i$ ,  $j = 1, 2$ ) and price (*spread*, *tick*) have large and significant effects on the probability of an aggressive order-submission strategy. Both conditions that make a market order strategy costly, a large spread and a large tick-size, reduce the probability of using that strategy, while long expected execution-times for the alternative (less aggressive) strategies increase that probability. This result supports the non-trivial direction of the

---

<sup>9</sup>Note that the estimated conditional lifetimes for the entire population do not include a correction for selection bias, simply because they are the relevant measures when one tries to compare the expected values for all observations in the sample.

<sup>10</sup>We assume that the lifetime of market orders is zero by definition.

correlation between execution time and order submission strategy, and provides empirical evidence for the assertion that the expected execution-times influence the trader's choice of order-strategy.

The variables that control for order submission time (*daytime* and *weekday*) do not affect the order submission strategy in the same way. While there are no significant differences on daily basis, the intraday variables are rather influential. The probability of an aggressive order (say market) increases at the second half of the day and is expected to be highest on the last hour of the continuous trading phase.

Other variables have smaller effects than those described above. Volume and the volatility have a positive effect on the probability of an aggressive order, yet the cross sectional variables (*volume*, *volatility*) are more influential than the daily ones, and volume is more influential than volatility. The probability of an aggressive order also increases if the price level of the stock (*price*) is high, the stock is not included in the TA25 index and if it is traded on some other exchange. The size of the submitted order (*ordersize*) has a negative effect on order aggressiveness, and so does the tendency of traders to prefer round numbers (*round*). Finally, the effect of the order side (*sell*) is very small and insignificant.

## 5.2 Bonds

The results of the appropriate lifetime regression models, for each one of the two sub-samples of bonds, are presented in **Table 6B**. Most variables affect the expected time-to-execution of bonds almost the same way they affect that of stocks. The measures of competition between suppliers of liquidity (*compete* and *opposed*), the instantaneous measure of competition for order

flow within one side of the market (*sameside*) and the proxy for arrival rate of traders (*arrivalrate*) have practically the same effect on the lifetimes of both types of limit orders in the bonds sample. That is if the relative magnitude and significance of the effects are considered. The only variable which behaves differently, and in an opposite direction to our prediction, is the second measure of competition for order flow, *imbalance*.

The coefficients of the weight variables, which correct for selection bias, are negative and significant. For bonds, as well as for stocks, we conclude that the correction for selection bias was required, and that there is a positive correlation between the expected execution-time and the level of price-aggressiveness.

The estimates of limit order lifetime are presented in **Table 7**. Just like for the sample of stocks, the distribution of the estimated lifetime stochastically dominates that of the realized one for all types of limit orders, and the difference between the lifetime estimates, of improving and non-improving limit order strategies, is positive for all the observations in the sample. There is a large difference between the realized and estimated distributions of lifetime that, to some extent, may be attributed to the larger percentage of censored observations in the sample of bonds. The size of the bonds sample, which is considerably smaller than that of the stocks, may be another cause for less accurate estimates. There is also a large difference in magnitude between the two estimates produced by our model, for the expected execution times of improving and non-improving limit orders.

The results of the third stage (structural probit) estimation are presented in **Table 8B**. All results for the bonds sample are practically the same as those for the sample of stocks. Yet, even though the relative importance

of the variables stays more or less the same, the effects of the explanatory variables on the probabilities (say the probability of a market order) are somewhat lower. An important result for this sample, as well as for the sample of stocks, is that the variables representing the trade-off between expected execution-time ( $\widehat{\beta}_j' X_i$ ,  $j = 1, 2$ ) and price (*spread*) have relatively large and significant effects on the probability of submitting an aggressive order (say market order). A large spread, which makes a market order strategy costly, reduces the probability of a market order while longer expected execution-times for the alternative (less aggressive) strategies increase that probability.

In summary, the results for the two samples provide answers for our three main questions of interest. First, the variables representing the trade-off between expected execution-time and price have relatively large and significant effects on the probability of submitting aggressive orders. Indeed, long execution times of non-aggressive orders increase the probability of using a price-aggressive strategy, while a large spread or a high tick-size decrease that probability. Second, the measures of competition between suppliers of liquidity and the arrival rate of traders have significant effects on the lifetime of all types of limit orders. As predicted by the theoretical models, intense competition prolongs the expected execution-time while a high arrival rate tends to shorten that time. Finally, our results are very much the same for both samples, stocks and government bonds. This finding supports our claim that liquidity based considerations have a significant effect on the price process in order driven markets, regardless of the degree of information asymmetry associated with the instrument.

## 6 Conclusion and Further Research

In this paper we ask whether the expected time-to-execution of the different order strategies, each characterized a different level of price-aggressiveness, affect the eventual choice of the order strategy. We present a framework for modeling order submission strategies and time-to-execution as a set of simultaneous equations, where both elements are endogenous and affect one another. Our empirical results confirm the existence of the expected trade-off between the expected time-to-execution and the level of price-aggressiveness, and offer a characterization of the relationship between those variables and other exogenous explanatory variables. Finally, the qualitatively similar results that we get for stocks and government bond strengthen our claim that liquidity based considerations have a significant effect on the price process regardless of information asymmetry.

The econometric model is relatively parsimonious, designed to capture the trade-off between the endogenous variables of interest yet, in the same time, allow for a feasible estimation procedure. In extensions we plan to add features to improve the model fit, such as non-naive censoring, since we suspect this assumption causes substantial bias in the quantitative fit of the lifetime models, especially for non-improving limit orders. Even though there is no way to identify strategic cancellations in our sample, we believe that price-change and time-of-the-day are important determinants of the decision whether to cancel an order. We think that introducing an additional endogenous variable for the decision to cancel and order, and introducing price and time as explanatory variables for that decision, will produce a richer yet tractable model.

The analyses presented in this paper are targeting one possible time variable, that was defined as time-to-first-fill, but one could argue that time-to-completion or some measure of execution probability are of more interest to traders. We plan to repeat the analyses for the modified definitions of the trade-off between time and price, and compare the results to our initial findings. This type of analysis will require a formulation where order size is an additional endogenous variable, which will complicate the estimation but will also capture the broader meaning of order aggressiveness. Finally, we also plan to compare our results for the equity and government bonds markets to those for options and derivatives, which have practically the same trading platform on the TASE.

## References

- [1] Biais, B., Hillion, P., and C. Spatt (1995), An Empirical Analysis of the Limit Order Book and the Order Flow in the Paris Bourse, *Journal of Finance*, 50, 1655-1689.
- [2] Bloomfield R., O'Hara M., and G. Saar (2002), The "Make or Take" Decision in Electronic Markets: Evidence on the Evolution of Liquidity, Working Paper.
- [3] Chakravarty, S. and C. Holden (1995), An Integrated Model of Market and Limit Orders, *Journal of Financial Intermediation*, 4, 213-241.
- [4] Copeland, T. and D. Galai, (1983), Information Effects of the Bid-Ask Spread, *Journal of Finance*, December 1983, 1457-1469.
- [5] Cox, D. R., and D. Oakes (1984), *Analysis of Survival Data*, Chapman and Hall.
- [6] Demsetz, H., (1968), The Costs of Transacting, *Quarterly Journal of Economics*, 82, 33-53.
- [7] Easley D. and M. O'Hara, (1992), Time and the Process of Security Price Adjustment, *Journal of Finance*, 47, 577-605.

- [8] Ellul, A., Holden C. W., Jain P. and R. Jennings (2003), A Comprehensive Test of Order Choice Theory: Recent Evidence form the NYSE, working paper, Indiana University.
- [9] Engle, R., (2000), The Econometrics of Ultra-High-Frequency Data, *Econometrica*, 68, 1-22.
- [10] Engle, R., and J. R. Russel (1998), Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data, *Econometrica*, 66, 1127-1162.
- [11] Foucault, T., (1999), Order Flow Composition and Trading Costs in a Dynamic Limit Order Market, *Journal of Financial Markets*, 2, 99-134.
- [12] Foucault, T., Kadan, O., and E. Kandel, (2003), Limit Order Book as a Market for Liquidity, *Review of Financial Studies*, forthcoming.
- [13] Glosten, L., (1994), Is the Electronic Order Book Inevitable, *Journal of Finance*, 49, 1127-1161.
- [14] Glosten, L. and P. Milgrom, (1985), Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders, *Journal of Financial Economics*, 14, 71-100.
- [15] Goettler, R., C. Parlour, and U. Rajan, (2004), Equilibrium in a Dynamic Limit Order Book, Carnegie Mellon working paper.
- [16] Griffith, M., Smith, B., Turnbull, D., and R. W. White (2000), The Costs and Determinants of Order Aggressiveness, *Journal of Financial Economics*, 56, 65-88.
- [17] Handa, P. and R. Schwartz (1996), Limit Order Trading, *Journal of Finance*, 51, 1835-1861.
- [18] Harris, L., (1990), Liquidity, Trading Rules and Electronic Trading Systems, NYU Solomon Center Series in Finance and Economics, 1990-4.
- [19] Harris, L., (1998), Optimal Dynamic Order Submission Strategies in Some Stylized Trading Problems, *Financial Markets, Institutions and Instruments*, Vol 7, 2.
- [20] Harris, L. and J. Hasbrouck (1996), Market versus Limit orders: the Superdot evidence on Order Submission Strategy, *Journal of Financial and Quantitative Analysis*, 31, 213-231.
- [21] Hasbrouck, J. and G. Saar (2002), Limit Orders and Volatility in a Hybrid Market: The Island ECN, working paper, NYU.

- [22] Hau, H., (2003), The Role of Transaction Costs for Financial Volatility: Evidence from the Paris Bourse, working paper, INSEAD.
- [23] Hollifield, B., Miller, A. and P. Sandas (2001), Empirical Analysis of Limit Order Markets, working paper, Carnegie-Mellon University.
- [24] Hollifield, B., Miller, A., P. Sandas and J. Slive (2003), Liquidity Supply and Demand in Limit Order Markets, working paper, Carnegie-Mellon University.
- [25] Kotz S., N. Balakrishnan and N. L. Johnson (2000) Continuous Multivariate distributions, John Wiley & Sons.
- [26] Kyle, A.S., (1985), Continuous Auctions and Insider Trading, *Econometrica* 53, 1315-1335.
- [27] Lo, A., C. MacKinlay, and J. Zhang (2002), Econometric Models of Limit Order Executions, *Journal of Financial Economics*, 65, 31-71.
- [28] Maddala, G. S., (1983), Limited-dependent and Qualitative variables in Econometrics, Cambridge University Press.
- [29] O'Hara, M., (2003), Liquidity and Price Discovery, AFA 2003 Presidential Address.
- [30] Parlour, C. (1998), Price Dynamics in Limit Order Markets, *Review of Financial Studies*, 11, 789-816.
- [31] Parlour, C., and D. Seppi (2001), Liquidity-Based Competition for Order Flow, Working Paper, Carnegie Mellon University.
- [32] Rock, K., (1996), The Specialists Order Book and Price Anomalies, Working Paper.
- [33] Rosu, I., (2004), A Dynamic Model of the Limit Order Book, working paper, MIT.
- [34] Seppi, D.(1997), Liquidity Provision with Limit Orders and a Strategic Specialist, *Review of Financial Studies*, 10, 103-150.
- [35] Stoll, H., (1978), The Supply of Dealer Services in Securities Markets, *Journal of Finance*, September 1978, 1133-1151.
- [36] U.S. Securities and Exchange Commission (SEC), 1997.Report on the practice of preferencing.

## Appendix 1: The Econometric Model

In this paper we developed and estimated an econometric model of switching lifetime regressions with endogenous switching, to model the trade-off between order execution time and price and, at the same time, investigate exogenous influences on our endogenous variables. The LMZ (2002) econometric estimation of limit orders lifetime made it clear that survival analysis is the appropriate modeling methodology for the positive and censored lifetime variables. In addition, the FKK (2004) theoretical model setting, where order price-aggressiveness and lifetime are endogenously determined in equilibrium, required an econometric model of simultaneous equations with endogenous switching. In particular, the endogenous switching is the tool we use to model the trade-off between time and price, where the choice of order price affects its execution time and the expected execution time influences the chosen price.

The formulation of our econometric model offers solutions to the modeling of censored lifetime variables, to the endogeneity of the price and time variables and also to the selection bias problem, caused by the fact that each lifetime variable is observed only for a sub-sample. In this section, we would like to present the econometric model and the estimation procedure in more detail, and discuss some of the modeling assumptions.

### The Setup of the Model

Consider a sequence of  $i = 1, \dots, n$  limit orders, and denote by  $C_i$  a latent variable representing the price-aggressiveness level of order  $i$ . We assume that the order submission strategy, and in particular its aggressiveness level, depends on the expected execution-time and on a set of exoge-

nous explanatory variables denoted by  $Z$ . We denote by  $Y_1$  the log expected time-to-execution of improving limit orders, and by  $Y_2$  the log expected time-to-execution of limit orders that do not improve on the quoted price. We assume that both lifetime variables depend on the same set of exogenous explanatory variables, denoted by  $X$ , which may overlap the set  $Z$ .

Let

$$Y_{1i} = \beta_1' X_i + u_{1i},$$

$$Y_{2i} = \beta_2' X_i + u_{2i},$$

and

$$C_i = \gamma' Z_i + \delta_1 Y_{1i} + \delta_2 Y_{2i} - u_i,$$

where

$$\begin{bmatrix} u_{1i} \\ u_{2i} \\ u_i \end{bmatrix} \sim N(0, \Sigma), \quad \Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{1u} \\ & \sigma_2^2 & \sigma_{2u} \\ & & \sigma_u^2 \end{bmatrix}.$$

The system of equation above is defined to capture the trade-off between execution time and price. We have one equation for each (non-zero) expected log execution-time ( $Y_1$  and  $Y_2$ ), each is observed only in the case the specific order strategy is chosen, and an equation for the price-aggressiveness level  $C$ , which is affected by the expected log execution-times of the different strategies. Given the priority rules in order driven markets, it is obvious that traders expect the execution-time of an aggressive limit order to be shorter than that of a non-aggressive one. This property should translate into the relationship  $\widehat{Y}_{1i} \leq \widehat{Y}_{2i}$  between the estimates of the expected lifetimes for every order  $i$ . Furthermore, given the FKK (2004) theoretical model setting, we expect traders' price-aggressiveness level to increase in the expected time-to-execution of a less aggressive order strategy. Thus, for example, the

probability of a market order should be higher when the trader expects relatively high execution-times for limit orders. This assertion should result in positive signs of the coefficients of the expected lifetime variables,  $\widehat{\delta}_1, \widehat{\delta}_2 > 0$ .

All the dependent variables defined above,  $Y_1$ ,  $Y_2$  and  $C$ , are not observable. Instead, we observe the following:

$$I_i = \begin{cases} 0 & \text{if } 0 < C_i < +\infty \text{ (market order - most aggressive)} \\ 1 & \text{if } -\mu_2 < C_i \leq 0 \text{ (price-improving limit order)} \\ 2 & \text{if } -\infty < C_i \leq -\mu_2 \text{ (non-improving limit order)} \end{cases},$$

or, if we define it in terms of the distribution of the residual  $u_i$ ,

$$I_i = \begin{cases} 0 & \text{if } -\infty < u_i < \gamma'Z_i + \delta_1Y_{1i} + \delta_2Y_{2i} \\ 1 & \text{if } \gamma'Z_i + \delta_1Y_{1i} + \delta_2Y_{2i} \leq u_i < \gamma'Z_i + \delta_1Y_{1i} + \delta_2Y_{2i} + \mu_2 \\ 2 & \text{if } \gamma'Z_i + \delta_1Y_{1i} + \delta_2Y_{2i} + \mu_2 \leq u_i < +\infty \end{cases}$$

(note that if we define the discrete variable that way the vector of parameters  $\gamma$  contains a constant term (say  $\mu_1$ ), which along with  $\mu_2$ , determines the cutoff point of the variable  $C$ ), and

$$Y_i = \begin{cases} 0 & \text{if } I_i = 0 \\ Y_{1i} & \text{if } I_i = 1 \\ Y_{2i} & \text{if } I_i = 2 \end{cases}.$$

Moreover, since a significant percentage of the orders are cancelled before execution, we do not observe  $Y_i$ , but rather a mixture of truncated and non-truncated lifetime values.

We can substitute the lifetime equations into the criterion function and present the reduced form of the criterion function as

$$C_i = \gamma'Z_i + (\delta_1\beta_1 + \delta_2\beta_2)'X_i - [u_i - \delta_1u_{1i} - \delta_2u_{2i}].$$

Taking into account the discrete nature of the observable variable, which

does not allow us to estimate the variance of the error term, the parameters  $\gamma$ ,  $\delta_1$ , and  $\delta_2$  are estimable only up to a proportionality factor. Let us denote  $(\sigma^*)^2 = \text{Var}[u_i - \delta_1 u_{1i} - \delta_2 u_{2i}]$ , and use it to write an estimable form of the criterion function as follows:

$$C_i^* = \gamma^* Z_i + (\delta_1^* \beta_1 + \delta_2^* \beta_2)' X_i - u_i^*,$$

where

$$C_i^* = \frac{C_i}{\sigma^*}, \gamma^* = \frac{\gamma}{\sigma^*}, \delta_j^* = \frac{\delta_j}{\sigma^*} \quad (j = 1, 2) \text{ and } u_i^* = \frac{[u_i - \delta_1 u_{1i} - \delta_2 u_{2i}]}{\sigma^*}.$$

Now we have an estimable set of equations for  $Y_1$ ,  $Y_2$  and  $C^*$  and, using the properties of the multivariate normal distribution, we can show that

$$\begin{bmatrix} u_{1i} \\ u_{2i} \\ u_i^* \end{bmatrix} \sim N(0, \Sigma^*), \quad \Sigma^* = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{1u^*} \\ & \sigma_2^2 & \sigma_{2u^*} \\ & & 1 \end{bmatrix}.$$

### Maximum Likelihood Estimation

Following the study of LMZ (2002), we assume that all cancellations, which occur before any fraction of the order is executed, represent naive censoring of lifetime values. Under this assumption, the likelihood function for the model is as follows:

$$\begin{aligned} L(\gamma^*, \mu_2^*, \delta_1^*, \delta_2^*, \beta_1, \beta_2, \sigma_1^2, \sigma_2^2, \sigma_{1u^*}, \sigma_{2u^*}) = & \\ & \prod_{\substack{I_i=0 \\ S_i=0}} \int_{-\infty}^{\gamma^* z_i + (\delta_1^* \beta_1 + \delta_2^* \beta_2)' x_i} \phi(u_i^*) du_i^* \\ & \prod_{\substack{I_i=1 \\ S_i=0}} \int_{\gamma^* z_i + (\delta_1^* \beta_1 + \delta_2^* \beta_2)' x_i}^{\gamma^* z_i + (\delta_1^* \beta_1 + \delta_2^* \beta_2)' x_i + \mu_2^*} f_1(y_i - \beta_1' x_i, u_i^*) du_i^* \\ & \prod_{\substack{I_i=2 \\ S_i=0}} \int_{\gamma^* z_i + (\delta_1^* \beta_1 + \delta_2^* \beta_2)' x_i + \mu_2^*}^{+\infty} f_2(y_i - \beta_2' x_i, u_i^*) du_i^* \end{aligned}$$

$$\prod_{\substack{I_i=1 \\ S_i=1}} \int_{(y_i - \beta_1' x_i)}^{+\infty} \int_{\gamma^{*'} z_i + (\delta_1^* \beta_1 + \delta_2^* \beta_2)' x_i}^{\gamma^{*'} z_i + (\delta_1^* \beta_1 + \delta_2^* \beta_2)' x_i + \mu_2^*} f_1(u_{1i}, u_i^*) du_i^* du_{1i}$$

$$\prod_{\substack{I_i=2 \\ S_i=1}} \int_{(y_i - \beta_2' x_i)}^{+\infty} \int_{\gamma^{*'} z_i + (\delta_1^* \beta_1 + \delta_2^* \beta_2)' x_i + \mu_2^*}^{+\infty} f_2(u_{2i}, u_i^*) du_i^* du_{2i}$$

where

$$S_i = \begin{cases} 1 & \text{if observation } i \text{ is censored} \\ 0 & \text{if observation } i \text{ is not censored} \end{cases} ,$$

and  $f_j(\cdot, \cdot)$  are the bivariate normal *pdfs* of  $(u_{ji}, u_i^*)$ ,  $j = 1, 2$ . We have two indicator variables,  $I$  for the order submission strategy and  $S$  for censoring, with three possible outcomes for the first variable and two for the second. Consequently, the likelihood function is constructed of seven expression, one for every possible combination of outcomes for the indicator variables (note that there is no possibility for censored lifetime when the order strategy is market, since the lifetime equals zero by definition).

It is possible to program an algorithm that maximizes this likelihood function, but the process is cumbersome, especially since it involves the evaluation of either an integral of the bivariate normal *pdf* or the bivariate normal *cdf* in each step. Thus, we will use a three stage estimation procedure to get consistent estimates for the unknown parameters in our equations, those of the criterion function ( $C$ ) that determines the order submission strategy, and those in each conditional lifetime equation ( $Y_1$  and  $Y_2$ ).

### Three Stage Estimation Procedure

Consider our model, as it is at this stage

$$Y_i = \begin{cases} 0 & \text{if } I_i = 0 \\ \beta_1' X_i + u_{1i} & \text{if } I_i = 1 \text{ and } S_i = 0 \\ \beta_2' X_i + u_{2i} & \text{if } I_i = 2 \text{ and } S_i = 0 \end{cases} ,$$

$$I_i = \begin{cases} 0 & \text{if } -\infty < u_i^* < \gamma^{*'}Z_i^* + \delta_1^*Y_{1i} + \delta_2^*Y_{2i} \\ 1 & \text{if } \gamma^{*'}Z_i^* + \delta_1^*Y_{1i} + \delta_2^*Y_{2i} \leq u_i^* < \gamma^{*'}Z_i^* + \delta_1^*Y_{1i} + \delta_2^*Y_{2i} + \mu_2^* \\ 2 & \text{if } \gamma^{*'}Z_i^* + \delta_1^*Y_{1i} + \delta_2^*Y_{2i} + \mu_2^* \leq u_i^* < +\infty \end{cases}$$

and

$$S_i = \begin{cases} 1 & \text{if observation } i \text{ is censored} \\ 0 & \text{if observation } i \text{ is not censored} \end{cases}$$

(note that if  $S_i = 1$  we can only say that  $Y_i > \beta_j'X_i + u_{1j}, j = 1, 2$ ),

where

$$\begin{bmatrix} u_{1i} \\ u_{2i} \\ u_i^* \end{bmatrix} \sim N(0, \Sigma^*), \quad \Sigma^* = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{1u^*} \\ & \sigma_2^2 & \sigma_{2u^*} \\ & & 1 \end{bmatrix}.$$

We do not observe the variables  $Y_{1i}$  and  $Y_{2i}$  for the entire sample, so we turn our attention to the observable variables and correct for selection bias. For example, we can observe  $Y_{1i}$  only for the sub-sample where  $I_i = 1$ . The main problem with an OLS estimation of the equation  $Y_{1i} = \beta_1'X_i + u_{1i}$  for that sub-sample (in the case of an uncensored sample), or with survival analysis using the maximum likelihood estimation technique (in our case, where some of the observations are censored), is that  $E(u_{1i}|I_i = 1) \neq 0$ . Fortunately, our assumption of multivariate normal distribution enables us to calculate this expectation

$$E(u_{1i}|I_i = 1) = E(u_{1i}|a < u_i^* < b) = \sigma_{1u^*} \frac{\phi(a) - \phi(b)}{\Phi(b) - \Phi(a)},$$

where

$$a = \gamma^{*'}Z_i^* + \delta_1^*Y_{1i} + \delta_2^*Y_{2i} \text{ and } b = \gamma^{*'}Z_i^* + \delta_1^*Y_{1i} + \delta_2^*Y_{2i} + \mu_2^*.$$

This result enables us to rewrite the lifetime equation for the observable sub-sample ( $I_i = 1$ ), defining a new residual term which has zero mean

$$(Y_i|I_i = 1) = \beta_1'X_i + \sigma_{1u^*} \frac{\phi(a) - \phi(b)}{\Phi(b) - \Phi(a)} + \varepsilon_{1i}, \quad E(\varepsilon_{1i}) = 0.$$

In the second stage estimation we are able to use survival analysis techniques on the modified equation. Correcting for the selection bias, we can rewrite all our original second stage equations as:

$$Y_i = \begin{cases} 0 & \text{if } I_i = 0 \\ \beta_1' X_i + \sigma_{1u^*} W_{1i}^* + \varepsilon_{1i} & \text{if } I_i = 1 \\ \beta_2' X_i + \sigma_{2u^*} W_{2i}^* + \varepsilon_{2i} & \text{if } I_i = 2 \end{cases}$$

where

$$W_{0i}^* = \frac{-\phi(\gamma^{*'} Z_i^* + \delta_1^* Y_{1i} + \delta_2^* Y_{2i})}{\Phi(\gamma^{*'} Z_i^* + \delta_1^* Y_{1i} + \delta_2^* Y_{2i})},$$

$$W_{1i}^* = \frac{\phi(\gamma^{*'} Z_i^* + \delta_1^* Y_{1i} + \delta_2^* Y_{2i}) - \phi(\gamma^{*'} Z_i^* + \delta_1^* Y_{1i} + \delta_2^* Y_{2i} + \mu_2^*)}{\Phi(\gamma^{*'} Z_i^* + \delta_1^* Y_{1i} + \delta_2^* Y_{2i} + \mu_2^*) - \Phi(\gamma^{*'} Z_i^* + \delta_1^* Y_{1i} + \delta_2^* Y_{2i})},$$

$$W_{2i}^* = \frac{\phi(\gamma^{*'} Z_i^* + \delta_1^* Y_{1i} + \delta_2^* Y_{2i} + \mu_2^*)}{1 - \Phi(\gamma^{*'} Z_i^* + \delta_1^* Y_{1i} + \delta_2^* Y_{2i} + \mu_2^*)},$$

and

$$E(\varepsilon_{1i} | I_i = 1) = E(\varepsilon_{2i} | I_i = 2) = 0.$$

We use the theoretical results obtained above in the following description of the estimation process.

*Stage I:*

First, we focus on the reduced form of the criterion function, which was normalized to have a unit variance:

$$C_i^* = \gamma^{*'} Z_i + (\delta_1^* \beta_1 + \delta_2^* \beta_2)' X_i - u_i^*,$$

where

$$C_i^* = \frac{C_i}{\sigma^*}, \quad \gamma^* = \frac{\gamma}{\sigma^*}, \quad \delta_1^* = \frac{\delta_1}{\sigma^*}, \quad \delta_2^* = \frac{\delta_2}{\sigma^*}, \quad \sigma^* = \text{Var}[u_i - \delta_1 u_{1i} - \delta_2 u_{2i}],$$

and

$$u_i^* = \frac{[u_i - \delta_1 u_{1i} - \delta_2 u_{2i}]}{\sigma^*}.$$

We use an ordered probit method to get consistent estimates of the parameters  $\gamma^*$ ,  $\mu_2^*$  and  $(\delta_1^*\beta_1 + \delta_2^*\beta_2)$ , which are obtained by maximizing the following likelihood function:

$$\begin{aligned}
L(\gamma^*, \mu_2^*, (\delta_1^*\beta_1 + \delta_2^*\beta_2)) = & \\
& \prod_{I_i=0} \Phi(\gamma^{*'}z_i + (\delta_1^*\beta_1 + \delta_2^*\beta_2)'x_i) \\
& \prod_{I_i=1} [\Phi(\gamma^{*'}z_i + (\delta_1^*\beta_1 + \delta_2^*\beta_2)'x_i + \mu_2^*) - \Phi(\gamma^{*'}z_i + (\delta_1^*\beta_1 + \delta_2^*\beta_2)'x_i)] \\
& \prod_{I_i=2} [1 - \Phi(\gamma^{*'}z_i + (\delta_1^*\beta_1 + \delta_2^*\beta_2)'x_i + \mu_2^*)].
\end{aligned}$$

*Stage II:*

As we already mentioned, each lifetime variable is observable only for a sub-sample, and we have to use the results of the estimation in the first stage to correct for selection bias of the second stage expected-lifetime estimates. We calculate the values of the appropriate weights ( $\widehat{W}_{1i}^*$  and  $\widehat{W}_{2i}^*$ ) for each observations, and use them as additional explanatory variables in the two lifetime equations:

$$Y_i = \begin{cases} \beta_1'X_i + \sigma_{1u^*}W_{1i}^* + \varepsilon_{1i} & \text{if } I_i = 1 \\ \beta_2'X_i + \sigma_{2u^*}W_{2i}^* + \varepsilon_{2i} & \text{if } I_i = 2 \end{cases},$$

where cancellations represent naive censoring of the lifetime variables. For the  $j$ -th equation ( $j = 1, 2$ ) we estimate the parameters  $\beta_j$  and  $\sigma_{ju^*}$  by maximizing the following likelihood function

$$L(\beta_j, \sigma_{ju^*}) = \prod_{S_i=0} f_j(\beta_j'x_i + \sigma_{ju^*}\widehat{w}_{ji}^*) \prod_{S_i=1} [1 - F_j(\beta_j'x_i + \sigma_{ju^*}\widehat{w}_{ji}^*)],$$

where  $f_j(\cdot)$  denotes the *pdf* of  $\varepsilon_{ji}$ , and  $F_j(\cdot)$  is the appropriate *cdf*. We assume that both distributions are approximately normal.

*Stage III:*

Finally, we can use the parameter estimates, obtained at the second stage, to evaluate  $\widehat{\beta}'_1 x_i$  and  $\widehat{\beta}'_2 x_i$  (the estimates of the expected-lifetime values,  $y_{1i}$  and  $y_{2i}$ ). The structural normalized criterion function is

$$C_i^* = \gamma^* Z_i + \delta_1^* \widehat{\beta}'_1 X_i + \delta_2^* \widehat{\beta}'_2 X_i - u_i^*,$$

and the consequent observable discrete variable is  $I_i$ . Just like we did in the first stage, we use an ordered probit method to get consistent estimates for the parameters  $\gamma^*$ ,  $\mu_2^*$ ,  $\delta_1^*$  and  $\delta_2^*$ , by maximizing the following likelihood function:

$$\begin{aligned} L(\gamma^*, \mu_2^*, \delta_1^*, \delta_2^*) = & \\ & \prod_{I_i=0} \Phi\left(\gamma^* z_i + \delta_1^* \widehat{\beta}'_1 x_i + \delta_2^* \widehat{\beta}'_2 x_i\right) \\ & \prod_{I_i=1} \left[ \Phi\left(\gamma^* z_i + \delta_1^* \widehat{\beta}'_1 x_i + \delta_2^* \widehat{\beta}'_2 x_i + \mu_2^*\right) - \Phi\left(\gamma^* Z_i + \delta_1^* \widehat{\beta}'_1 x_i + \delta_2^* \widehat{\beta}'_2 x_i\right) \right] \\ & \prod_{I_i=2} \left[ 1 - \Phi\left(\gamma^* z_i + \delta_1^* \widehat{\beta}'_1 x_i + \delta_2^* \widehat{\beta}'_2 x_i + \mu_2^*\right) \right]. \end{aligned}$$

Note that the sets  $Z$  (exogenous variables characterizing the level of order price-aggressiveness) and  $X$  (exogenous variables characterizing the expected lifetime of the order) must not overlap completely for identification reasons. Otherwise, we face a perfect multicollinearity problem and the parameters  $\gamma^*$ ,  $\delta_1^*$  and  $\delta_2^*$  are not estimable.

### The Categories of Order Submission Strategy

The discrete variable  $I$ , representing order submission strategy in our model, has three categories. Given that we use a formulation that is ordered and discrete, one could argue that it is not complicated to slice the latent variable  $C$  into many categories and make finer distinctions between strategies. Obviously, the slicing cannot be too fine, since this action also

determines the size of each sub-sample that we use for the estimation of the expected lifetime in that group. Yet another reason for a low number of categories is that finer definitions are also costly in terms of the total sample size. Note that all the order strategies are possible, in the three category framework that we analyze in the paper, only for orders submitted when the spread size is at least two ticks. Thus, 140,520 (26%) of the original number of observations in the stocks sample are excluded from the analyses. If we use a definition with four categories, where we distinguish between orders that improve by one tick and orders that improve by more than that, an aggressively improving order is distinguishable from other strategies only if the inside spread is three ticks or more. This definition will result in an additional loss of 73,927 (13%) observations from the sample.

In spite of the reduction in the total sample size, a definition of four categories seems like a good compromise that captures all aspects of order strategy. The decision whether to queue or not is captured by the cutoff between non-improving and price-improving orders, the decision regarding price aggressiveness is captured by small versus large improvements on the quoted price, and the decision whether to supply or demand liquidity is captured by the distinctions between limit and market orders. Eventually, we decided to use a model with only three categories for the strategy variable for two main reasons: small sub-sample size and a multicollinearity problem.

The distribution presented in **Table A1**, where an order is considered to be an aggressive improvement if it reduces the spread size by more than 50%, indicates that the number of aggressively improving limit orders is relatively very low (only 2% of the orders in the sample of stocks and 4% in the sample of bonds). Even if we change the definition of this category

to include all limit orders that improve by more than one tick, the size of that category remains very small. In addition, we encountered a multicollinearity problem when we tried to estimate the four categories model. The structural form of the criterion function equation  $C^*$  includes the estimates of the expected lifetime for each strategy as explanatory variables, and the correlations between those variables are high. Plugging three such lifetime variables into the structural criterion equation resulted in a sign reversal for some of them. This way, when all three lifetime variables were included in the equation, we got the unreasonable result that long execution time of aggressively improving limit orders reduces the probability of submitting a market order.

Finally, we would like to discuss our definition of the category of the most aggressive order submission strategy. An inspection of the frequencies of the different strategies in our sample (**Table A1**) indicates that there are two clear cutoff points: one is between non-improving and improving limit orders and the other is between improving limit orders and market orders. Yet, we decided to include very aggressive limit orders in the same category as marketable limit orders. Those orders also represent a very aggressive strategy, their percentage is very small, and we suspect that in many cases the traders actually meant to submit a marketable limit order yet the prices moved. The distribution of the size of spread-improvement of improving limit orders, for a range of spread-size categories, is presented in **Table A2**, and it supports our decision to define the cutoff points between the categories as we did. There is a relatively small percentage of improving limit orders that reduce the size of the spread by more than 25% but less than 75%, and as the spread gets larger the orders tend to concentrate closer to the quoted

prices. For the sample of stocks, there are 8,617 aggressively improving limit orders that were included in the category of marketable limit orders  $I = 0$ . That is merely 3% of the total number of limit orders, 10% of the number of price-improving limit orders and 6% of the total number of orders included in the category  $I = 0$ . For the sample of bonds the picture is practically the same.

### Cancellations as Naive Censoring

One of our important assumptions is that order cancellations are random, so we can treat cancellations as naive censoring. This assumption is prevalent in the literature, and both studies we cited above (LMZ (2002) and HMSS (2003)), that estimate the lifetime of orders, use it. Common sense, but also the estimation results for our econometric model, indicate that the naive censoring assumption may be problematic. We plan to develop a model that will relax that assumption by allowing strategic cancellations. Unfortunately, such a model will increase the dimensions of our estimation problem and we are not yet sure if it's estimation is feasible. If, for example, we assume that strategic cancellations depend on the time of the day and the price at the time of cancellation, we have to add an equation to our original system. Unless we assume that at least order strategy and cancellation time are correlated, this addition will not affect the results of our original model but, if we allow for non-zero correlation, we will get a three rather than two dimensional distribution in each expression in the likelihood function. A direct estimation of the model parameters, by maximizing the likelihood function, becomes even more complicated than before, and we are not sure how one can extend the multistage estimation process to suit this

new formulation. Thus, we plan to address this problem in future research.

## Appendix 2: Robustness

We analyze TASE data since it is a pure limit order market, which is also the assumption in the theoretical model that we use to get intuition. However, there is always a concern that our results are driven by some features of the data that we did not take into account. To avoid this concern we reproduce the analyses of Lo, MacKinlay and Zhang (2002) in our data and show that the results of our estimation using TASE data are very close to those obtained by LMZ for the S&P500 data.

LMZ (2002) utilize survival analysis techniques to estimate the expected lifetime of limit orders. They use an accelerated failure time specification, and in particular assume the generalized gamma distribution in their model of time-to-execution. They estimate time-to-first-fill and time-to-completion, separately for buy and sell side limit orders. Cancellations before execution and, as customary in lifetime regression models, are assumed to be the result of naive censoring. The data set in the LMZ (2002) study is unique, provided by an institutional brokerage firm (ITG), and it contains all the limit orders submitted through that firm during one year for the 100 largest stocks in the S&P500 index. LMZ provide an empirical estimation and characterization of the lifetime of limit orders, but their analysis doesn't rely on a theoretical model and thus, they end up ignoring the endogeneity of order lifetime and strategy. Going over the list of the explanatory variables in their model, one can easily observe aspects of order strategy aggressiveness incorporated in the model as if they were exogenous. In the next section we present our results for the estimation of the LMZ (2002)

model for time-to-first-fill, followed by a detailed discussion of the explanatory variables.

### The Model of LMZ (2002) Reproduced for the TASE Sample of Stocks

A comparison of the LMZ (2002) lifetime regression estimates of limit orders time-to-first-fill, for our dataset and the original results of LMZ for their dataset, is presented in **Table A3**<sup>11</sup>. Most of the results we get are the same as those of LMZ, but there are a couple of differences. First, there is a significant difference in the magnitude of the coefficient of *MQLP* and *MKD1X*, but the fact that spreads are not as tight for TASE stocks as they are for US most active stocks may be the reason of that. As for the interaction variable *MKD1X*, we had to truncate some of its extreme values so the estimation algorithm will converge. On the other hand, two estimates that assume the "wrong" sign in the LMZ analysis behave as expected in our dataset. If the last transaction was buyer initiated and the current order is a limit buy (sell), we expect it to have a positive (negative) effect on order lifetime. This prediction is confirmed by our dataset while the LMZ analysis shows the reverse effect for the variable *BSID*. High values of the variable *SZSD* characterize non-aggressive limit orders, thus the variable is predicted to have a positive effect on lifetime for both buy and sell limit orders. In our dataset it is exactly the case but it is not so for sell orders in the LMZ dataset.

---

<sup>11</sup>Three of the original variables are modified, but we get practically the same results if we use the original LMZ definitions. The variable *BSIDm* is a modified version of *BSID*, where we are able to extract the exact initiator of the last trade. The variables *LPRm* and *LVOm* are based on monetary volumes rather than the number of shares, and the variable *LSO* (log of shares outstanding) is omitted since we don't have that information for our sample.

The use of generalized Gamma distribution, to model limit order lifetime in the LMZ analyses, is justified by the fact that both distribution parameters (shape and scale) are significant. However, in our dataset the shape parameter is not significantly different from zero for sell side orders. In **Table A4** we compare the estimates under the assumption of a lognormal distribution of limit order lifetimes and those under the assumption of generalized gamma distribution. The results for the sell side models are practically the same, as we could expect given the insignificant value of the shape parameter. For buy side orders the shape parameter has some explanatory power, but this fact doesn't have a qualitative effect on the interpretation of the model, since the estimates in both models are practically the same. Comparing the two models we conclude that even though we make a simplifying distributional assumption, one that will enable the estimation of our structural model, qualitatively we may expect to get similar results to those of a richer model based on the generalized gamma distribution.

The results presented in **Table A3** also indicate that, except for a couple of variables with opposite signs, the estimated parameter values for buy and sell side orders are very much the same. We defined the variables *BSID* and *MKD1X* in a symmetrical way, so they will have the same effect on buy and sell side orders, and added an order side dummy variable to the model. The estimates of the symmetrical version of the LMZ model, for the whole sample (buy and sell limit orders), are presented in **Table A5**. The values of the estimates are slightly different from those in the separate models, but qualitatively they are the same. The order side dummy is not significant and, practically, has no effect of the values of other explanatory variables. These results enable us to conclude that the separation of orders by order

side is unnecessary. Given that such an estimation strategy is costly, since the sample size in each model is half the number of observations, we prefer to use both buy and sell orders in one model.

### Definitions of Explanatory Variables in the LMZ (2002) Model

In this section we list the variables used in the LMZ (2002) model, for the estimation of limit orders lifetime, and point out how those variables are connected to measures of price-aggressiveness.

#### *Notation*

$P$  = market price (most recent transaction)

$P_l$  = limit order price

$P_b$  = bid price (best bid)

$P_a$  = ask price (best ask)

$P_q$  = mid quote price (  $(P_b + P_a)/2$  )

$S_l$  = limit order size (quantity of the submitted limit order)

$S_b$  = bid size (probably quantity of best bid)

$S_a$  = ask size (probably quantity of best ask)

*The dependent variables, defined for a limit-buy order:*

$$MQLP = P_q - P_l$$

$MQLP$  is the distance between the mid-quote and the limit price. A large value for  $MQLP$  means that little or no improvement of the quoted price was made. Large values of  $MQLP$  indicate low aggressiveness.

$$BSID = \begin{cases} +1 & \text{if prior trade occurred above } P_q \\ 0 & \text{if prior trade occurred at } P_q \\ -1 & \text{if prior trade occurred below } P_q \end{cases}$$

*BSID* indicates who was the initiator of the previous trade. If *BSID* equals +1 then the trade price was above the mid-quote, closer to the ask price, which is more probable if it was a buyer initiated trade. Our data set enables us to detect the exact initiator of the last trade, so we will modify the definition of this variable in our analyses. Note that this way we get only two values, either buyer or seller initiated trade.

$$MKD1 = \begin{cases} (1 + P_b - P_l) \times \log(S_b) & \text{if } P_l \leq P_b \\ 0 & \text{otherwise} \end{cases}$$

*MKD1* is different from zero only for buy limit orders that do not improve on the best bid, and it measures an interaction between two terms. The first term is the quantity on the bid side, that has higher price priority than the limit-buy order. The second one is the distance between the limit-buy order price and the best bid. *MKD1* gets high values if either the quantity or the distance are large, i.e. for non-aggressive orders.

$$MKD1X = \begin{cases} (P - P_l) \times MKD1 & \text{if } P > P_l \\ 0 & \text{otherwise} \end{cases}$$

*MKD1X* is different from zero only for buy limit orders submitted below the last transaction price, and it measures the interaction between two terms. The first term  $(P - P_l)$  is the distance between the last transaction price and the limit-buy order price (positive). The second term is the variable *MKD1*. High values of both expressions can be interpreted as versions of low aggressiveness.

$$MKD2 = \begin{cases} \log(S_a)/(1 + P_a - P_l) & \text{if } P_a \geq P_l \\ \log(S_a) & \text{otherwise} \end{cases}$$

*MKD2* is a measure of the interaction between liquidity offered on the ask side and limit-buy order price aggressiveness. The first expression

$(\log(S_a))$  measures the liquidity offered on the ask side, while the second one  $(1/(1 + P_a - P_l))$  measures the price aggressiveness of a limit-buy order, this time as the inverse of their distance from the ask price.  $MKD2$  is large if the liquidity offered on the ask side is high, especially if the order makes a large price improvement. High values of this variable may capture an exogenous high level of impatience, since the trader makes an aggressive price improvement in spite of the "crowding out" effect expected on the ask side.

$$SZSD = \begin{cases} \log(S_l) \times (1 + P_a - P_l) & \text{if } P_a > P_l \\ \log(S_l - S_a) & \text{if } P_a = P_l \text{ and } S_l > S_a \\ 0 & \text{otherwise} \end{cases}$$

$SZSD$  is a measure of the interaction between liquidity demanded by the limit-buy order and its price aggressiveness. For marketable limit-buy orders, the liquidity demand is lowest when they are price aggressive (the conditions  $P_a \leq P_l$  and  $S_l \leq S_a$  ensure immediate execution of the total quantity, thus the value of  $SZSD$  is zero). It's relatively low, for marketable limit-buy orders, when the quantity offered at the ask price is large (the conditions  $P_a = P_l$  and  $S_l > S_a$  ensure immediate but only partial execution of the order, thus the value of  $SZSD$  depends on the difference of quantities). For limit-buy orders, liquidity demand is defined by the interaction between two terms. The first term  $(\log(S_l))$  is high when the trader demands a large quantity, while the second term  $(1 + P_a - P_l)$  is large when the order aggressiveness level is low. High values of this variable may capture an exogenous high level of patience, since the trader is demanding a large quantity and isn't willing to use an aggressive order strategy.

$$STKV = (\# \text{ of trades last half hour} / \# \text{ of trades last hour})$$

$STKV$  is a short-term measure, capturing shifts in trading activity. The

value of the variable is high if there is an increase in the activity level in the market.

$$TURN = \log(\# \text{ of trades last hour})$$

*TURN* is a trading activity measure, providing an absolute measure of volatility. When activity is high the value of the variable is large and volatility is probably also high.

$$LSO = \log(\text{previous month-end shares outstanding, in thousands})$$

*LSO* is a measure of the potential market size (since trade activity is usually measured relative to the number of shares outstanding). When the activity potential is large the value of the variable is large. Since our data set is of Israeli stocks, and for some of the companies we observe large proportions of shares outstanding that are held by a few large investors, the more relevant measure of activity potential is the float.

$$LPR = \log(\text{previous month's average daily closing price})$$

*LPR* is a measure of the historical price level of the asset. Since the closing price at the TASE is not a good representative of the stock's true value, it could be better to use the historical open price.

$$LVO = \log(\text{previous month's average daily share volume})$$

*LVO* is a measure of the historical rate of activity or liquidity in the stock.

The last four variables: *TURN*, *LSO*, *LPR* and *LVO* are control variables designed to capture differences across stocks.

Table 1S

Summary statistics for the sample of 32 most liquid stocks traded on the TASE, with no changes in the tick-size, for the sample period from May 1, 2000 to July 31, 2000. The value of 'TA25 Indicator' is 1 if the stock is included in the TA25 index and the value of 'Dual Listing' is 1 if the stock is traded on some other venue in the US or Europe. We report the median daily NIS volume, the tick-size in 1/100 NIS, the average size of the spread (ticks), the average price (NIS) of the stock at the opening stage and the standard deviation of the price across days.

#	Ticker	ID	TA25 Indicator	Dual Listing	Median Daily Volume	Tick Size	Average Spread	Average Price	Std of Price
1	ISRA.L	232017	0	0	7,890,501	0.1	1.1	0.16	0.01
2	AVNR.L	268011	0	0	3,432,130	0.1	1.9	0.33	0.02
3	MGDL	1081165	1	0	1,829,317	0.1	30.6	4.03	0.25
4	DEDR.L	475020	0	0	1,687,587	0.1	16.9	2.23	0.13
5	POLI	662577	1	0	23,892,661	1	2.0	11.95	0.24
6	LUMI	604611	1	0	18,765,993	1	1.7	8.72	0.32
7	BEZQ	230011	1	0	17,611,433	1	4.7	23.06	1.01
8	MAIN	1081819	1	0	4,208,252	1	3.7	8.86	0.30
9	MZRH	695437	1	0	3,765,695	1	4.9	12.22	0.40
10	MSHV	169011	0	0	2,732,504	1	19.6	29.12	3.63
11	ELBT	768051	0	1	2,492,884	1	22.4	39.26	2.11
12	SAE	777037	1	1	2,364,396	1	6.8	14.56	0.58
13	AGIS	311019	0	0	1,267,826	1	28.1	36.34	2.50
14	TEVA	629014	1	1	24,171,905	10	2.9	215.66	19.10
15	KOR	649012	1	1	12,540,958	10	11.1	413.68	24.64
16	DISI	639013	1	0	11,645,293	10	6.6	224.37	15.46
17	NICE	273011	1	1	10,950,828	10	6.8	291.05	29.01
18	FORT	256016	0	1	7,944,436	10	6.4	219.00	11.74
19	IDBD	798017	1	0	7,828,695	10	7.8	177.67	10.88
20	ELRN	749077	1	1	4,825,355	10	6.3	142.46	9.74
21	POIN	1080670	0	1	4,376,808	10	5.5	93.89	11.62
22	IDBH	736579	1	0	4,104,989	10	8.9	152.26	8.89
23	LPMA	428011	0	0	3,720,620	10	8.9	142.68	18.39
24	MATV	510016	1	1	2,395,000	10	6.3	82.48	4.04
25	CLIS	224014	1	0	1,900,726	10	6.6	64.31	4.21
26	ELEI5	746099	0	0	1,565,533	10	17.2	213.56	13.48
27	ESLT	1081124	0	1	1,420,616	10	3.6	58.57	2.14
28	BRAN	286013	0	0	1,256,225	10	10.6	118.51	7.40
29	DELT	627034	0	1	1,238,641	10	7.0	87.28	3.02
30	BLSQ	223016	1	0	1,016,786	10	5.6	56.76	2.87
31	CLEI	808014	1	0	3,169,510	100	5.4	919.18	41.84
32	ILCO1	576017	0	0	2,120,941	100	6.1	767.63	73.46

Table 1B

Summary statistics for the sample of 45 most liquid government bonds traded on the TASE, for the sample period from May 1, 2000 to July 31, 2000. The value of 'Bond Type' is 'Nom' if the payments are nominal, 'Ind' if payments are indexed and 'Dollar' if they are linked to the US Dollar. We report the median daily NIS volume, the average size of the spread (ticks), the average price (NIS) of the bond at the opening stage and the standard deviation of the price across days. All bonds in the sample have a tick size of 0.0001 NIS.

#	Ticker	ID	Bond Name	Bond Type	Median Daily Volume	Average Spread	Average Price	Std of Price
1	TB810	8000812	Makam	Nom	11,348,764	46.2	0.9890	0.0063
2	TB910	8000911	Makam	Nom	6,160,309	6.7	0.9813	0.0062
3	TB1010	8001018	Makam	Nom	7,225,325	3.4	0.9742	0.0064
4	TB1110	8001117	Makam	Nom	9,165,360	5.3	0.9679	0.0064
5	TB1210	8001216	Makam	Nom	4,422,852	8.8	0.9603	0.0063
6	TB111	8010118	Makam	Nom	9,329,499	5.9	0.9536	0.0064
7	TB211	8010217	Makam	Nom	3,565,230	9.0	0.9454	0.0066
8	TB311	8010316	Makam	Nom	3,427,885	6.9	0.9398	0.0060
9	TB411	8010415	Makam	Nom	6,829,468	6.4	0.9335	0.0059
10	TB511	8010514	Makam	Nom	3,777,397	8.6	0.9298	0.0067
11	GN2250	9225038	Gilon	Nom	776,864	17.8	1.0427	0.0124
12	GN2251	9225137	Gilon	Nom	545,998	130.7	1.0167	0.0215
13	GN2270	9227034	Gilon	Nom	4,432,087	10.2	1.0231	0.0045
14	NGN2301	9230137	Gilon	Nom	14,718,775	10.2	1.0252	0.0119
15	NGN2302	9230236	Gilon	Nom	11,095,042	13.7	1.0286	0.0116
16	SH2633	9263336	Shahar	Nom	3,180,348	58.6	1.0538	0.0085
17	SH2634	9263435	Shahar	Nom	3,039,223	63.0	1.0923	0.0090
18	SH2635	9263534	Shahar	Nom	4,102,391	60.1	1.0538	0.0079
19	SH2636	9263633	Shahar	Nom	2,007,076	16.4	1.0871	0.0085
20	SH2660	9266032	Shahar	Nom	3,210,500	15.7	1.1158	0.0103
21	SH2661	9266131	Shahar	Nom	16,824,409	11.9	1.1117	0.0459
22	KF1399	9139932	Kfir	Ind	150,988	644.7	1.2312	0.0209
23	KF1511	9151135	Kfir	Ind	933,401	562.0	2.2987	0.0163
24	GL3870	9387036	Galil	Ind	1,504,207	191.0	1.0400	0.0089
25	SG4256	9425638	Sagi	Ind	2,520,046	247.7	1.2540	0.0093
26	SG4257	9425737	Sagi	Ind	1,725,618	54.9	1.1484	0.0106
27	GL4703	9470337	Galil	Ind	227,666	254.5	1.0899	0.0120
28	GL5419	9541939	Galil	Ind	73,999	764.5	1.4590	0.0191
29	GL5420	9542036	Galil	Ind	325,234	635.0	1.3844	0.0402
30	GL5422	9542234	Galil	Ind	56,262	306.1	1.2963	0.0192
31	GL5424	9542432	Galil	Ind	65,851	235.6	1.2530	0.0302
32	GL5425	9542531	Galil	Ind	287,093	240.6	1.2612	0.0252
33	GL5426	9542630	Galil	Ind	548,219	205.8	1.2398	0.0112
34	GL5427	9542739	Galil	Ind	4,407,816	32.1	1.0292	0.0258
35	GL5451	9545138	Galil	Ind	1,444,975	161.5	0.9695	0.0072
36	GL5470	9547035	Galil	Ind	2,052,085	88.2	1.0251	0.0101
37	GL5471	9547134	Galil	Ind	390,702	65.2	0.9961	0.0106
38	GB6535	9653536	Gilboa	Dollar	412,169	119.7	1.3640	0.0106
39	GB6536	9653635	Gilboa	Dollar	1,108,129	58.7	1.3060	0.0132
40	GB6537	9653734	Gilboa	Dollar	1,082,653	53.7	1.2786	0.0105
41	GB6538	9653833	Gilboa	Dollar	2,011,695	49.6	1.2843	0.0233
42	GB6539	9653932	Gilboa	Dollar	3,768,712	23.8	1.1614	0.0097
43	GB6540	9654039	Gilboa	Dollar	11,287,801	67.9	1.0790	0.0099
44	GB6541	9654138	Gilboa	Dollar	7,968,599	69.2	1.0127	0.0179
45	GB6542	9654237	Gilboa	Dollar	4,557,639	80.2	1.0214	0.0163

Table 2

The joint distribution of order type by order side, for non-error observations in the sample of 32 most liquid stocks traded on the TASE, with no changes in the tick-size, and 45 most liquid government bonds, for the sample period from May 1, 2000 to July 31, 2000. Order type may be MKT, which includes market and marketable limit orders, or LMT which includes limit orders.

Order type	Stocks			Bonds		
	Buy	Sell	Total	Buy	Sell	Total
<b>LMT</b>	167,878	166,110	333,988	30,435	26,825	57,260
	50%	50%	100%	53%	47%	100%
	59%	63%	61%	66%	67%	66%
<b>MKT</b>	116,397	98,572	214,969	15,572	13,413	28,985
	54%	46%	100%	54%	46%	100%
	41%	37%	39%	34%	33%	34%
<b>Total</b>	284,275	264,682	548,957	46,007	40,238	86,245
	52%	48%	100%	53%	47%	100%
	100%	100%	100%	100%	100%	100%

Table 3

The frequency of cancellations, calculated for non-error limit orders in the sample of 32 most liquid stocks traded on the TASE, with no changes in the tick-size, and 45 most liquid government bonds, for the sample period from May 1, 2000 to July 31, 2000. Cancellation type is 'Active' if the limit order was actively cancelled by the trader before any transaction took place. The type is 'End of the day' if the limit order was automatically cancelled by the exchange, at the end of the trading day, before any transaction took place.

Cancellation type	Stocks		Bonds	
	Cancelled LMT	Total LMT	Cancelled LMT	Total LMT
<b>Active</b>	124,440 65%		16,959 42%	
<b>End of the day</b>	67,795 35%		23,664 58%	
<b>Total</b>	192,235 58%	333,988 100%	40,623 47%	86,245 100%

Table 4

Summary statistics for time-to-cancellation (hours : minutes : seconds) by cancellation type, of non-error limit orders that were cancelled before any transaction took place, in the sample of 32 most liquid stocks traded on the TASE, with no changes in the tick-size, and 45 most liquid government bonds, for the sample period from May 1, 2000 to July 31, 2000. Cancellation type is 'Active' if the limit order was actively cancelled by the trader before any transaction took place. The type is 'End of the day' if the limit order was automatically cancelled by the exchange, at the end of the trading day, before any transaction took place.

	Stocks			Bonds		
	Cancellation type		All	Cancellation type		All
	Active	End of the day	Cancellations	Active	End of the day	Cancellations
<b>mean</b>	0:45:05	4:41:32	2:08:28	1:00:36	3:47:20	2:37:44
<b>std</b>	1:15:50	2:23:31	2:34:07	1:14:25	1:53:09	2:08:34
<b>min</b>	0:00:03	0:14:21	0:00:03	0:00:03	0:14:35	0:00:03
<b>Q25</b>	0:02:26	2:32:01	0:05:50	0:07:06	2:06:10	0:32:04
<b>median</b>	0:11:16	5:23:21	0:41:58	0:29:03	4:15:34	2:09:35
<b>Q75</b>	0:47:59	7:02:09	4:02:37	1:26:07	5:32:34	4:45:52
<b>max</b>	7:07:23	7:14:45	7:14:45	6:00:55	6:04:28	6:04:28
<b># of obs</b>	124,440	67,795	192,235	16,959	23,664	40,623
<b>pct</b>	65%	35%	100%	42%	58%	100%

Table 5

Summary statistics for time-to-first-hit (minutes) by order strategy, of non-error limit orders, for the sample of 32 most liquid stocks traded on the TASE, with no changes in the tick-size, and 45 most liquid government bonds, for the sample period from May 1, 2000 to July 31, 2000. Time-to-first-hit is defined to be the minimum of two: the execution-time of the first share for the limit order and the cancellation-time of the order (if cancelled).

	Stocks			Bonds		
	Price-Improving LMT orders	Non-Improving LMT orders	All LMT Orders	Price-Improving LMT orders	Non-Improving LMT Orders	All LMT Orders
<b>mean</b>	33.26	105.79	86.89	95.81	147.64	123.58
<b>std</b>	74.69	142.21	131.99	113.42	128.60	124.50
<b>min</b>	0.00	0.00	0.00	0.00	0.00	0.00
<b>Q25</b>	1.43	4.83	3.16	8.40	26.12	15.39
<b>median</b>	5.64	28.08	17.51	39.86	107.11	69.49
<b>Q75</b>	22.75	164.64	107.81	156.54	273.87	231.59
<b>max</b>	434.63	434.75	434.75	364.39	364.47	364.47
<b># of obs</b>	87,030	246,958	333,988	26,585	30,675	57,260
<b>pct</b>	26%	74%	100%	46%	54%	100%

Table 6S

The results of the second stage lifetime regression estimation, for non-error limit orders that were submitted when the spread was two ticks or more, in the sample of 32 most liquid stocks traded on the TASE, with no changes in the tick-size, for the sample period from May 1, 2000 to July 31, 2000. We assume that the expected lifetime (minutes) is log-normally distributed and treat cancellations as naive (random) censoring. The explanatory variables W1 and W2 are the appropriate weights, calculated to correct for selection bias due to the use of the relevant sub-sample of limit orders in each model estimation.

## Panel a – Price Improving LMT Orders (I=1)

<u>Parameter</u>	<u>Estimate</u>	<u>Chisq.</u>	<u>P Value</u>
Intercept	17.425	4,307	<.0001
daytime 7	0.135	11	0.0010
daytime 6	0.226	29	<.0001
daytime 5	0.309	54	<.0001
daytime 4	0.292	50	<.0001
daytime 3	0.026	0	0.5214
daytime 2	0.031	1	0.4193
weekday 5	0.213	48	<.0001
weekday 3	0.034	2	0.1809
compete	0.230	554	<.0001
opposed	-0.133	341	<.0001
arrivalrate	-0.036	49	<.0001
sameside	1.118	1,687	<.0001
imbalance	0.099	16	<.0001
price	0.167	802	<.0001
volume	-1.046	3,823	<.0001
volatility	-0.130	106	<.0001
Zvolume	-0.358	1,688	<.0001
Zvolatility	0.029	9	0.0026
TA25	0.263	97	<.0001
duallist	-0.056	5	0.0194
sell	-0.242	149	<.0001
W1	-1.906	1,029	<.0001
Scale	2.293		

## Panel b - Non-Improving LMT Orders (I=2)

<u>Parameter</u>	<u>Estimate</u>	<u>Chisq.</u>	<u>P Value</u>
Intercept	13.635	2,490	<.0001
daytime 7	-0.884	493	<.0001
daytime 6	-0.542	195	<.0001
daytime 5	-0.358	85	<.0001
daytime 4	-0.342	83	<.0001
daytime 3	-0.315	77	<.0001
daytime 2	-0.193	32	<.0001
weekday 5	0.136	21	<.0001
weekday 3	-0.039	3	0.1084
compete	0.562	4,856	<.0001
opposed	-0.125	309	<.0001
arrivalrate	-0.025	21	<.0001
sameside	0.627	592	<.0001
imbalance	0.199	67	<.0001
price	0.045	58	<.0001
volume	-0.694	1,709	<.0001
volatility	-0.026	5	0.0332
Zvolume	-0.331	1,664	<.0001
Zvolatility	-0.016	3	0.0847
TA25	-0.119	21	<.0001
duallist	-0.042	3	0.0678
sell	0.125	44	<.0001
W2	-1.872	553	<.0001
Scale	2.833		

# of Observations	67,311		
# of non-censored	40,202		
# of right censored	27,109	40%	
Log Likelihood	-107,091		
# of parameters	24		

# of Observations	142,846		
# of non-censored	51,778		
# of right censored	91,068	64%	
Log Likelihood	-169,415		
# of parameters	24		

Table 6B

The results of the second stage lifetime regression estimation, for non-error limit orders that were submitted when the spread was two ticks or more, in the sample of 45 most liquid government bonds traded on the TASE, for the sample period from May 1, 2000 to July 31, 2000. We assume that the expected lifetime (minutes) is log-normally distributed and treat cancellations as naive (random) censoring. The explanatory variables W1 and W2 are the appropriate weights, calculated to correct for selection bias due to the use of the relevant sub-sample of limit orders in each model estimation.

## Panel a - Price Improving LMT Orders (I=1)

<u>Parameter</u>	<u>Estimate</u>	<u>Chisq.</u>	<u>P Value</u>
Intercept	11.681	245	<.0001
daytime 7	-0.249	5	0.0241
daytime 6	-0.232	5	0.0214
daytime 5	-0.097	1	0.3296
daytime 4	-0.211	5	0.0250
daytime 3	-0.272	9	0.0022
weekday 5	-0.002	0	0.9793
weekday 3	0.122	3	0.0668
compete	0.521	315	<.0001
opposed	-0.056	13	0.0003
arrivalrate	-0.044	62	<.0001
sameside	1.238	313	<.0001
imbalance	-0.405	66	<.0001
price	-0.341	1	0.2230
volume	-0.758	336	<.0001
volatility	-0.027	2	0.1284
Zvolume	-0.390	293	<.0001
Zvolatility	0.018	1	0.4297
bondtype Nom	-0.804	107	<.0001
bondtype Ind	0.243	6	0.0142
sell	-0.207	16	<.0001
W1	-4.178	321	<.0001
Scale	2.845		
# of Observations		17,583	
# of non-censored		6,881	
# of right censored		10,702	61%
Log Likelihood		-22,298	
# of parameters		23	

## Panel b - Non-Improving LMT Orders (I=2)

<u>Parameter</u>	<u>Estimate</u>	<u>Chisq.</u>	<u>P Value</u>
Intercept	18.443	448	<.0001
daytime 7	-0.613	18	<.0001
daytime 6	-0.783	39	<.0001
daytime 5	-0.818	45	<.0001
daytime 4	-0.547	23	<.0001
daytime 3	-0.442	17	<.0001
weekday 5	0.144	2	0.1561
weekday 3	0.027	0	0.7493
compete	0.488	241	<.0001
opposed	-0.010	0	0.6225
arrivalrate	-0.042	31	<.0001
sameside	0.656	57	<.0001
imbalance	-0.192	9	0.0027
price	0.828	4	0.0421
volume	-0.824	231	<.0001
volatility	-0.094	15	<.0001
Zvolume	-0.572	462	<.0001
Zvolatility	-0.001	0	0.9753
bondtype Nom	0.306	9	0.0025
bondtype Ind	0.490	12	0.0006
sell	-0.055	1	0.4015
W2	-2.981	62	<.0001
Scale	3.318		
# of Observations		20,189	
# of non-censored		4,352	
# of right censored		15,837	78%
Log Likelihood		-16,893	
# of parameters		23	

Table 7

A comparison between the realized time-to-first-hit (minutes) and the estimated lifetime (minutes) of limit order, by limit order strategy, is presented in the first two sections of the table. Only the relevant sub-sample observations are used for these calculations: non-error improving limit orders for the first section and non-error non-improving limit order for the second. A comparison between the estimated lifetimes (minutes), conditional on improving and non-improving limit order strategy, across all non-error observations. All calculations use the relevant non-error orders, that were submitted when the spread was two ticks or more, in the sample of 32 most liquid stocks traded on the TASE, with no changes in the tick-size, and in the sample of 45 most liquid government bonds, for the sample period from May 1, 2000 to July 31, 2000.

### Stocks

#### Sub-sample of Price Improving Limit Orders (I=1)

	# of obs	q25	median	mean	q75
Realized	67,311	1.61	6.27	31.38	24.10
Estimated	67,311	7.60	15.54	25.39	31.57

#### Sub-sample of Non-improving Limit Orders (I=2)

	# of obs	q25	median	mean	q75
Realized	142,846	4.53	23.61	79.32	110.11
Estimated	142,846	87.63	174.13	265.18	334.95

#### Estimated Lifetime for All Observation in the Sample (I=0,1,2)

Conditioning Strategy	# of obs	q25	median	mean	q75
Improving LMT (I=1)	343,265	5.42	12.03	22.63	26.58
Non-Improving LMT (I=2)	343,265	555.05	1131.28	1768.84	2193.65

### Bonds

#### Sub-sample of Price Improving Limit Orders (I=1)

	# of obs	q25	median	mean	q75
Realized	17,583	8.79	36.66	79.80	121.82
Estimated	17,583	82.51	172.99	333.37	388.88

#### Sub-sample of Non-improving Limit Orders (I=2)

	# of obs	q25	median	mean	q75
Realized	20,189	22.68	81.70	117.14	206.86
Estimated	20,189	759.45	1697.45	2942.08	3523.86

#### Estimated Lifetime for All Observation in the Sample (I=0,1,2)

Conditioning Strategy	# of obs	q25	median	mean	q75
Improving LMT (I=1)	59,472	54.19	143.77	400.46	389.61
Non-Improving LMT (I=2)	59,472	20861.11	49267.82	96007.46	110363.18

Table 8S

The results of the third stage structural-form probit estimation, for non-error orders that were submitted when the spread was two ticks or more, in the sample of 32 most liquid stocks traded on the TASE, with no changes in the tick-size, for the sample period from May 1, 2000 to July 31, 2000. In this table we present the marginal effects, on the probability of each order submission strategy and the significance levels of the marginal effects of the explanatory variables on the MKT order strategy. The marginal effects are calculated for each variable with continuous variables at their mean value and discrete variables at the baseline level. The explanatory variables Xb1 and Xb2 are the estimates of log lifetime, conditional on price-improving and non-improving limit order strategies. Those variables are calculated for each order using the parameter estimates obtained in the second stage.

Parameter	Order Strategy			t*	P_Value*
	MKT	Improving LMT	Non-Improving LMT		
daytime 7	0.123	-0.021	-0.103	31.731	<.0001
daytime 6	0.048	-0.006	-0.042	11.730	<.0001
daytime 5	0.017	-0.002	-0.015	4.026	<.0001
daytime 4	0.017	-0.002	-0.015	4.165	<.0001
daytime 3	0.022	-0.002	-0.019	5.581	<.0001
daytime 2	0.007	-0.001	-0.006	1.854	0.0637
weekday 5	0.005	-0.001	-0.005	1.894	0.0583
weekday 3	0.002	0.000	-0.002	1.118	0.2637
spread	-0.071	0.007	0.064	-59.786	<.0001
tick 100	-0.212	-0.009	0.221	-21.398	<.0001
tick 10	-0.132	0.002	0.130	-26.431	<.0001
tick 1	-0.066	0.004	0.062	-15.425	<.0001
Xb1	0.079	-0.008	-0.071	39.914	<.0001
Xb2	0.037	-0.004	-0.034	23.119	<.0001
price	0.092	-0.009	-0.083	52.282	<.0001
volume	0.066	-0.007	-0.059	32.791	<.0001
volatility	0.011	-0.001	-0.010	10.223	<.0001
Zvolume	0.035	-0.004	-0.032	42.929	<.0001
Zvolatility	0.002	0.000	-0.002	2.429	0.0152
ordersize	-0.035	0.003	0.031	-42.512	<.0001
round	-0.045	0.003	0.042	-27.056	<.0001
TA25	-0.034	0.003	0.031	-15.124	<.0001
duallist	0.016	-0.002	-0.014	7.771	<.0001
sell	-0.001	0.000	0.001	-0.809	0.4183
# of Observations		343,265			
Log Likelihood		-351,319			
# of parameters		26			
# of (MKT) obs		133,107	39%		
# of (Improving LMT) obs		67,312	20%		
# of (Non-Improving LMT) obs		142,846	42%		

\* Note that the significance levels are calculated only for the effects on the probability of a MKT order.

Table 8B

The results of the third stage structural-form probit estimation, for non-error orders that were submitted when the spread was two ticks or more, in the sample of 45 most liquid government bonds traded on the TASE, for the sample period from May 1, 2000 to July 31, 2000. In this table we present the marginal effects, on the probability of each order submission strategy and the significance levels of the marginal effects of the explanatory variables on the MKT order strategy. The marginal effects are calculated for each variable with continuous variables at their mean value and discrete variables at the baseline level. The explanatory variables Xb1 and Xb2 are the estimates of log lifetime, conditional on price-improving and non-improving limit order strategies. Those variables are calculated for each order using the parameter estimates obtained in the second stage.

Parameter		Order Strategy			t*	P_Value*
		MKT	Improving LMT	Non-Improving LMT		
daytime	7	0.078	-0.009	-0.069	8.777	0.0000
daytime	6	0.050	-0.005	-0.046	5.649	0.0000
daytime	5	0.048	-0.004	-0.044	4.901	0.0000
daytime	4	0.029	-0.002	-0.027	3.763	0.0002
daytime	3	0.026	-0.002	-0.024	3.444	0.0006
weekday	5	-0.002	0.000	0.002	-0.284	0.7766
weekday	3	-0.006	0.000	0.006	-1.198	0.2311
spread		-0.044	0.002	0.042	-20.628	0.0000
Xb1		0.039	-0.002	-0.038	6.295	0.0000
Xb2		0.039	-0.002	-0.038	5.058	0.0000
price		0.040	-0.002	-0.038	1.848	0.0646
volume		0.054	-0.002	-0.051	15.356	0.0000
volatility		0.008	0.000	-0.008	6.124	0.0000
Zvolume		0.050	-0.002	-0.048	19.428	0.0000
Zvolatility		0.000	0.000	0.000	0.269	0.7883
ordersize		-0.021	0.001	0.020	-15.114	0.0000
round		-0.024	0.001	0.023	-6.620	0.0000
bondtype	Nom	-0.074	-0.002	0.076	-7.419	0.0000
bondtype	Ind	-0.016	0.000	0.015	-1.739	0.0821
sell		0.007	0.000	-0.007	2.001	0.0454
# of Observations			59,472			
Log Likelihood			-63,531			
# of parameters			22			
# of (MKT) obs			21,699	36%		
# of (Improving LMT) obs			17,584	30%		
# of (Non-Improving LMT) obs			20,189	34%		

\* Note that the significance levels are calculated only for the effects on the probability of a MKT order.

Table A1

The distribution of order strategies, for non-error limit orders that were submitted when the spread was two ticks or more, in the sample of 32 most liquid stocks traded on the TASE, with no changes in the tick-size, 45 most liquid government bonds, for the sample period from May 1, 2000 to July 31, 2000. The values of the column 'Category' are the values of the discrete variable I, that we use in the ordered probit model to describe the aggressiveness level of an order.

Order Submission Strategy	Category	Stocks		Bonds	
		Frequency	Percentage	Frequency	Percentage
Non-improving Limit	2	129,401	32%	20,334	26%
At-the-quote Limit	2	51,669	13%	7,718	10%
Up to 50% improving Limit	1	78,413	19%	23,272	30%
More than 50% improving Limit	0	8,617	2%	3,313	4%
Market and Marketable Limit	0	140,337	34%	23,563	30%
Total		408,437	100%	78,200	100%

Table A2

The distribution of the percentage price improvement by spread size (in ticks), for non-error price-improving limit orders that were submitted when the spread was two ticks or more, in the sample of 32 most liquid stocks traded on the TASE, with no changes in the tick-size, 45 most liquid government bonds, for the sample period from May 1, 2000 to July 31, 2000.

Stocks Spread size in ticks	Percentage of spread improvement			Frequency
	1% - 25%	26% - 75%	76% - 99%	
2	0%	100%	0%	7,914
3 - 5	47%	51%	2%	20,643
6 - 10	68%	28%	4%	22,461
11 - 20	74%	23%	3%	19,636
21 - 50	79%	18%	3%	12,969
51+	85%	12%	3%	3,407
Frequency	52,591	32,098	2,341	87,030

Bonds Spread size in ticks	Percentage of spread improvement			Frequency
	1% - 25%	26% - 75%	76% - 99%	
2	0%	100%	0%	841
3 - 5	51%	48%	1%	3,391
6 - 10	64%	34%	2%	4,878
11 - 20	64%	33%	3%	5,430
21 - 50	71%	25%	5%	6,005
51+	70%	23%	7%	6,040
Frequency	16,777	8,790	1,018	26,585

Table A3

An estimation of the LMZ (2002) lifetime regression model of time-to-first-fill with our sample of the 32 most liquid stocks traded on the TASE, with no changes in the tick-size, for the sample period from May 1, 2000 to July 31, 2000, and a comparison to the LMZ results as published in their paper. To be consistent with the LMZ study, only limit orders are included in the analysis, which is carried out separately for buy and sell side orders, we assume that the expected lifetime (minutes) has a generalized gamma distribution and we treat cancellations as naive (random) censoring. We did not have data regarding the number of shares outstanding, thus the variable LSO is excluded from our analyses.

<u>Parameter</u>	<u>DF</u>	<u>Buy limit orders</u>		<u>Sell limit orders</u>	
		<u>LMZ data</u>	<u>Our data</u>	<u>LMZ data</u>	<u>Our data</u>
		<u>Estimate</u>	<u>Estimate</u>	<u>Estimate</u>	<u>Estimate</u>
Intercept	1	6.507 ***	6.297 ***	4.979 ***	6.895 ***
MQLP	1	8.989 ***	0.093 ***	-13.674 ***	-0.077 ***
BSID(m)	1	-5.613 ***	0.328 ***	6.852 ***	-0.348 ***
MKD1	1	0.641 ***	0.247 ***	0.476 ***	0.278 ***
MKD1X	1	-0.920 ***	-0.015 ***	0.903 ***	0.019 ***
MKD2	1	-0.353 ***	-0.328 ***	-0.171 ***	-0.387 ***
SZSD	1	-0.015 ***	0.019 ***	0.091 ***	0.021 ***
STKV	1	-0.414 ***	-0.168 ***	-0.563 ***	-0.155 ***
TURN	1	-0.252 ***	-0.527 ***	-0.331 ***	-0.463 ***
LSO	1	0.278 ***		0.187 ***	
LPR(m)	1	-0.529 ***	-0.365 ***	-0.272 ***	-0.515 ***
LVO(m)	1	-0.082 ***	0.043 ***	-0.000 x	0.065 ***
Scale	1	1.927 ***	2.459 ***	1.804 ***	2.531 ***
Shape	1	-0.404 ***	-0.084 ***	-0.526 ***	0.000 x

Table A4

An estimation of the LMZ (2002) lifetime regression model of time-to-first-fill, using our sample of the 32 most liquid stocks traded on the TASE, with no changes in the tick-size, for the sample period from May 1, 2000 to July 31, 2000. To be consistent with the LMZ study, only limit orders are included in the analysis, which is carried out separately for buy and sell side orders, and cancellations are treated as naive (random) censoring. We compare the model estimates for the LMZ original assumption with the assumption of generalized gamma distribution with an assumption of lognormal lifetime distribution. We did not have data regarding the number of shares outstanding, thus the variable LSO is excluded from our analyses.

<u>Buy limit orders</u>		<u>Gamma</u>	<u>Lognormal</u>	<u>Sell limit orders</u>		<u>Gamma</u>	<u>Lognormal</u>
<u>Parameter</u>	<u>DF</u>	<u>Estimate</u>	<u>Estimate</u>	<u>Parameter</u>	<u>DF</u>	<u>Estimate</u>	<u>Estimate</u>
Intercept	1	6.297 ***	6.390 ***	Intercept	1	6.895 ***	6.9061 ***
MQLP	1	0.093 ***	0.091 ***	MQLP	1	-0.077 ***	-0.1443 ***
BSIDm	1	0.328 ***	0.326 ***	BSIDm	1	-0.348 ***	-0.3492 ***
MKD1	1	0.247 ***	0.247 ***	MKD1	1	0.278 ***	0.277 ***
MKD1X	1	-0.015 ***	-0.015 ***	MKD1X	1	0.019 ***	0.020 ***
MKD2	1	-0.328 ***	-0.329 ***	MKD2	1	-0.387 ***	-0.386 ***
SZSD	1	0.019 ***	0.020 ***	SZSD	1	0.021 ***	0.015 ***
STKV	1	-0.168 ***	-0.167 ***	STKV	1	-0.155 ***	-0.154 ***
TURN	1	-0.527 ***	-0.523 ***	TURN	1	-0.463 ***	-0.461 ***
LSO				LSO			
LPRm	1	-0.365 ***	-0.367 ***	LPRm	1	-0.515 ***	-0.520 ***
LVOm	1	0.043 ***	0.041 ***	LVOm	1	0.065 ***	0.068 ***
Scale	1	2.459 ***	2.421 ***	Scale	1	2.531 ***	2.5219 ***
Shape	1	-0.084 ***		Shape	1	0.000 x	
Log likelihood		-158,727	-158,749	Log likelihood		-148,503	-148,499
# of observations		117,397	117,397	# of observations		112,036	112,036

Table A5

An estimation of the LMZ (2002) lifetime regression model of time-to-first-fill, using our sample of the 32 most liquid stocks traded on the TASE, with no changes in the tick-size, for the sample period from May 1, 2000 to July 31, 2000. To be consistent with the LMZ study, only (non-error) limit orders are included in the analysis, which is carried out for all limit orders and also separately for buy and sell side limit orders, we assume that the expected lifetime (minutes) has a generalized gamma distribution and we treat cancellations as naive (random) censoring. We compare the estimates obtained from a pooled model for all limit orders that includes an explanatory indicator variable for order side, the same pooled model without an such explanatory variable and two separate models, one for each side. We did not have data regarding the number of shares outstanding, thus the variable LSO is excluded from our analyses.

Parameter	DF	<u>All limit orders - I</u>	<u>All limit orders - II</u>	<u>Buy limit orders</u>	<u>Sell limit orders</u>
		Estimate	Estimate	Estimate	Estimate
Intercept	1	6.441 ***	6.446 ***	6.297 ***	6.557 ***
MQLPs	1	0.121 ***	0.121 ***	0.093 ***	0.144 ***
BSIDms	1	0.335 ***	0.336 ***	0.328 ***	0.349 ***
MKD1	1	0.262 ***	0.262 ***	0.247 ***	0.277 ***
MKD1Xs	1	-0.018 ***	-0.018 ***	-0.015 ***	-0.020 ***
MKD2	1	-0.357 ***	-0.357 ***	-0.328 ***	-0.386 ***
SZSD	1	0.016 ***	0.016 ***	0.019 ***	0.015 ***
STKV	1	-0.166 ***	-0.166 ***	-0.168 ***	-0.154 ***
TURN	1	-0.497 ***	-0.497 ***	-0.527 ***	-0.461 ***
LSO					
LPRm	1	-0.441 ***	-0.441 ***	-0.365 ***	-0.520 ***
LVOm	1	0.056 ***	0.056 ***	0.043 ***	0.068 ***
Scale	1	2.496 ***	2.496 ***	2.459 ***	2.522 ***
Shape	1	-0.049 ***	-0.049 ***	-0.084 ***	0.000 x
oside=B	1		-0.013 x		
Log likelihood		-307,507	-307,506	-158,727	-148,499
# of observations		229,433	229,433	117,397	112,036

\*\*\* P\_value < 0.01  
x Not Significant at 0.1