ELSEVIER

Journal of Financial Economics 65 (2002) 31–71

JOURNAL OF Financial ECONOMICS

www.elsevier.com/locate/econbase

Econometric models of limit-order executions ☆

Andrew W. Lo[a,*], A. Craig MacKinlay[b], June Zhang[c]

[a] *MIT Sloan School of Management, 50 Memorial Drive, Cambridge, MA 02142-1347, USA*

[b] *Finance Department, 2300 Steinberg-Dietrich Hall, Wharton School, University of Pennsylvania, Philadelphia, PA 19104-6367, USA*

[c] *AlphaSimplex Group, One Cambridge Center, Cambridge, MA 02142, USA*

Received 25 September 1997; received in revised form 24 November 2000

Abstract

We develop and estimate an econometric model of limit-order execution times using survival analysis and actual limit-order data. We estimate versions for time-to-first-fill and time-to-completion for both buy and sell limit orders, and incorporate the effects of explanatory variables such as the limit price, limit size, bid/offer spread, and market volatility. Execution times are very sensitive to the limit price, but are not sensitive to limit size. Hypothetical limit-order executions, constructed either theoretically from first-passage times or empirically from transactions data, are very poor proxies for actual limit-order executions.

JEL classification: G23

Keywords: Market microstructure; Transactions costs; Portfolio management

☆ This research was partially supported by the MIT Laboratory for Financial Engineering, the Rodney White Center for Financial Research, and the National Science Foundation (Grant No. SBR-9709976). We are grateful to ITG for providing us with their limit-order dataset, and to the New York Stock Exchange for their TORQ dataset. We thank Jushan Bai, David Cushing, Robert Ferstenberg, Harvey Fram, Joachim Grammig, Larry Harris, Joel Hasbrouck, Stephanie Hogue Mauricio Karchmer, David Modest, Bill Schwert, George Sofianos, the referee, and participants at the Finance Seminar at Johann Wolfgang Goethe University (Frankfurt), the Berkeley Program in Finance, the Western Finance Association's 1998 Annual Meeting, and the 1998 Conference on the International Competition for Order Flow for helpful comments and discussion.

*Corresponding author.

E-mail address: alo@mit.edu (A.W. Lo).

0304-405X/02/$ - see front matter
PII: S0304-405X(02)00134-4

1. Introduction

One of the most important tools for trading equity securities is the limit order, which is an order to transact a prespecified number of shares at a prespecified price. Indeed, limit orders constitute a significant fraction of stock market trading activity, accounting for approximately 45% of total NYSE orders (Harris and Hasbrouck, 1996). The primary advantage of a limit order is the absence of price risk—a transaction occurs only if the *limit price* is attained. However, this advantage does not come without a cost: execution is not guaranteed, and the time-to-execution is a random function of many factors, such as the limit price, the number of shares, market conditions, and private information. For some trades, the uncertainty in execution time is unimportant, but for others, the opportunity cost of waiting can be significant.

If immediacy is critical, the market order is the appropriate instrument to use. However, market orders can be subject to significant price risk, particularly for large orders and in volatile markets. In practice, traders submit both market and limit orders, with an eye towards balancing the risks of delaying execution against the risks associated with immediate execution.[1] A prerequisite for any quantitative approach to making such tradeoffs is an econometric model of limit-order execution times and the associated execution probabilities.

Limit orders play another important role in determining trading costs: they influence bid/offer quotes and, therefore, spreads. Chung et al. (1997) estimate that 21% of the quotes in their sample originate from limit orders on both the bid and offer sides without any direct participation from the specialist. Therefore, limit-order execution times affect the frequency with which quotes are updated and are likely to be a major factor in the dynamics of bid/offer spreads. Moreover, limit-order execution times have been used to measure the overall quality of equity markets (e.g., Battalio et al., 1999; SEC, 1997), hence their determinants can have important implications for the economic consequences of market fragmentation, the practice of "preferencing", and the relative merits of specialist vs. multiple-dealer market structures.

In this paper, we propose and estimate an econometric model of limit-order execution times using actual historical limit-order data. Using survival analysis, which is a well-known statistical technique for modeling failure times and other nonnegative random variables, we are able to estimate the conditional distribution of limit-order execution times as a function of economic variables such as the limit price, order size, and current market conditions. Because limit-order execution times can be interpreted quite naturally as failure times—they are nonnegative, random, and temporally ordered—survival analysis is the most appropriate method for modeling their evolution.

[1] See, for example, Cohen et al. (1981), O'Hara and Oldfield (1986), Glosten (1989, 1994), Easley and O'Hara (1991), Parlour (1998), Chakravarty and Holden (1995), Keim and Madhavan (1995), Belonsky (1996), Harris and Hasbrouck (1996), Kavajecz (1998), Rock (1996), and Seppi (1997).

Moreover, survival analysis can accommodate an important feature of limit-order execution times that existing models have ignored: *censored* observations, i.e., limit orders that expire or are canceled before they are executed. There is great temptation to ignore censored observations since they seem to provide little information about execution times. However, the fact that a limit order is canceled after, say, 30 minutes yields a piece of useful information: the limit order "survived" for at least 30 minutes. Therefore, censored observations do affect the conditional distribution of execution times despite the fact that they are not executions. Ignoring censored observations can dramatically bias the estimator of the conditional distribution of execution times.

Using a sample of limit orders for the 100 largest stocks in the S&P 500 from August 1994 to August 1995, we construct models of limit-order execution times based on survival analysis and show that they fit the data remarkably well. In particular, we estimate separate models for time-to-first-fill and time-to-completion for both buy and sell limit orders, hence four models in all. Each of these four models yields a conditional distribution that closely matches the data's and passes several diagnostic tests of goodness-of-fit. The parameter estimates show that execution times can be quite sensitive to certain explanatory variables, such as market depth, the spread between the limit price and the quote midpoint, and market volatility, implying that the kind of strategic order-placement strategies described by Angel (1994), Foucault (1996), Harris (1994), Hollifield et al. (1999), Kumar and Seppi (1993), and Parlour (1998) could well be feasible in practice. Limit-order execution times can be accurately modeled, hence controlled.

In Section 2 we review the literature on limit orders, and in Section 3 we discuss some of the institutional features of limit orders and describe our limit-order dataset. We present a simple but powerful application of this dataset in Section 4 in which we compare actual limit-order execution times to their *hypothetical* counterpart, constructed theoretically (from the first-passage times of Brownian motion) and empirically (from transactions data). We present a brief review of survival analysis in Section 5 and turn to our empirical analysis in Section 6. We conclude in Section 7.

2. Literature review

There is a large and growing theoretical literature that considers the economic role of limit orders in the price discovery process. Foucault (1993), Glosten (1989, 1994), Easley and O'Hara (1991), Parlour (1998), Chakravarty and Holden (1995), Kavajecz (1998), Rock (1996), Sandås (1999), and Seppi (1997) are just a few recent examples. The general focus of these papers is the effect of limit orders on the market, the interaction between limit and market orders, and the role of the market maker. None of these studies is set in a continuous-trading environment, hence they provide little direct guidance for modeling limit-order execution times.

However, several other studies explore the probability of limit-order execution. For example, under a number of rather strong assumptions, Angel (1994) derives an analytical expression for the probability of limit-order execution, conditional upon an investor's information set. His result applies to batch trading of one round lot of the stock for informed traders, assuming that traders know the entire limit-order book. Within his analytical framework, Angel also conducts some simulations for continuous-trading environments.

Hollifield et al. (1999) build a structural model of a pure limit-order market. The model captures the tradeoff between order price and probability of execution. They estimate their model nonparametrically and derive implications for traders' order-submission strategies.

A number of studies compare the use of market orders to limit orders empirically. In particular, using the NYSE transactions, orders, and quotes (TORQ) data, Petersen and Fialkowski (1994) find that limit orders placed at the quote outperform market orders in markets that quote in $1/8 increments, but underperform market orders in wider markets. Harris and Hasbrouck (1996) use the TORQ data to compare the profitability of order-submission strategies using limit orders vs. market orders. They find that in some cases the use of limit orders can reduce execution costs. Handa and Schwartz (1996) assess the profitability of limit-order trading by comparing unconditional expected returns of market orders vs. limit orders. Their analysis is based on hypothetical limit-order executions, or fictitious executions constructed from transactions data (see Section 4 for further discussion and an empirical critique). Biais et al. (1995) present an empirical analysis of the order flow of the Paris Bourse which is a pure limit-order market. They find that traders' strategies vary with market conditions, with more limit orders at times when spreads are wide and more market orders at times when spreads are narrow.

Other empirical studies have focused on the role of limit orders in determining execution costs and overall market quality. Keim and Madhavan (1995) examine a unique dataset containing information for 62,000 equity orders (each order generating one or more trades) by 21 institutional investors from January 1991 to March 1993, and one of the many issues they consider is the selection of order type (limit order, market order, working order, or crossing network). McInish and Wood (1995) argue that there are "hidden" limit orders on the NYSE, orders that would improve the posted quotes but which are not always displayed by specialists. Using the NYSE TORQ data, Chung et al. (1997) find that bid/offer spreads are heavily influenced by limit orders—posted spreads are widest when there is no competition from the limit-order book, and narrowest when the quotes originate exclusively from the limit-order book. Battalio et al. (1999) compare limit-order fill rates and execution times of primary and regional exchanges to gauge execution quality across markets. The SEC (1997) performs a similar analysis. And Belonsky (1996) provides an extensive cross-sectional analysis of limit orders submitted to the NYSE during the month of February 1994 for all common stocks with SuperDOT activity and prices between $1.00 and $150.00.

Although none of these papers attempts to model the determinants of limit-order execution times, it is apparent that such a model might provide important insights into each of the issues they address.

3. Limit-order data

Although limit orders differ slightly in their institutional features from one exchange to another, we shall focus on those characteristics that are common to the largest exchanges, e.g., the New York and American Stock Exchanges. Upon submission to a designated exchange, a limit order enters the specialist's display book, known as the *order book* or the *queue*. The queue gives top priority to the highest limit buy price and to the lowest limit sell price. Limit orders with the same limit price are prioritized by time of submission, with the oldest order given the highest priority.[2] An order's execution often involves several *partial fills* before it is completed, but partial fills do not change the time priority. A limit order is not binding—it can be *canceled* or *corrected* at any time. Correcting an order does not imply that a mistake has been made, but merely that the original limit order has been revised—either in price or size or both—and resubmitted (and, as a result, has lost its time priority).

When a limit order is submitted, a number of parameters must be specified, including the limit price, whether the order is to buy or to sell, the order size (in shares), the designated exchange, and the *time-in-force*. The time-in-force is the period during which a limit order can be filled. For example, a *day-order* is a limit order that can be filled anytime until the market closes; a *good-till-canceled* order is a limit that can be filled anytime prior to cancellation. The majority of limit orders—82% of the NYSE's TORQ database considered by Harris and Hasbrouck (1996)—are day-orders, although the number of good-till-canceled limit orders is also substantial, about 17% in the same sample.

In addition to the parameters of the order, we would expect the time-to-execution to depend on current market conditions for the stock itself as well as for the market as a whole. Thus, in modeling the time-to-execution it is necessary to specify relevant measures to capture any interaction with market conditions. (We shall return to these issues in Section 6.)

The limit-order data used in this study were provided by Investment Technology Group (ITG), an institutional brokerage firm that provides technology-based equity trading services such as POSIT (an electronic crossing system), QuantEX (a decision-support and routing system), and a full-service trading desk. The ITG limit-order dataset consists of all limit orders submitted through the ITG trading desk from August 1, 1994 to August 31, 1995 for the 100 largest stocks (in market capitalization as of the end of September 1995) in the S&P 500. This dataset is unique in several respects. Each limit order is time-stamped and tracked from submission to termination. After submission, a limit order can be partially or completely filled, canceled or corrected by its initiator, or it can expire if its time-in-force is reached.

[2] On the NYSE, time priority applies only to orders in the limit-order book; an order given to a floor broker may trade ahead of a pre-existing limit-order. Also, time priority is given only to the first order at a given price—after that, all members in the crowd have equal time priority, and the limit-order book may be considered a single floor trader. We are grateful to the referee for pointing out this interesting aspect of the NYSE order-handling rules.

Every action relating to the order during its life is time-stamped, reported, and recorded in the dataset. The submission time is the time when the order departs electronically from its submitter, usually an Exchange-member firm, to the designated exchange. For example, the order can be submitted to the NYSE via the NYSE SuperDOT System. The order is transmitted from the submitter to the exchange almost instantaneously, with a typical delay of less than a second. Once the order is received by the specialist, it is placed in the queue, ready for execution. When the specialist fills the order, a time-stamped report is sent to the submitter. This time stamp is the *report time* and considered the time of execution. When the investor requests cancellation or correction of an order, the submitter informs the exchange and the exchange sends back a cancellation report and the time of cancellation is recorded.

To illustrate the dynamics of typical limit orders, Fig. 1 provides two examples of paths a limit order can follow from submission to termination using data for AT&T. The first panel follows the path of a buy order on December 29, 1994. The order is first submitted at a limit price of $51.250 and then canceled. It is resubmitted at $51.375, corrected, and resubmitted again at $51.500. It is executed at this price. Our analysis treats this sequence as three observations: limit orders at $51.250 and $51.375 that are not executed and a limit order at $51.500 that is executed. The second panel presents the path of a sell order on November 11, 1994. The order is submitted at a limit price $54.750, corrected and resubmitted at $54.625, and then executed. Our analysis treats this sequence as two observations, one at $54.750 that is not executed and one at $54.625 that is filled.

These examples illustrate three possible execution times that we shall distinguish in our subsequent analysis: (1) time-to-cancellation/correction, (2) time-to-first-fill, and (3) time-to-completion. We shall develop separate models for the second and third—they have markedly different properties—and incorporate the first into our estimation procedures for both models.

In addition to presenting aggregate results for the entire sample of 100 stocks, we also provide detailed results for the following 16 individual stocks: Abbott Labs (ABT), American Express Co (AXP), Anheuser Busch Cos Inc (BUD), Chrysler Corp (C), Colgate Palmolive Co (CL), Dean Witter Discover (DWD), General Elec Co (GE), General Mtrs Corp (GM), International Business Machines (IBM), JP Morgan & Co Inc (JPM), Mobil Corp (MOB), Pacific Telesis Group (PAC), Procter & Gamble Co (PG), Sara Lee Corp (SLE), Seagram Ltd (VO), and Xerox Corp (XRX). These 16 stocks will be identified by their ticker symbols hereafter, and we shall refer to the pooled sample of 100 stocks as "POOL".

3.1. TORQ vs. ITG

Because our limit-order dataset comes from a single source, ITG, it might contain certain biases that are not present in the TORQ dataset.[3] For example, ITG's clients are almost exclusively institutional investors and other broker/dealers, hence its

[3] We are grateful to the referee for some of these observations.

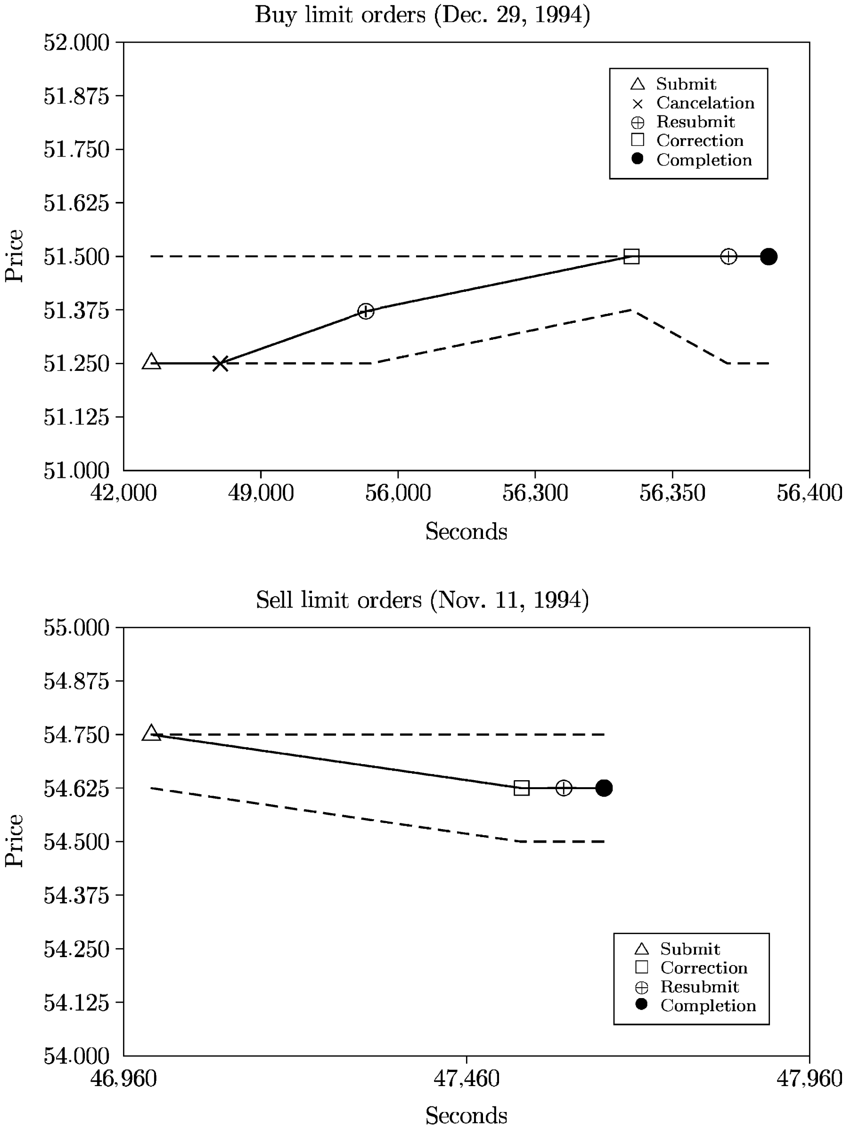


Fig. 1. Sample paths for two particular limit orders. The first panel shows a buy limit order on December 29, 1994 for AT&T and the second panel shows a sell limit order on November 11, 1994 for AT&T. In the upper panel, a buy limit order is submitted at $51\frac{1}{4}$, canceled, resubmitted at $51\frac{3}{8}$, corrected and resubmitted at $51\frac{1}{2}$, and executed at $51\frac{1}{2}$. In the lower panel, a sell limit order is submitted at $54\frac{3}{4}$, corrected and resubmitted at $54\frac{5}{8}$, and executed at $54\frac{5}{8}$.

trading desk sees few retail orders (e.g., small or odd lots). Therefore, the ITG limit-order data reflect this institutional focus which, in turn, affects the econometric models estimated from the data.

Also, ITG's trading platform, QuantEX, is a real-time event-driven interface to multiple sources of liquidity, such as SuperDOT, POSIT, and Instinet. A common function of the ITG trading desk is to handle residual (unfilled) trades from some of these liquidity sources—for example, orders that were unfilled in a POSIT match—hence a portion of the orders in the ITG limit-order dataset arise from an absence of liquidity.

QuantEX also provides data and analytics to support order-routing and order-handling decisions; it is an "expert system" that allows portfolio managers to program trading rules and optimization algorithms to fully automate their trading decisions. Therefore, the conditioning information behind an ITG limit order is likely to be richer than the typical TORQ limit order.

Despite these factors, ITG's limit-order dataset is likely to be an important one, particularly for institutional investors, since virtually every major institution uses POSIT to some degree. Nevertheless, this dataset is only one of many possible datasets, each reflecting the style and customs of a particular set of traders. We hope to show that our application of survival analysis to limit-order execution times is promising enough to motivate others to apply the same techniques to their own datasets, and to compare their findings to ours. We conducted a preliminary comparison of the ITG and TORQ datasets but quickly decided that it was unlikely to be very informative because of the difference in time periods (the TORQ dataset spans November 1, 1990 to January 31, 1991).

3.2. Summary statistics

Summary statistics for the limit-order dataset are reported in Table 1. The number of limit orders per stock ranges from 1,160 (DWD) to 11,298 (GE), and is almost evenly split between buy orders (52.42%) and sell orders (48.58%) for the pooled sample of 375,998 limit orders. Among the sell orders, short sales account for the majority (32.83%). Because shortsale orders are subject to the up-tick rule (whereby a short sale can be executed only if it occurs at a price higher than the preceding transaction at a different price), we expect their dynamics to differ from pure sell orders. For this reason, we omit them from our empirical analysis. Hereafter, by "sell order" we shall mean pure sell orders only.

Once an order is submitted, it can be partially filled, completely filled, or not filled at all due to cancellation or correction (we do not distinguish between these last two conditions). The last three columns of Table 1 report the percentage of orders that are partially filled, completely filled in the first fill, and completely filled, respectively. The orders not included in the "partially filled" category either expire or are canceled. Approximately half the orders are at least partially filled and 37% are completely filled. About 30% are completely filled on the first fill.

Although most of the completed limit orders are completed with the first fill, a number require multiple fills. Table 2 reports the percentage of completed limit orders that are completed with a given number of fills. Over 80% of completed orders are completed with the first fill, and only 1% require seven or more fills.

Table 1
Summary statistics for limit-order data from August 1994 to August 1995 for a pooled sample of 100 stocks (POOL) and for 16 individual stocks listed by ticker symbol.

Stock	Number of observations	% Buy orders	% Sell orders	% Short sales	% Partially filled	% Completed one fill	% Completed multiple fills
POOL	375,998	52.42	14.75	32.83	53.85	30.51	37.47
ABT	4,208	52.23	17.21	30.56	55.44	34.74	41.24
AXP	3,600	49.44	16.31	34.25	51.08	31.81	37.98
BUD	2,640	50.98	16.93	32.08	49.97	31.62	38.32
C	5,606	48.59	12.67	38.74	46.83	23.01	29.32
CL	4,544	51.74	8.19	40.07	43.74	29.38	34.67
DWD	1,160	56.98	32.16	10.86	73.60	35.20	45.84
GE	11,298	50.24	10.28	39.48	48.25	22.44	27.98
GM	6,284	51.00	13.88	35.12	55.70	27.15	34.04
IBM	8,331	55.80	10.89	33.31	52.00	28.29	35.78
JPM	5,485	43.92	20.42	35.66	62.26	35.59	44.15
MOB	6,524	54.52	10.06	35.42	48.94	34.18	39.09
PAC	1,457	56.62	33.70	9.68	70.06	38.60	47.80
PG	6,619	52.97	9.13	37.91	47.74	26.76	33.09
SLE	2,207	60.13	17.90	21.98	58.30	28.16	34.79
VO	1,618	49.94	11.25	38.81	48.69	30.20	36.46
XRX	8,646	54.73	3.19	42.08	33.83	23.92	27.66

Table 2
Percentage breakdown of the total number of completed orders by the number of fills required for completion, for a pooled sample of 100 stocks (POOL) and for 16 individual stocks. The sample period of the data is August 1994 to August 1995.

Stock	Number of fills to completion						
	1	2	3	4	5	6	≥7
POOL	81.42	10.79	3.64	1.66	0.92	0.51	1.05
ABT	84.23	8.96	3.15	1.66	0.66	0.33	1.00
AXP	83.76	8.68	3.34	1.89	1.00	0.89	0.44
BUD	82.53	11.79	3.20	1.02	0.58	0.58	0.29
C	78.45	10.92	4.27	2.38	1.39	0.89	1.69
CL	84.75	10.28	2.75	1.38	0.21	0.21	0.42
DWD	76.79	13.08	5.06	1.90	0.63	1.05	1.48
GE	80.19	9.51	3.66	1.78	1.05	1.20	2.61
GM	79.76	10.37	3.53	1.51	1.51	0.94	2.38
IBM	79.07	11.72	4.48	2.21	1.11	0.40	1.01
JPM	80.62	11.17	3.98	2.12	0.83	0.64	0.64
MOB	87.43	8.14	2.49	1.03	0.36	0.24	0.30
PAC	80.76	10.33	3.18	2.38	1.75	0.95	0.64
PG	80.88	10.00	3.38	2.50	1.18	0.81	1.25
SLE	80.97	8.51	3.67	1.00	2.00	0.50	3.34
VO	82.83	10.80	3.60	1.66	0.55	0.28	0.28
XRX	86.50	10.25	2.02	0.79	0.14	0.14	0.14

Summary statistics for time-to-execution and time-to-censoring are reported in Tables 3a and b. The buy orders are separated from the sell orders. Table 3a reports the mean and standard deviation for time-to-first-fill and time-to-completion and Table 3b reports the mean and standard deviation for time-to-censoring. The mean time-to-execution varies considerably across stocks. The average time-to-first-fill and time-to-completion of PG buy orders is 36.54 and 37.88 minutes, respectively. (Note that more observations are used to calculate the first fill numbers since they include partially filled orders; one should thus be cautious comparing the two times.) For PG sell orders the average times are 10.51 and 10.75 minutes. The PG numbers are representative, although there is some variability. The mean time-to-first-fill for the entire sample of buy orders is 29.22 minutes and for sell orders it is 11.37 minutes. The corresponding completion averages are 30.40 and 12.37 minutes.

The means and standard deviations for time until expiration or cancellation in Table 3b are presented for the orders not included in the time-to-execution statistics. The "no fills" columns consist of orders not executed at all and the "partial fills" columns consist of orders that are partially but not completely filled. As would be expected, the time-to-expiration or time-to-cancellation of the non-executed orders is longer than the fill times. For example, for PG, the average time-to-expiration or time-to-cancellation is 61.2 minutes for buy orders vs. an average time of 36.54 minutes to first fill. One consistent trend is that the average time for a buy limit order to be executed (first fill or completion) is longer than that for a sell order. An examination of the limit order dataset explains this result. As we shall see, on average sell limit orders are submitted closer to the bid price than are buy orders to the ask price. This order submission pattern is consistent with sellers being more concerned about immediacy than buyers.

4. Hypothetical limit orders

Before turning to our econometric analysis of limit-order data in Sections 5 and 6, we explore the prospect of studying limit-order execution times indirectly via theoretical and empirical methods for constructing *hypothetical* limit-order executions. In our theoretical approach, we model stock prices as a geometric Brownian motion and capture limit-order execution times as the first-passage time to the limit-price "boundary". The corresponding empirical approach, first proposed by Handa and Schwartz (1996), is based on the same principle but uses transactions data to determine when the limit-price boundary is hit. Although both methods have the virtue of simplicity, a comparison with actual limit-order data reveals some severe biases that make hypothetical limit-order executions unreliable indicators of actual execution times.

4.1. A theoretical approach: first-passage times

From a purely statistical perspective, the execution time of a limit order can be modeled as a "first-passage time" of the stock price process $P(t)$, i.e., the first time

Table 3

(a) Summary statistics for limit-order execution times for a pooled sample of 100 stocks (POOL) and for 16 individual stocks, for the sample period from August 1994 to August 1995. Columns labeled "First fill" report statistics for the time-to-first-fill (in minutes) for limit orders with at least one fill. Columns labeled "Completions" report statistics for the time-to-completion (in minutes) for completed limit orders.

Stock	First fill				Completions			
	Buy orders		Sell orders		Buy orders		Sell orders	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
POOL	27.92	54.91	11.30	26.84	29.07	55.59	12.29	28.10
ABT	25.39	55.84	8.66	16.11	25.79	56.18	9.98	21.00
AXP	29.61	61.58	15.47	37.20	31.42	62.09	16.57	38.58
BUD	28.48	49.02	11.21	23.39	29.38	49.20	12.76	24.78
C	27.99	54.01	13.66	33.76	29.52	55.06	14.50	35.88
CL	31.84	56.42	9.25	17.81	32.84	56.97	10.81	21.17
DWD	9.71	16.29	13.79	27.24	11.27	18.78	15.71	30.64
GE	39.41	65.75	9.47	20.87	40.31	65.72	10.93	22.80
GM	32.98	56.17	12.01	27.56	34.58	57.29	12.69	28.26
IBM	23.41	50.48	6.25	17.09	24.27	51.07	6.51	17.35
JPM	27.11	54.86	7.51	18.56	28.52	55.67	8.99	22.39
MOB	34.50	64.32	7.25	17.23	34.79	63.97	7.92	17.90
PAC	12.25	30.62	11.69	24.65	14.33	31.28	12.22	25.58
PG	36.42	59.94	10.10	28.51	37.64	60.00	10.33	24.74
SLE	21.14	37.32	16.83	34.51	23.28	38.39	20.42	42.34
VO	40.88	73.97	14.28	32.21	42.84	75.00	16.85	36.24
XRX	51.90	67.23	6.43	13.94	53.10	67.37	7.23	15.72

(b) Summary statistics for limit-order time-to-censoring for a pooled sample of 100 stocks (POOL) and for 16 individual stocks, for the sample period from August 1994 to August 1995. Columns labeled "No fills" report statistics for the time-to-censoring (in minutes) for limit orders without any fills. Columns labeled "Partial fills" report statistics for the time-to-censoring (in minutes) for limit orders partially but not completely filled.

Stock	No fills				Partial fills			
	Buy orders		Sell orders		Buy orders		Sell orders	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
POOL	46.92	72.31	34.15	53.94	41.14	55.65	49.16	65.71
ABT	44.50	71.33	35.43	48.06	63.57	75.54	17.92	24.93
AXP	32.34	52.71	47.92	66.82	28.19	22.25	36.63	37.00
BUD	62.04	83.26	39.93	63.61	57.65	86.04	29.38	26.21
C	60.29	80.34	41.13	60.46	33.27	60.83	24.33	21.23
CL	45.57	70.20	21.45	38.43	19.31	13.43	54.10	60.58
DWD	30.78	38.71	40.07	59.78	33.85	24.82	72.06	63.79
GE	31.55	62.27	22.13	36.27	39.46	54.88	49.64	79.32
GM	63.99	88.62	39.93	59.49	49.36	63.33	99.54	65.14
IBM	48.88	77.51	16.82	27.72	32.67	50.93	6.23	5.56
JPM	36.40	62.89	29.61	47.86	31.74	44.38	51.01	55.56
MOB	52.89	77.26	23.10	37.11	80.55	103.36	38.55	39.72
PAC	36.13	46.44	49.70	65.21	23.82	20.55	64.18	73.19

Table 3 (*continued*)

Stock	No fills				Partial fills			
	Buy orders		Sell orders		Buy orders		Sell orders	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
PG	71.41	93.91	37.76	65.70	42.52	70.44	48.90	100.54
SLE	41.92	70.81	36.66	50.77	47.55	38.69	43.04	10.43
VO	48.45	65.97	45.00	57.74	29.33	22.09	65.14	77.39
XRX	51.48	78.14	32.31	55.71	43.82	71.24	35.95	40.12

the transaction price reaches or crosses the limit price.[4] By placing structure on the stochastic process for transaction prices, the statistical properties of execution times can be derived explicitly.

In particular, let the dynamics of $P(t)$ be given by the leading continuous-time specification for stock prices, i.e., geometric Brownian motion with drift:

$$\mathrm{d}P(t) = \alpha P(t)\,\mathrm{d}t + \sigma P(t)\,\mathrm{d}W(t), \tag{4.1}$$

where $W(t)$ is a standard Brownian motion and α and σ are constants. Let t_0 denote the current time and P_0 denote the current stock price. Let $P_{\min}$ denote the lowest price observed in the time interval $[t_0, t_0 + t]$, so that t is the length of the interval.

We assume that a buy limit order with limit price P_l will be executed in the interval $[t_0, t_0 + t]$ if and only if $P_{\min}$ is less than or equal to P_l. Thus, the probability of a limit-order execution is simply the probability that $P_{\min}$ is less than or equal to P_l in $[t_0, t_0 + t]$, i.e., the probability that the first passage of $P(t)$ to P_l occurs within $[t_0, t_0 + t]$. By modifying a formula given in Harrison (1990, p. 14), this probability can be derived exactly under Eq. (4.1) and is given by

$$\Pr(P_{\min} \leqslant P_l | P(t_0) = P_0) = 1 - \Phi\left(\frac{\log(P_0/P_l) + \mu t}{\sigma\sqrt{t}}\right) + \left(\frac{P_l}{P_0}\right)^{2\mu/\sigma^2} \Phi\left(\frac{\log(P_l/P_0) + \mu t}{\sigma\sqrt{t}}\right), \quad P_l \leqslant P_0, \tag{4.2}$$

where $\mu \equiv \alpha - \frac{1}{2}\sigma^2$, $\Phi(\cdot)$ is the standard normal cumulative distribution function (CDF), and the limit buy price P_l is less than or equal to P_0, the current price. A similar expression can be obtained for sell limit orders in the same manner.

Now if we denote by T the limit-order execution time—a nonnegative real-valued random variable—then (4.2) yields the CDF, $F(t)$, for T, i.e.,

[4] Related financial applications of first-passage times include Gottlieb and Kalay (1985), Marsh and Rosenfeld (1986), Ball (1988), Cho and Frees (1988), and Harris (1990).

$$F(t) = \Pr(T \leqslant t | P(t_0) = P_0, P_l, \mu, \sigma) = \Pr(P_{\min} \leqslant P_l) \quad \text{(limit buys)},$$

$$F(t) = \Pr(T \leqslant t | P(t_0) = P_0, P_l, \mu, \sigma) = \Pr(P_{\max} \geqslant P_l) \quad \text{(limit sells)}.$$

The performance of the first-passage time (FPT) model of limit-order executions can then be evaluated by comparing the theoretical CDF, $F(t)$, with the empirical distribution of actual limit-order execution times from our limit-order dataset.

In particular, if actual limit-order execution times T_i are distributed according to (4.2) with CDF $F(\cdot)$, then the random variables $F(T_i)$ must be uniformly distributed on the unit interval $[0, 1]$. Therefore, by tabulating the frequency counts of $F(T_i)$ within, say, each of the deciles of the uniform distribution on $[0, 1]$, i.e., $[0, 0.10), [0.10, 0.20), \ldots, [0.90, 1]$, we can see how closely the empirical behavior of limit-order execution times matches the theoretical predictions of the FPT model.

To do this, we require estimates of the parameters of $F(t)$, i.e., μ and σ. These parameters can be easily estimated from historical data via maximum likelihood:

$$\hat{\mu} = \frac{1}{N\tau} \sum_{j=1}^{N} r_j, \quad \hat{\sigma}^2 = \frac{1}{N} \sum_{i=1}^{N} \frac{(r_j - \hat{\mu}\tau)^2}{\tau}, \tag{4.3}$$

where N is the number of observations in the sample, $r_j \equiv \log P_j - \log P_{j-1}$ is the continuously compounded stock return over a time interval of τ units, and τ is a fixed sampling interval. Over the estimation period from August 1, 1994 to August 31, 1995, and for each of the 16 stocks in our individual-stock sample (see Table 1), we divide each trading day into 13 half-hour trading intervals and calculate the continuously compounded return $r_j = \log P_j - \log P_{j-1}$ over each interval j, $j = 1, \ldots, 13$, where P_j is the average of the bid and ask prices at the end of the jth interval. Then for each stock we calculate the maximum likelihood estimators of μ and σ^2, scaled by 30 to yield per-minute parameter estimates.

By inserting $\hat{\mu}$ and $\hat{\sigma}^2$ into (4.2), we obtain an estimate of the first-passage time CDF $F(t)$ as a function of t, P_0, and P_l. Therefore, for each limit order in our dataset that is executed, we insert its parameters T_i, P_0, and P_l into F to obtain a random variable u_i that contributes towards the frequency count of one of the ten uniform deciles. We use only time-to-first-fill for the FPT model since this most closely matches the notion of a first-passage time. This underscores an important shortcoming of the FPT model: the inability to distinguish among time-to-first-fill, time-to-completion, and time-to-censoring.

But what about the many limit orders that are *not* executed, i.e., those that are canceled or corrected (see Table 3b)? Eliminating them from our frequency count would clearly bias the empirical distribution towards shorter execution times (because we are discarding limit orders that have "survived"), but since they are unexecuted, we cannot evaluate the CDF for these censored observations. Fortunately, a well-known technique for handling censored observations has been developed by Kaplan and Meier (1958), now known as the Kaplan-Meier estimator, and we use this procedure to incorporate limit-order cancellation/correction times into our decile counts.[5]

Table 4 reports the percentage frequency counts of each of the ten uniform deciles for the limit orders of the 16 individual stocks in our sample. It is apparent from the entries in the last column—the tenth decile—that the limit-order data fit the FPT model very poorly. For example, 41.0% of the limit-order execution times of ABT fall into the tenth decile of the FPT model; if the FPT model were correct, this value should be close to 10%. For PG, the empirical value of the tenth decile is even higher at 74.6%. Even the smallest entry in this column—26.4% for GE—is still over twice the theoretical value of 10%, and all of the entries are statistically significantly different from 10%.[6]

The fact that there is a far higher proportion of execution times in the tenth decile than predicted by the FPT distribution (and a correspondingly lower proportion of execution times in the lower deciles) implies that the FPT model vastly *under-estimates* limit-order execution times. In fact, what the FPT model predicts as the 90th percentile of execution times is less than the empirical *median* execution time for DWD, PAC, PG, and SLE limit orders.

Of course, the FPT model (4.2) is predicated on the geometric Brownian motion specification (4.1) for stock prices, and if this specification is not appropriate, it can lead to the kind of inconsistencies documented in Table 4. If, for example, stock prices exhibit short-term mean reversion, e.g., an Ornstein-Uhlenbeck process (see Lo and Wang, 1995), execution times will be longer than predicted by geometric Brownian motion. Unfortunately, explicit expressions for the distribution of first-passage times are unavailable for these more interesting stochastic processes.

The FPT model suffers from several other important limitations. It allows no explicit role for price discreteness, it does not accommodate the impact of limit-order size, it cannot accommodate time priority, it makes no distinction between time-to-first-fill, time-to-completion, and time-to-censoring, and it cannot easily incorporate the effects of explanatory variables such as price volatility, spreads, and market conditions. Therefore, although the FPT model is a natural theoretical framework in which to model limit-order executions, it leaves much to be desired from a practical point of view.

[5] The Kaplan-Meier estimator is a nonparametric method of redistributing the probability mass of censored observations. Specifically, a censored observation indicates that the corresponding uncensored observation must lie to its right, but how far to the right is unknown due to the censoring. The Kaplan-Meier estimator redistributes the probability mass of the censored observation evenly over the portion of the empirical distribution function to the right of the censored observation. In the case of no censoring, the Kaplan-Meier estimator coincides with the conventional empirical distribution function, which assigns a mass of $1/n$ to each observation. See Kaplan and Meier (1958) and Miller (1981) for further discussion.

[6] The asymptotic z-statistics in Table 4 are calculated under the null hypothesis that the FPT model correctly describes the data. In that case, each of the percentage frequency counts $\hat{\pi}_j$ is a consistent estimator of the value 10%, and $N \times \hat{\pi}_j$ is a binomial random variable with mean $N \times 10\%$ and variance $N \times 10\% \times 90\%$ where N is the sample size. Therefore, the z-statistic $\sqrt{N}(\hat{\pi}_j - 10\%)/\sqrt{10\% \times 90\%}$ is asymptotically standard normal.

Table 4
Goodness-of-fit diagnostics for the first-passage time (FPT) model for a sample of 16 individual stocks, for the sample period from August 1994 to August 1995. For each stock, the percentage of execution times that fall within each of the ten theoretical deciles of the FPT model are tabulated. If the FPT model is correct, the expected percentage falling in each decile is 10%. Test statistics that are asymptotically standard normal under the FPT model are given in parentheses.

Stock	Decile									
	1	2	3	4	5	6	7	8	9	10
ABT	10.0	5.6	5.9	6.5	6.2	6.5	6.1	6.1	5.7	41.0
	(0.1)	(−8.2)	(−7.3)	(−6.1)	(−6.7)	(−6.1)	(−6.9)	(−6.9)	(−7.9)	(26.9)
AXP	11.1	6.0	7.0	6.0	6.2	6.0	6.0	5.8	5.6	40.3
	(1.3)	(−6.6)	(−4.7)	(−6.6)	(−6.1)	(−6.5)	(−6.6)	(−7.0)	(−7.4)	(24.0)
BUD	10.6	7.7	6.0	6.3	5.6	6.4	6.5	5.2	5.4	39.4
	(0.7)	(−2.8)	(−5.4)	(−4.9)	(−6.1)	(−4.8)	(−4.6)	(−6.9)	(−6.5)	(19.2)
C	5.6	4.6	4.7	5.3	5.4	7.0	8.3	8.7	7.5	42.9
	(−6.9)	(−9.2)	(−9.1)	(−7.6)	(−7.3)	(−4.3)	(−2.3)	(−1.6)	(−3.4)	(24.0)
CL	5.0	5.2	5.7	7.1	7.4	8.5	9.1	9.3	9.4	33.3
	(−10.2)	(−9.6)	(−8.2)	(−5.0)	(−4.5)	(−2.3)	(−1.4)	(−1.1)	(−0.9)	(21.9)
DWD	5.9	1.8	3.0	3.3	4.1	4.8	3.5	3.3	2.9	66.9
	(−3.8)	(−13.0)	(−8.9)	(−8.1)	(−6.4)	(−5.2)	(−7.5)	(−8.0)	(−9.0)	(25.9)
GE	12.1	7.9	8.4	7.7	8.0	8.2	7.2	7.1	7.0	26.4
	(4.0)	(−4.7)	(−3.6)	(−5.3)	(−4.6)	(−4.0)	(−6.7)	(−7.0)	(−7.4)	(23.1)
GM	11.2	6.5	6.1	6.5	5.7	6.6	6.9	6.5	6.0	38.2
	(1.6)	(−6.1)	(−7.0)	(−6.2)	(−8.1)	(−6.0)	(−5.2)	(−6.0)	(−7.2)	(24.8)
IBM	8.6	5.8	5.9	6.5	6.8	6.9	6.8	6.2	6.4	40.2
	(−2.6)	(−9.6)	(−9.3)	(−7.5)	(−6.8)	(−6.4)	(−6.7)	(−8.3)	(−7.9)	(32.5)
JPM	9.0	6.0	6.6	6.3	5.4	5.4	5.6	5.5	5.1	45.1
	(−1.5)	(−7.3)	(−5.9)	(−6.6)	(−8.8)	(−8.8)	(−8.2)	(−8.4)	(−9.6)	(30.4)
MOB	10.3	6.7	7.1	6.3	6.9	6.3	7.1	7.2	7.3	34.8
	(0.6)	(−6.6)	(−5.7)	(−7.5)	(−6.2)	(−7.5)	(−5.6)	(−5.3)	(−5.2)	(25.9)
PAC	7.0	2.3	2.2	2.1	3.2	3.2	2.6	2.9	2.5	71.9
	(−2.9)	(−12.7)	(−12.8)	(−13.4)	(−9.3)	(−9.3)	(−11.5)	(−10.3)	(−11.9)	(33.5)
PG	0.0	0.0	0.3	0.5	1.1	1.6	2.4	4.6	13.8	74.6
	(—)	(—)	(−82.1)	(−66.6)	(−42.2)	(−33.3)	(−25.0)	(−12.8)	(5.4)	(73.4)
SLE	7.6	6.4	5.5	5.4	5.6	5.5	4.9	4.7	5.0	49.1
	(−2.9)	(−4.7)	(−6.4)	(−6.6)	(−6.1)	(−6.3)	(−7.6)	(−8.0)	(−7.4)	(25.0)
VO	9.2	7.7	6.0	8.5	5.4	6.7	6.7	6.0	6.3	36.9
	(−0.8)	(−2.2)	(−4.4)	(−1.4)	(−5.3)	(−3.5)	(−3.4)	(−4.4)	(−4.0)	(14.6)
XRX	7.2	6.1	6.7	7.6	9.1	9.3	10.3	10.3	11.1	22.3
	(−5.7)	(−8.7)	(−7.0)	(−4.8)	(−1.6)	(−1.2)	(0.6)	(0.4)	(1.8)	(15.6)

4.2. *An empirical approach: transactions data*

The empirical counterpart to the FPT model is based on first-passage times determined by the historical time series of transactions data. For example, consider a stock XYZ that trades at \$50.875 at 10:37 a.m. on April 19, 1995, and suppose that a buy limit order for XYZ is submitted at that time at a limit price of \$50.500. The first time after 10:37 a.m. that a transaction is observed at a price of \$50.500 or lower, the limit order is considered executed, and the time between this transaction and 10:37 a.m. is considered the limit-order execution time. This approach has been used by Angel (1994), Handa and Schwartz (1996), Battalio et al. (1999), and others.

The primary advantage of such hypothetical limit-order executions over the FPT model is the fact that executions are determined by the historical time series of transactions data, not by geometric Brownian motion. Therefore, if the stochastic process for stock prices exhibits mean reversion or more complex forms of temporal dependence and heterogeneity, this will be incorporated into the empirical model.

To compare actual limit orders with hypothetical ones generated by the empirical model, we apply the following procedure to the limit orders of the 16 individual stocks from August 1994 to August 1995. For every buy limit order in our limit-order database that has at least one fill, we create a matching hypothetical limit, i.e., the submission time and limit price are set to equal those of the actual limit order. The time-to-execution of the hypothetical order is determined by the transaction and quotation (TAQ) database distributed by the NYSE, and involves searching for the first time after submission when the transaction price is less than or equal to the limit buy price. The difference between this time and the submission time is recorded as the time-to-execution for the hypothetical limit order. This time-to-execution will obviously be a lower bound for the actual time-to-execution, hence we shall refer to it as the *lower-bound execution time*. It will equal the actual execution time only if the actual limit order is at the top of the queue or close enough to the top that it is filled with the first incoming sell order. However, Handa and Schwartz (1996) treat this lower bound as the execution time itself.

If we continue to track the stock price after its first-passage time, we can obtain an upper bound to the execution time. The *upper-bound execution time* is either the first time during the day when the transaction price falls *below* the limit buy price or the last time of the day the market price is equal to the limit buy price. This last condition will lead to a downward bias in the upper bound. However, given that it applies less than one percent of the time, any bias will be negligible. If neither of these two conditions is met, we treat the observation as missing.

The means and standard deviations of the lower-bound and upper-bound execution times, as well as those of the actual limit-order execution times (time-to-first-fill), are reported in Table 5. Histograms of the times are presented in Figs. 2a and b. Together, they provide conclusive evidence that lower-bound and upper-bound execution times are poor proxies for actual limit-order execution times. In particular, the distance between bounds is large and the mean actual execution time is not consistently close to either bound. For example, ABT's lower-bound mean is 15.58 minutes and its upper-bound mean is 60.12 minutes, yet its actual mean is 25.39 minutes. The standard deviations also disagree: ABT's lower-bound standard deviation is 50.61 minutes and its upper-bound standard deviation is 83.37 minutes, but the actual standard deviation is 55.84 minutes. Even for a very liquid stock such as IBM, the differences between the moments of hypothetical and actual execution times are substantial: its lower-bound mean is 16.80 minutes, its upper-bound mean is 43.26 minutes, and its actual mean is 23.41 minutes.

Table 5
Comparison of hypothetical time-to-first-fill (lower and upper bounds, in minutes) for limit orders simulated using TAQ data with actual time-to-first-fill for limit orders for 16 stocks, for the sample period from August 1994 to August 1995. The "Actual minus TAQ" column reports the difference between the actual time-to-first-fill and the TAQ lower bound. The z-statistics are asymptotically standard normal under the null hypothesis that the expected difference is zero.

Stock	TAQ hypothetical								
	Lower bound		Upper bound		Actual		Actual minus TAQ		
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	z
ABT	15.58	50.61	60.12	83.37	25.39	55.84	9.81	22.88	12.18
AXP	18.21	57.34	66.12	89.05	29.61	61.58	11.41	23.42	13.21
BUD	18.04	41.34	67.69	85.99	28.48	49.02	10.44	23.68	9.34
C	18.44	48.59	56.48	82.54	27.99	54.01	9.55	24.90	9.94
CL	25.88	51.51	66.71	77.51	31.84	56.42	5.96	19.31	8.44
DWD	5.05	10.49	44.33	61.65	9.71	16.29	4.66	12.35	6.52
GE	24.27	57.46	65.08	82.44	39.41	65.75	15.14	31.93	18.18
GM	19.97	49.47	55.56	74.76	32.98	56.17	13.02	26.95	14.95
IBM	16.80	44.57	43.26	72.17	23.41	50.48	6.61	20.37	12.52
JPM	20.27	49.31	55.41	77.27	27.11	54.86	6.84	22.35	9.67
MOB	28.49	61.53	61.38	83.18	34.50	64.32	6.01	21.94	9.40
PAC	1.82	4.82	57.09	81.41	12.25	30.62	10.43	28.64	7.83
PG	27.65	55.10	66.32	81.25	36.42	59.94	8.77	24.80	11.35
SLE	6.21	24.02	67.25	86.86	21.14	37.32	14.94	27.22	12.25
VO	32.64	70.98	89.32	100.99	40.88	73.97	8.24	23.29	5.93
XRX	44.47	66.47	70.81	80.69	51.90	67.23	7.43	13.46	8.84

Table 5 also reports more formal statistical inferences in the last three columns in which the significance of the difference between the actual-time and lower-bound means is evaluated. The differences are strongly significant for all 16 stocks as shown by the asymptotically standard normal z-statistics—they range from 5.93 (VO) to 18.18 (GE). A similar test using differences between the upper-bound and actual-time means also yields strong rejections, hence we omit them to conserve space.

Figs. 2a and b plot the entire distributions of the lower-bound, upper-bound, and actual execution times of four of the 16, and a comparison of these three distributions reveals that they differ not only in one or two moments but over their entire support. In fact, we have attempted to "shift" the distributions of the hypothetical execution times by using "nth-passage" times in place of first-passage times (as n increases, the mean of the hypothetical execution time also increases). That is, instead of determining the execution time as the first time the transaction price reaches the limit price, let it be the nth time that the transaction price reaches the limit price. This is tantamount to assuming a lower position in the queue, and yields intermediate executions to the lower-bound (top of the queue) and upper-bound (bottom of the queue) cases. But even selecting an n that minimizes the difference between the mean hypothetical execution time and the mean actual time does not yield similar distributions.

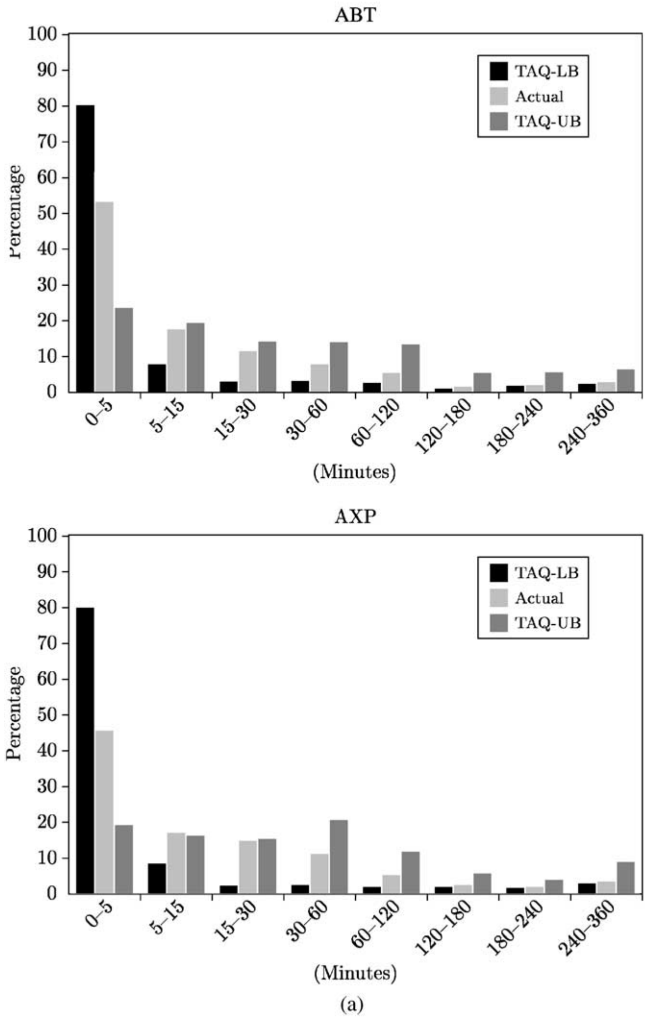
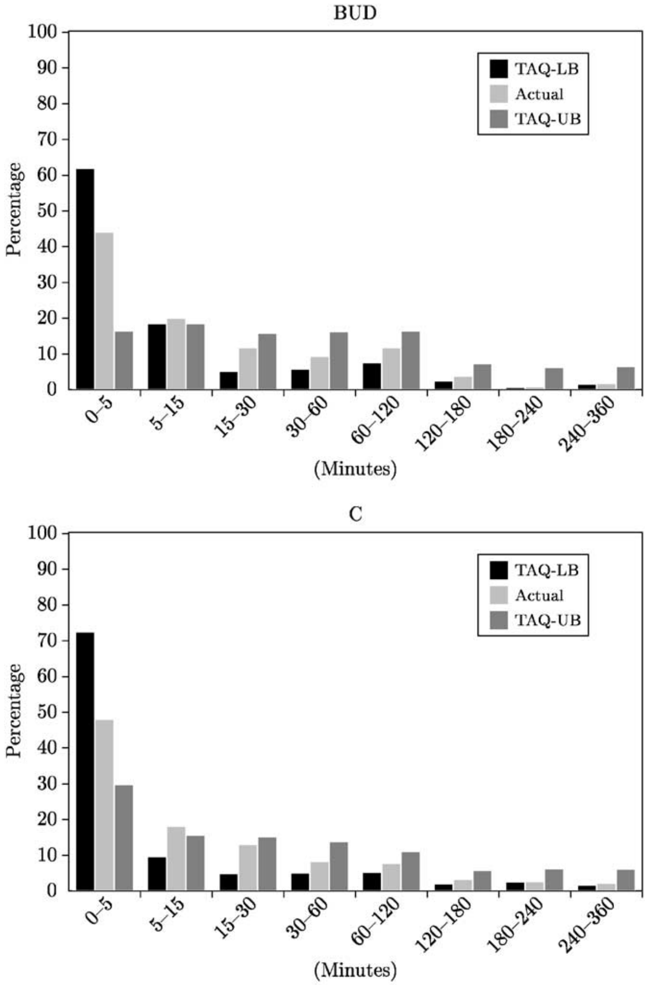


Fig. 2. (a) Histograms of time-to-execution for limit orders for four stocks: ABT, AXP, BUD, and C. "TAQ-UB" and "TAQ-LB" are the upper and lower bounds determined using transaction data, and "Actual" is the actual time to execution.

These results underscore several important weaknesses of the empirical model, the most obvious being the assumption that the hypothetical limit order is executed when the limit price is first attained. Such an assumption implicitly presumes that there are no other limit orders with the same limit price and higher time priority, i.e., the hypothetical limit order is assumed to be at the "top of the queue". However, even intermediate hypothetical execution times such as the nth-passage time and lower-bound models cannot match the empirical distribution of actual limit-order execution times. Moreover, as in the theoretical FPT model, the empirical model



(b)

Fig. 2. (*continued*) (b) Histograms of time-to-execution for limit orders for four stocks: ABT, AXP, BUD, and C. "TAQ-UB" and "TAQ-LB" are the upper and lower bounds determined using transaction data, and "Actual" is the actual time to execution.

cannot easily handle varying limit-order sizes, explanatory variables, and the distinction between time-to-first-fill, time-to-completion, and time-to-censoring.

In summary, hypothetical limit-order execution times are very poor substitutes for actual limit-order data.

5. Survival analysis

To develop an econometric model of limit-order execution times, it is important not only to distinguish between the various execution possibilities, but to incorporate all the characteristics of the order and capture the influence of market conditions. We accomplish this through the application of a well-known statistical technique called *survival analysis*.

Since a limit order can require multiple fills, we must distinguish between time-to-first-fill and time-to-completion. Recognizing this distinction, we estimate two separate models, one for first fills and one for completions. Moreover, since market conditions can affect execution times differently for buy limit orders and sell limit orders, we also estimate separate models for buy orders and sell orders. Thus, we estimate four separate models in all.

For each model, we seek to estimate the following conditional probability, essentially the CDF of the execution time T_k of the kth limit order:

$$\Pr(T_k \leqslant t \mid \mathbf{X}_k, P_{lk}, S_k, I_k), \tag{5.1}$$

where $\mathbf{X}_k$ is a vector of "explanatory" variables that captures market conditions and other conditioning information at the time of submission for the kth limit order, and P_{lk}, S_k, and I_k are the limit-order price, size (in shares), and side indicator (buy or sell), respectively, of the kth limit order.

5.1. A brief review of survival analysis

Survival analysis is a statistical technique for analyzing positive-valued random variables such as lifetimes, failure times, or, in our case, time-to-execution. It is particularly useful for modeling the time-to-execution of limit orders because censored observations (orders terminated prior to execution) can be easily and correctly accommodated. In this section we present a brief review of survival analysis; readers interested in a more detailed exposition should consult Cox and Oakes (1984), Kalbfleisch and Prentice (1980), and Miller (1981).

Let T denote a nonnegative random variable that represents the lifetime of an item, also known as the *failure time*—in our application, it is a limit-order execution time. Let $f(t)$ and $F(t)$ denote the probability density function (PDF) and CDF, respectively, of T. The *instantaneous failure rate* or *hazard rate* of T at time t, denoted by $h(t)$, is defined as

$$h(t) = \frac{f(t)}{1 - F(t)}$$

since $h(t)\,\mathrm{d}t$ is the probability that an item that has survived through time t will fail in the interval $[t, t + \mathrm{d}t)$. Alternatively, we can define the *survivor* function, $S(t) \equiv 1 - F(t)$, which is the probability that an item's lifetime will be at least t. Any one of these four quantities—the PDF, the CDF, the hazard rate, and the survival function—uniquely determines the other three, and all are the focus of survival analysis.

There are two general approaches to estimating these functions: parametric and nonparametric. Parametric survival analysis, described below, assumes a specific parametric family for the distribution of failure times, such as the generalized gamma distribution. Given the distributional assumption, maximum likelihood estimation can be performed. Nonparametric survival analysis involves estimating the survival function without resorting to any parametric assumptions. In this paper, we use the parametric approach. Its dominance over the nonparametric approach for this application is shown in Lo et al. (1999).

The parametric approach to survival analysis begins with the specification of the distribution of the random variable T, from which the likelihood function is obtained. Let $(t_1, \ldots, t_n)$ denote a sequence of n realizations of T, possibly with censoring. We assume that we know which observations have been censored (limit-order cancellations and corrections are reported) and let $(\delta_1, \ldots, \delta_n)$ denote censoring indicators:

$$\delta_i \equiv \begin{cases} 1 & \text{if observation } i \text{ is censored,} \\ 0 & \text{if observation } i \text{ is not censored.} \end{cases} \tag{5.2}$$

If the pairs (t_i, δ_i) are statistically independent, then the likelihood function for the data is given by

$$\prod_{i=1}^{n} f(t_i)^{\delta_i} S(t_i)^{1-\delta_i} = \prod_{U} f(t_i) \prod_{C} S(t_i), \tag{5.3}$$

where U and C denote the indexes of the uncensored and censored observations, respectively. Given the likelihood function (5.3), the parameters of the distribution of T can be estimated via maximum likelihood.

The assumption of independence of (t_i, δ_i) is a restrictive one for limit-order execution times. Therefore, it is useful to note that the likelihood function (5.3) is appropriate under more general assumptions for the dependence structure of the data. In particular, as long as the censoring mechanism for each observation (t_i, δ_i) is independent of the probability that the limit order is executed (conditional on a vector of explanatory variables $\mathbf{X}_i$), the likelihood function of the sample is given by

$$\prod_{i=1}^{n} f(t_i; \mathbf{X}_i)^{\delta_i} S(t_i; \mathbf{X}_i)^{1-\delta_i} = \prod_{U} f(t_i \mathbf{X}_i) \prod_{C} S(t_i; \mathbf{X}_i). \tag{5.4}$$

Therefore, execution times t_i can be dependent and related to $\mathbf{X}_i$, but at any time t and for a given $\mathbf{X}$, the censoring mechanism must be independent of the likelihood that the limit order is executed. This type of censoring, known as *independent censoring*, includes cases in which the censoring mechanism depends on previous execution times, previous censoring times, or on the explanatory variables $\mathbf{X}_i$. However, censoring as a result of prices moving away from the limit price would be a violation of the underlying assumption since prices at the time of censoring are not included in $\mathbf{X}_i$. Kalbfleisch and Prentice (1980, Chapter 3.2) provide a detailed discussion of the role that explanatory variables play in survival analysis including the specific assumptions underlying (5.4).

There are several widely used distributions for failure times, such as the exponential, gamma, Weibull, lognormal, and inverse Gaussian (e.g., Cox and Oakes, 1984, Table 2.1). We choose the *generalized gamma* distribution, which nests a number of other distributions as special cases. Given this nesting, we can test the restrictions imposed on the generalized gamma specification by the simpler cases to see if the other specifications are adequate.

The generalized gamma distribution has three parameters: two shape parameters κ and p, and one scale parameter λ. Its PDF is

$$f(t) = \frac{\lambda|p|\kappa^{\kappa}(\lambda t)^{p\kappa-1}\exp(-(\lambda t)^{p}\kappa)}{\Gamma(\kappa)} \tag{5.5}$$

and the corresponding survival function is

$$S(t) = \begin{cases} \Gamma(\kappa,(\lambda t)^{p}\kappa)/\Gamma(\kappa) & \text{if } p<0, \\ 1-\Gamma(\kappa,(\lambda t)^{p}\kappa)/\Gamma(\kappa) & \text{if } p>0, \end{cases} \tag{5.6}$$

where $\Gamma(a,b)$ denotes the incomplete gamma function and $\Gamma(a)$ denotes the complete gamma function.

When $\kappa = 1$, the generalized gamma distribution reduces to a Weibull distribution, which has PDF

$$f(t) = \lambda|p|(\lambda t)^{p-1}\exp(-(\lambda t)^{p}). \tag{5.7}$$

When $\kappa = 1$ and $p = 1$, the generalized gamma reduces to an exponential distribution, and when $\kappa = 0$, it reduces to a lognormal distribution.

5.2. Incorporating explanatory variables

As we have noted, the incorporation of explanatory variables into the likelihood function (5.4) poses no difficulties, and allows execution times to be dependent as long as restrictions are placed on the censoring mechanism. The dependence of failure times on explanatory variables is addressed by assuming that the effect can be captured by rescaling time. This formulation is commonly called the *accelerated failure time* specification, and an exponential factor is often used to rescale time.[7]

Specifically, an accelerated failure time model has the form

$$T = \mathrm{e}^{\mathbf{X}'\boldsymbol{\beta}}T_0,$$

where T is the time-to-execution, $\mathbf{X}$ is a vector of explanatory variables, $\boldsymbol{\beta}$ is a parameter vector, and T_0 is called the *baseline failure time* and its distribution the *baseline distribution*. The time-to-execution T is then a scaled transformation of the baseline time T_0, where the explanatory variables and coefficients determine the

[7] Another popular alternative is to assume that the hazard rate $h(t)$ satisfies

$$h(t;\mathbf{X}) = h_0(t)\mathrm{e}^{-\mathbf{X}'\boldsymbol{\beta}},$$

where $h_0(t)$ is called the *baseline hazard rate*. For obvious reasons, this is known as the *proportional hazard rate* specification. In most applications, the functional form of $h_0(t)$ can be estimated nonparametrically, hence the proportional hazard rate specification falls within the nonparametric framework. Lo et al. (1999) report estimates of this proportional hazard rate specification and conclude that it does not fit the data as well as the generalized gamma model with an accelerated failure time specification. For example, in contrast to the Q–Q plots of Fig. 3 for the generalized gamma model (see below), the corresponding Q–Q plots for the nonparametric specification in Lo et al. show significant deviations from a 45° line, implying model misspecification.

scaling. Because the baseline distribution is typically specified parametrically, the accelerated failure time approach falls within the parametric framework. We next investigate this specification empirically.

6. Empirical analysis

We now turn to the empirical analysis of our limit-order data using the generalized gamma model described above. We first define the explanatory variables, and then present the parameter estimates. To evaluate the specifications, we consider several measures of goodness-of-fit. Finally, we discuss the economic significance of our estimates.

6.1. Explanatory variables

The dependence of time-to-execution on the limit order's characteristics and on current market conditions is captured through the inclusion of explanatory variables. These variables measure the limit order's price relative to the most recent market price and quotes, the size of the limit order, market depth, and other stock-specific characteristics relating to volatility and liquidity. In particular, let $P \equiv$ market price (most recent transaction), $P_l \equiv$ limit price, $P_b \equiv$ bid price, $P_o \equiv$ offer price, $P_q \equiv$ mid-quote price, $S_o \equiv$ offer size, $S_b \equiv$ bid size, $S_l \equiv$ limit-order size.

Then the following are the explanatory variables included in the buy limit-order models (all variables are measured at the time of submission):

$$\mathrm{MQLP} = P_q - P_l,$$

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

$$\mathrm{MKD1} = \begin{cases} (1 + P_b - P_l)\log S_b & \text{if } P_l \leqslant P_b, \\ 0 & \text{if } P_l > P_b, \end{cases}$$

$$\mathrm{MKD1X} = \begin{cases} (P - P_l)\mathrm{MKD1} & \text{if } P \geqslant P_l, \\ 0 & \text{if } P < P_l, \end{cases}$$

$$\mathrm{MKD2} = \begin{cases} \log S_o/(1 + P_o - P_l) & \text{if } P_o \geqslant P_l, \\ \log S_o & \text{if } P_o < P_l, \end{cases}$$

$$\mathrm{SZSD} = \begin{cases} \log(S_l)(1 + P_o - P_l) & \text{if } P_o > P_l, \\ \log(S_l - S_o) & \text{if } P_o = P_l \text{ and } S_l > S_o, \\ 0 & \text{otherwise,} \end{cases}$$

$$\text{STKV} = \#\text{ trades last half hour}/\#\text{ trades last one hour},$$

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

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

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

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

The first eight variables accommodate the dynamic nature of the marketplace by capturing current market conditions. These are updated on a real-time basis. In contrast, the last three variables facilitate differences across stocks and are updated monthly.

The variable MQLP measures the distance between the limit buy price and the current quote midpoint. BSID is an indicator to measure whether the prior transaction was buyer-initiated or seller-initiated (see, e.g., Hausman et al., 1992). MKD1 is a measure of the minimum number of shares that have higher priority for execution scaled by the distance between the limit buy price and the bid price. The variable MKD1X is an interactive term to capture nonlinearities between market depth and the market price relative to the limit buy price. MKD2 is a measure of the liquidity available from the selling side of the market. The measure is constructed to decline as the limit buy price decreases below the offer price. SZSD is a measure of liquidity demanded by the limit order scaled by the distance between the limit buy price and the offer price. STKV is a short-term measure capturing shifts in trading activity; it proxies for high-frequency changes in volatility. TURN is a trading activity measure providing an absolute measure of volatility. LSO is the logarithm of the number of shares outstanding, LPR is the logarithm of share price, and LVO is the logarithm of average daily volume. These are primitive variables included to capture differences across stocks. They can be combined to form a number of measures one might consider including. For example, the log of price plus the log of shares outstanding is the log of market value, the log of volume minus the log of shares outstanding is the log of turnover, and the log of price plus the log of volume is approximately the log of dollar volume.

Four of the explanatory variables are redefined for the sell limit-order models. The definitions are altered so that the underlying economic interpretation of these variables is retained (although the direction of the effect may be reversed). The redefined variables are listed below:

$$\text{MKD1} = \begin{cases} (1 + P_l - P_o)\log S_o & \text{if } P_l \geqslant P_o, \\ 0 & \text{if } P_l < P_o, \end{cases}$$

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

$$\text{MKD2} = \begin{cases} \log S_b/(1 + P_l - P_o) & \text{if } P_o \leqslant P_l, \\ \log S_b & \text{if } P_o > P_l, \end{cases}$$

$$\text{SZSD} = \begin{cases} \log(S_l)(1 + P_l - P_b) & \text{if } P_l > P_b, \\ \log(S_l - S_b) & \text{if } P_l = P_b \text{ and } S_l > S_b, \\ 0 & \text{otherwise.} \end{cases}$$

Summary statistics and correlation matrices for the explanatory variables are available in Lo et al. (1999), but for completeness we provide a brief overview here. For the 100 stocks in aggregate, there is considerable variation in the explanatory variables as well as differences across buy and sell orders. For example, consider the variable MQLP. On average, the limit buy price is almost one-quarter below the quote midpoint. However, there is substantial variation: the standard deviation for this variable is over one-quarter. In contrast, the mean of MQLP for sell orders is -0.0373, indicating that the limit sell price is only slightly above the quote midpoint on average. Also, there is much less variation for sell limit orders: this variable has a standard deviation of only 0.0847. Similar observations hold for the other explanatory variables.

The cross-correlations of the explanatory variables are generally relatively small, with most being less than 30% in magnitude. For example, the highest correlation between the variable STKV, which captures changing volatility, and the variables related to the limit order is 8.6%. A similar observation holds for TURN, the other volatility-related variable. Exceptions to the low correlation are market-depth variables, which are more highly correlated with each other. For example, the correlation between MKD1 and MKD2 is 59.1% for sell limit orders. Most of the results are similar across buy and sell limit orders, with the exception of the correlation of BSID with the other market-depth variables, which is much higher in magnitude for sells than for buys.

6.2. Parameter estimates

We now present estimation results using our limit-order data. The generalized gamma distribution is used for the baseline distribution and the explanatory variables are incorporated using the accelerated failure time approach.

Recall that we estimate four different models: time-to-first-fill for buy limit orders, time-to-first-fill for sell limit orders, time-to-completion for buy limit orders, and time-to-completion for sell limit orders. We estimate each model using the pooled sample of 100 stocks, and we perform specification checks for both the pooled sample as well as for the 16 individual stocks (see below). The specification check using the individual stocks allows us to assess how well the models capture cross-sectional differences in execution times.

As discussed earlier, the accelerated failure time specification assumes that the effect of explanatory variables on the time-to-execution is to rescale the failure time itself. The sign of the coefficient of an individual explanatory variable indicates the direction of the (partial) effect of that variable on the conditional probability of executing the limit order and on the expected time-to-execution. With this specification, the time-to-execution has a generalized gamma distribution and the maximum likelihood approach is used for estimation.

Using the accelerated failure time specification and the generalized gamma for the baseline distribution, we obtain $f(t;\mathbf{X})$, $S(t;\mathbf{X})$, and the likelihood function by replacing λ by $\exp(-\mathbf{X}'\boldsymbol{\beta})$ in (5.5), (5.6), and (5.3). The density function is given by

$$f(t) = \frac{\exp(-\mathbf{X}'\boldsymbol{\beta})|p|\kappa^{\kappa}(\exp(-\mathbf{X}'\boldsymbol{\beta})t)^{p\kappa-1}\exp(-(\exp(-\mathbf{X}'\boldsymbol{\beta})t)^{p}\kappa)}{\Gamma(\kappa)}. \tag{6.8}$$

Under this specification the model has two parameters in addition to the parameter vector $\boldsymbol{\beta}$: κ and p. In our estimation procedure, we reparametrize the model with $\kappa = 1/v^2$ and $p = v/\sigma$, and estimate it by maximizing the likelihood in (5.3). This reparametrization entails no loss of generality and is purely an artifact of the SAS procedure LIFEREG.[8]

Given the parameter values, we can easily calculate implications of the model for the time-to-execution. For example, the conditional mean of time-to-execution is

$$\mathrm{E}[T|\mathbf{X}] = \exp(\mathbf{X}'\boldsymbol{\beta})(v^2)^{(\sigma/v)}\frac{\Gamma(v^{-2}+\sigma v^{-1})}{\Gamma(v^{-2})} \tag{6.9}$$

and the τth conditional quantile q_τ is given by

$$q_\tau = \exp(\boldsymbol{X}'\boldsymbol{\beta})(v^2)^{(\sigma/v)}(G^{-1}(\tau, v^{-2}))^{(\sigma/v)} \tag{6.10}$$

for $v>0$ (for $v<0$, replace τ by $1-\tau$ on the right side of (6.10)), where $G^{-1}(\tau, v^{-2})$ is the τth quantile of a gamma-distributed random variable with parameter v^{-2}. We shall make use of these formulas below.

Table 6 reports the estimated parameters, along with their corresponding standard errors. The estimates of the parameters associated with the conditioning variables, with only one exception, have the expected signs and generally are statistically significant for all four of the models.

The coefficient on the variable of MQLP is positive with z-statistics of 197 and 190 for the buy models. This indicates that the larger the gap between the mid-quote price and the limit buy price, the longer is the expected time-to-execution. The positive sign on the variable of BSID for buy orders indicates that if the prior transaction has been seller-initiated, a shorter time-to-execution is expected. The positive sign of the estimated coefficient of MKD1 is consistent with the expected time-to-execution increasing with the order size and decreasing with the limit-order price. The negative sign on the variable of MKD2, on the other hand, indicates that the greater the depth of the opposite side of the market and the closer the limit buy price is to the offer, the shorter is the expected time-to-execution. The variable of MKD1X captures a nonlinear relation between time-to-execution and the market price and its depth. The coefficient of SZSD is positive and statistically significant in three of the four models. In the first-fill buy model the coefficient is negative, but not

[8] The LIFEREG procedure fits parametric models to failure time data.

Table 6
Parameter estimates of the accelerated failure time specification of limit-order executions under the generalized gamma distribution for limit orders of a pooled sample of 100 stocks from August 1994 to August 1995. The variable "INTCP" denotes the intercept and the definitions of the remaining explanatory variables are given in the text. z-statistics are asymptotically standard normal under the null hypothesis that the corresponding coefficient is zero.

Variable	Buy limit order model			Sell limit order model		
	Estimate	S.E.	z	Estimate	S.E.	z
Time-to-first-fill						
INTCP	6.507	0.207	31.365	4.979	0.308	16.181
MQLP	8.989	0.046	197.180	−13.674	0.161	−85.034
BSID	−5.613	0.076	−74.168	6.852	0.154	44.543
MKD1	0.641	0.005	127.608	0.476	0.008	59.106
MDD1X	−0.920	0.012	−79.882	0.903	0.058	15.464
MKD2	−0.353	0.005	−66.409	−0.171	0.008	−22.617
SZSD	−0.015	0.005	−3.250	0.091	0.007	13.308
STKV	−0.414	0.052	−7.984	−0.563	0.080	−7.048
TURN	−0.252	0.012	−21.217	−0.331	0.018	−18.757
LSO	0.278	0.014	19.969	0.187	0.021	8.939
LPR	−0.529	0.019	−28.101	−0.272	0.028	−9.872
LVO	−0.082	0.015	−5.563	−0.000	0.021	−0.022
SCALE	1.927	0.006	344.736	1.804	0.008	224.781
SHAPE	−0.404	0.012	−33.781	−0.526	0.018	−29.574
Time-to-completion						
INTCP	6.468	0.212	30.560	5.052	0.317	15.959
MQLP	8.744	0.046	189.979	−13.307	0.163	−81.713
BSID	−5.517	0.077	−71.582	6.766	0.158	42.892
MKD1	0.620	0.005	121.052	0.457	0.008	55.189
MDD1X	−0.895	0.012	−76.409	0.943	0.060	15.708
MKD2	−0.334	0.005	−61.798	−0.148	0.008	−19.122
SZSD	0.069	0.005	14.581	0.186	0.007	25.974
STKV	−0.394	0.053	−7.451	−0.568	0.082	−6.911
TURN	−0.259	0.012	−21.409	−0.327	0.018	−18.045
LSO	0.281	0.014	19.787	0.181	0.022	8.427
LPR	−0.498	0.019	−25.950	−0.229	0.028	−8.069
LVO	−0.092	0.015	−6.139	−0.021	0.022	−0.940
SCALE	1.960	0.006	338.053	1.854	0.008	221.060
SHAPE	−0.410	0.012	−32.965	−0.566	0.018	−30.901

large in magnitude. This is not surprising, since in the case of first fills, we would expect the order size to be less important. The negative signs for the variables STKV and TURN imply that a shorter time-to-execution is expected when market conditions are more active and volatile.

The importance of the three variables included to capture cross-sectional differences is not consistent. This is not of concern, however, since these primitive variables are included to capture a number of composite cross-sectional effects including market value, turnover, and dollar volume. As far as the primitive variables are concerned, the log of share price is the most important. Its coefficient is consistently strongly negative. This is to be expected since higher priced stocks tend to be more liquid. We go beyond the statistical significance of the estimates below, when we consider the economic significance.

Simplifications of the generalized gamma to the Weibull or exponential distribution are strongly rejected. In Table 6, the estimated shape parameter for all models is more than two standard errors from one, the value consistent with the simpler distributions. For example, with the first-fill buy model, the estimate for the shape parameter is −0.404 with a standard error of 0.012. Thus, the estimate is more than 117 standard errors from one. Given the strength of this result, we proceed using the generalized gamma.

The survival function can be easily estimated given the parameters of the model. For the generalized gamma, the estimate of the survival function is

$$\hat{S}(t;X) = 1 - \Gamma(\hat{\kappa}, (\hat{\lambda}t)^{\hat{p}}\hat{\kappa})/\Gamma(\hat{\kappa}) \quad \text{if the estimated } p \text{ is positive,} \tag{6.11}$$

$$\hat{S}(t;X) = \Gamma(\hat{\kappa}, (\hat{\lambda}t)^{\hat{p}}\hat{\kappa})/\Gamma(\hat{\kappa}) \quad \text{if the estimated } p \text{ is negative,} \tag{6.12}$$

where $\hat{\lambda} = \exp(-\mathbf{X}'\hat{\boldsymbol{\beta}})$ for a given $\mathbf{X}$. We shall present diagnostics for each of these specifications below. Overall, the estimates for the generalized gamma accelerated failure time model are in line with our expectations.

6.3. Assessing goodness-of-fit

To check the goodness-of-fit of the versions of the generalized gamma model estimated above, we use two diagnostic measures: a graphical diagnostic (Q–Q plot) and a numerical diagnostic (decile statistics). Both suggest that the generalized gamma model does a good job of capturing the empirical properties of the limit-order data.

6.3.1. Q–Q plots for pooled data

If $S(t;\mathbf{X})$ is the true survival function of the random variable T, then $S(T;\mathbf{X})$, with S as a function of the random variable T, must be uniformly distributed on $[0,1]$. This implies that $-\log S(T;\mathbf{X})$ has an exponential distribution with density e^{-t} for $t>0$, hence we can regard $\{\eta_i \equiv -\log S(t_i;\mathbf{X}_i)\}$ as a sequence of realizations (with censoring) of an exponential random variable. Therefore, testing whether $\{\eta_i\}$ is drawn from an exponential distribution represents one test of whether $S(t;\mathbf{X})$ is the true survival function.

Since the survival function depends on unknown parameters, we cannot work with it directly. Instead, we substitute the sample estimates for the unknown parameters and use the estimated survival function for the analysis. If the model is correctly specified, the estimated survival function $\hat{S}$ will be close to the true survival function, and the sequence $\{\hat{\eta}_i \equiv -\log \hat{S}(t_i,\mathbf{X}_i)\}$ should have properties similar to $\{\eta_i\}$. That is, we can consider $\{\hat{\eta}_i\}$ as a (censored) sample from an exponential distribution, provided that the model is correctly specified. The sequence $\{\hat{\eta}_i\}$ is called the *generalized residuals* (see Cox and Oakes, 1984).

To check this hypothesis, we use Q–Q plots of the negative logarithm of the empirical survival function of the sample $\{\hat{\eta}_i\}$ against the negative logarithm of the theoretical survival function ($-\log e^{-t} = t$). If the model is correctly specified, the plot should be a straight line with a unit slope. Because the generalized residuals can be censored (in particular, whenever the original survival time is censored), we use the Kaplan-Meier estimator. For the gamma model, $\hat{\eta}_i = -\log \hat{S}(t_i, \mathbf{X}_i)$ for $\hat{S}$ given in (6.11).

Since the empirical survival function is subject to sampling variation, we do not expect to see an exact straight line; however, if the model is correctly specified, the plot should show points closely clustered about the 45° line. Q–Q plots that deviate from the 45° line are an indication of model misspecification. Fig. 3 contains the Q–Q plots for the generalized gamma model, and from the relatively straight Q–Q plots for all four of the gamma models, it is apparent that the model is a good fit.

Although the Q–Q plots indicate that the generalized gamma model fits the pooled data quite well, this says little about the performance of the model from stock to stock. Q–Q plots of the time-to-first-fill generalized gamma models for buy and sell limit orders, respectively, using limit-order data for the 16 individual stocks listed in Table 1, show that although there is some variation in the goodness-of-fit of the generalized gamma model across stocks, the pooled model also fits individual limit-order data quite well. The only stock to exhibit a poor fit for both models is GE; for practical purposes it might be worthwhile to estimate a separate model for this one stock. Nevertheless, the generalized gamma model performs admirably stock by stock (plots available upon request). Note that the generalized gamma models are estimated with the pooled data, *not* with individual stock data. The Q–Q plots are constructed stock-by-stock by calculating generalized residuals for each stock using the pooled model and stock-specific limit-order data.

6.3.2. The first-passage time model revisited

For comparison, Fig. 4 contains the estimated survival functions of the theoretical first-passage model (see Section 4.1) for the first four of the 16 individual stocks listed in Table 1—ABT, AXP, BUD, and C. These functions are evaluated at two randomly selected limit orders for each of the four stocks, yielding the eight panels in Fig. 4. For purposes of comparison, the estimated survival function of the generalized gamma model (evaluated for the same two randomly selected limit orders) for the time-to-first-fill of buy limit orders is also plotted. The results for the other 12 stocks are similar (we omit them to conserve space). In contrast to the generalized gamma model estimated on the entire pooled sample, the survival function is estimated individually based on each of the four stocks' estimated drift and diffusion coefficients.

Fig. 4 shows that when the limit buy price is close to or at the market price, the theoretical model underpredicts the time-to-execution. The FPT model predicts that such an order is executed almost immediately and this manifests itself in Fig. 4 as a horizontal line along the horizontal axis. In practice, such an order is typically not executed immediately. For example, transactions occurring at the market price could have been trades on the other side of the market, i.e., sells.

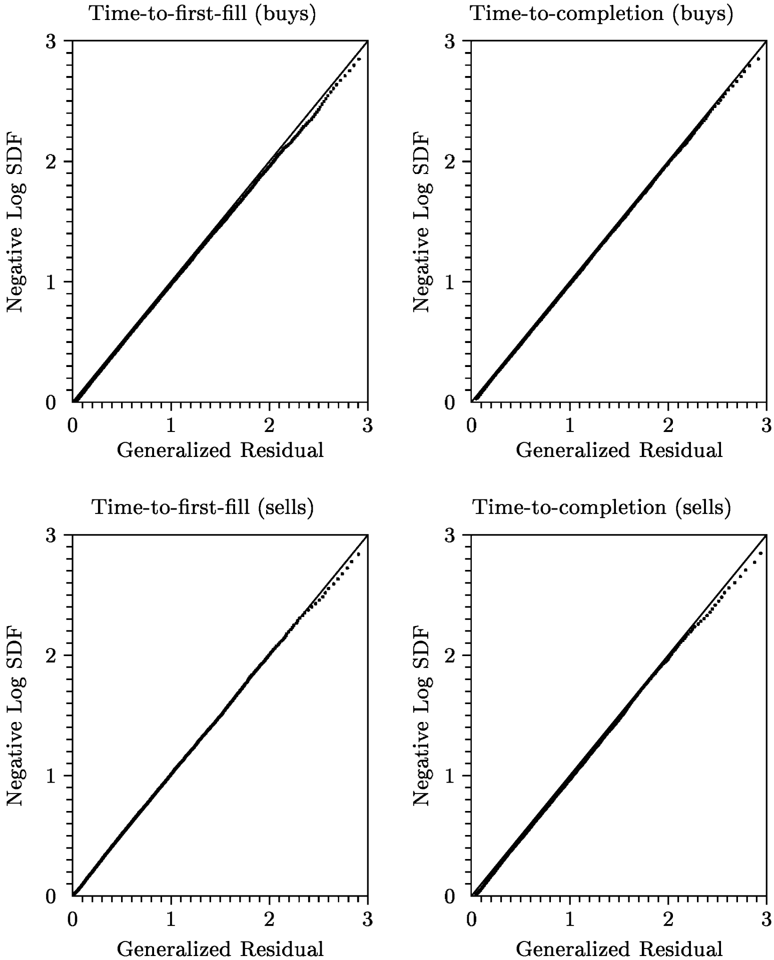
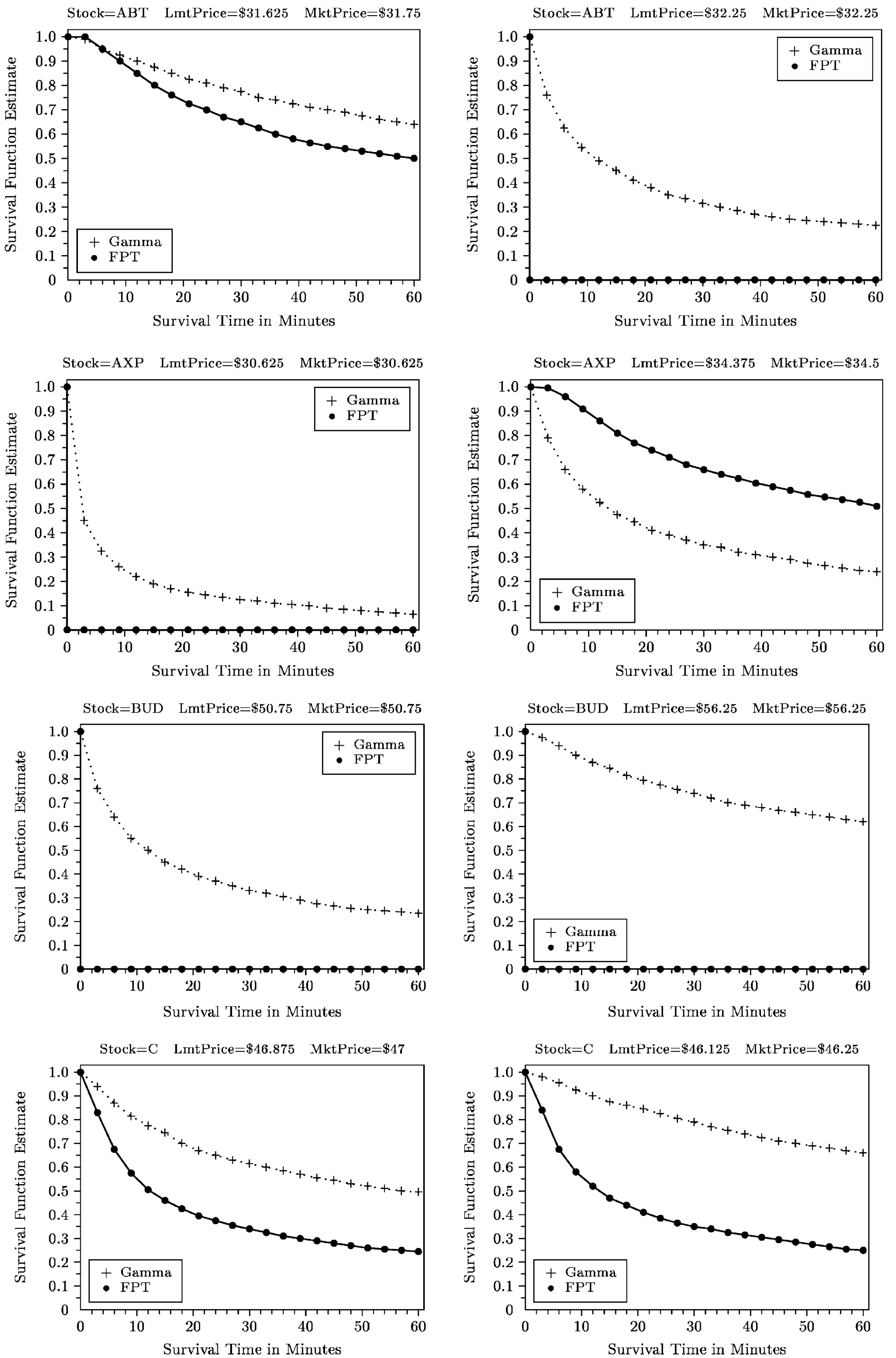


Fig. 3. Q–Q goodness-of-fit plots for four versions of the limit order model using limit orders from August 1994 to August 1995 for a pooled sample of 100 stocks: the time-to-first-fill model for buy limit orders, time-to-first-fill model for sell limit orders, the time-to-completion model for buy limit orders, and the time-to-completion model for sell limit orders.

Moreover, because the generalized gamma model incorporates market information into its model of survival probabilities, it yields more realistic execution times than the FPT model. Interestingly, the two methods have similar predictions when the limit buy price is one tick below the market. But the predictions from the two

→

Fig. 4. Estimated survival functions, gamma vs. first passage time (FPT). Estimated survival functions for two models of limit-order execution times: the generalized gamma model with an accelerated failure time specification, and the first-passage time model based on the assumption that prices follow a geometric Brownian motion. Plots are presented for two randomly selected limit orders for each of four stocks (ABT, AXP, BUD, and C).


Stock=ABT LmtPrice=$31.625 MktPrice=$31.75
Stock=ABT LmtPrice=$32.25 MktPrice=$32.25
Stock=AXP LmtPrice=$30.625 MktPrice=$30.625
Stock=AXP LmtPrice=$34.375 MktPrice=$34.5
Stock=BUD LmtPrice=$50.75 MktPrice=$50.75
Stock=BUD LmtPrice=$56.25 MktPrice=$56.25
Stock=C LmtPrice=$46.875 MktPrice=$47
Stock=C LmtPrice=$46.125 MktPrice=$46.25
Survival Function Estimate
Survival Time in Minutes
Gamma
FPT

models diverge again as the limit price moves away from the market price. The assumption of a geometric Brownian motion tends to imply smaller price changes over short intervals than are observed in the data.

6.3.3. Assessing statistical significance

Although the Q–Q plots in Fig. 3 (as well as those for the individual stocks) suggest that the generalized gamma model is well specified, graphical diagnostics are, of course, meant to be indicative, not conclusive. To gauge the performance of the generalized gamma model quantitatively, we follow the same procedure outlined in Section 4.1 in constructing decile statistics. In particular, we tabulate the frequency counts of the estimated CDF (evaluated at each of the failure times in our sample) for each of the deciles of the uniform distribution on $[0, 1]$, i.e., $[0, 0.10), \ldots, [0.90, 1]$. If the specification is correct, these frequency counts should be close to their theoretical value of 10%.

We report decile statistics for the 16 individual stocks for the buy and sell time-to-first-fill models in Tables 7a and b, respectively. Despite the large sample sizes, there are few decile statistics significantly different from 10% (asymptotic z-statistics are reported in parentheses). For example, in Table 7a the decile statistics range from 9.1% in decile 10 to 12.0% in decile 2 for ABT, and although the decile 2 statistic is statistically different from 10% (with a z-statistic of 2.6), the difference between 10% and 12% is not very meaningful from an economic standpoint. Moreover, when compared to the decile statistics of Table 4 for the FPT model, the statistics in Table 7 show that the generalized gamma model fits very well indeed. Results for the time-to-completion models are similar but are omitted to conserve space.

6.4. Implications of the generalized gamma model

In this section, we go beyond the statistical fit of the generalized gamma limit-order model and consider the implications of the parameter estimates for the specification for limit-order execution times.

To see if there is much variation in the estimated survival function from one limit order to another and with changes in $\mathbf{X}$, we plot in Fig. 5 the estimated survival function $\hat{S}(t)$ of the buy limit/time-to-completion model for three randomly selected buy limit orders for each of four stocks: ABT, AXP, BUD, and C. Each plot also contains the survival function evaluated at the average $\mathbf{X}$ (averaged across the $\mathbf{X}$'s for the three randomly chosen limit orders). From these plots, it is apparent that the estimated survival functions vary considerably from one observation to the next, implying that the distribution of execution times is quite sensitive to conditioning information represented by the explanatory variables.

Figs. 6 and 7 illustrate the sensitivity of the estimated survival function to the limit price and limit shares, respectively, and Table 8 documents the sensitivity of the forecast median execution time to the limit price. In Fig. 6, the estimated survival function is plotted for a single randomly selected limit order for each stock, and the limit price varies from down two ticks to up two ticks, holding all other explanatory variables fixed. Fig. 6 shows that, as expected, the higher the limit buy price, the

Table 7

(a) Goodness-of-fit diagnostics for the accelerated failure time specification of the limit-buy time-to-first-fill model under the generalized gamma distribution for a sample of 16 individual stocks, for the sample period from August 1994 to August 1995. For each stock, the percentages of execution times that fall within each of the ten theoretical deciles of the accelerated failure time specification are tabulated. If this specification is correct, the expected percentage falling in each decile is 10%. Test statistics which are asymptotically standard normal under this specification are given in parentheses

Stock	Decile									
	1	2	3	4	5	6	7	8	9	10
ABT	11.5	12.0	10.4	10.0	10.0	9.2	8.9	9.5	9.6	9.1
	(2.0)	(2.6)	(0.5)	(−0.0)	(0.0)	(−1.2)	(−1.7)	(−0.7)	(−0.7)	(−1.3)
AXP	8.0	10.0	10.2	8.5	10.0	10.8	9.5	11.5	11.1	10.6
	(−2.9)	(0.1)	(0.2)	(−2.2)	(−0.1)	(1.0)	(−0.7)	(1.8)	(1.3)	(0.8)
BUD	8.9	9.4	9.9	10.0	10.2	9.3	10.8	10.3	10.4	10.9
	(−1.3)	(−0.6)	(−0.2)	(−0.0)	(0.2)	(−0.8)	(0.8)	(0.3)	(0.4)	(0.9)
C	7.1	9.1	11.8	9.1	11.4	10.8	10.2	11.1	9.5	10.0
	(−4.1)	(−1.2)	(−2.0)	(−1.2)	(1.6)	(0.9)	(0.3)	(1.2)	(−0.6)	(0.0)
CL	9.0	9.4	8.8	9.5	11.3	10.0	10.0	10.9	10.0	11.0
	(−1.5)	(−0.9)	(−1.9)	(−0.7)	(1.9)	(−0.0)	(−0.0)	(1.3)	(0.0)	(1.4)
DWD	11.0	9.9	9.8	11.6	9.2	9.5	9.3	9.1	10.2	10.4
	(0.7)	(−0.1)	(−0.1)	(1.1)	(−0.6)	(−0.4)	(−0.5)	(−0.7)	(0.2)	(0.3)
GE	10.6	11.2	10.2	11.3	9.9	10.3	9.4	9.5	9.0	8.8
	(1.1)	(2.3)	(0.4)	(2.5)	(−0.2)	(0.6)	(−1.2)	(−1.1)	(−2.2)	(−2.7)
GM	9.2	9.7	10.5	10.3	11.2	10.2	10.4	9.1	9.2	10.3
	(−1.3)	(−0.4)	(0.7)	(0.4)	(1.6)	(0.3)	(0.5)	(−1.4)	(−1.2)	(0.5)
IBM	7.3	10.7	11.8	10.7	10.3	10.8	9.7	9.3	9.3	10.0
	(−5.4)	(1.2)	(2.9)	(1.2)	(0.4)	(1.4)	(−0.6)	(−1.2)	(−1.2)	(0.1)
JPM	9.0	11.2	10.9	11.5	11.2	9.9	10.1	9.0	8.7	8.5
	(−1.5)	(1.7)	(1.3)	(2.0)	(1.6)	(−0.1)	(0.1)	(−1.5)	(−2.0)	(−2.3)
MOB	11.1	11.9	10.2	9.6	9.2	8.9	9.2	9.8	10.0	10.0
	(1.8)	(2.9)	(0.4)	(−0.6)	(−1.4)	(−2.0)	(−1.3)	(−0.3)	(0.0)	(−0.0)
PAC	10.8	12.1	10.2	10.9	8.5	10.9	9.6	9.2	8.8	9.0
	(0.7)	(1.6)	(0.2)	(0.7)	(−1.3)	(0.7)	(−0.4)	(−0.7)	(−1.1)	(−0.9)
PG	11.1	9.0	9.6	10.1	9.9	9.5	10.3	9.6	10.2	10.8
	(1.8)	(−1.7)	(−0.8)	(0.1)	(−0.2)	(−0.9)	(0.5)	(−0.8)	(0.4)	(1.3)
SLE	8.9	9.7	10.4	11.0	10.1	10.7	9.4	10.6	9.9	9.3
	(−1.2)	(−0.3)	(0.5)	(1.0)	(0.1)	(0.7)	(−0.6)	(0.6)	(−0.1)	(−0.8)
VO	11.3	9.8	10.4	10.5	10.5	9.7	9.3	9.7	9.6	9.1
	(1.1)	(−0.2)	(0.4)	(0.4)	(0.4)	(−0.2)	(−0.6)	(−0.2)	(−0.4)	(−0.8)
XRX	9.4	9.4	8.9	10.1	9.4	10.3	9.9	10.3	10.8	11.6
	(−1.0)	(−1.0)	(−2.1)	(0.2)	(−1.1)	(0.5)	(−0.3)	(0.5)	(1.3)	(2.6)

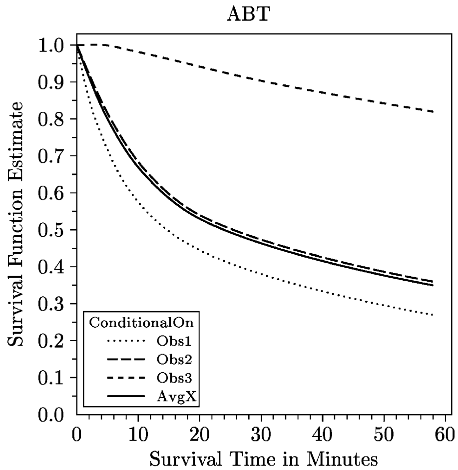
Table 7 (*continued*)

(b) Goodness-of-fit diagnostics for the accelerated failure time specification of the limit-sell time-to-first-fill model under the generalized gamma distribution for a sample of 16 individual stocks, for the sample period from August 1994 to August 1995. For each stock, the percentages of execution times that fall within each of the ten theoretical deciles of the accelerated failure time specification are tabulated. If this specification is correct, the expected percentage falling in each decile is 10%. Test statistics which are asymptotically standard normal under this specification are given in parentheses

Stock	Decile									
	1	2	3	4	5	6	7	8	9	10
ABT	9.9	10.6	13.1	11.7	10.2	9.6	10.3	9.3	7.2	8.0
	(−0.1)	(0.4)	(2.2)	(1.2)	(0.2)	(−0.3)	(0.2)	(−0.6)	(−2.5)	(−1.7)
AXP	7.8	9.8	9.2	9.3	11.0	10.2	8.8	11.8	9.9	12.3
	(−1.8)	(−0.2)	(−0.6)	(−0.5)	(0.7)	(0.1)	(−0.9)	(1.2)	(−0.1)	(1.5)
BUD	12.7	9.7	6.4	9.7	11.1	10.2	10.9	9.0	10.4	9.9
	(1.5)	(−0.2)	(−2.7)	(−0.2)	(0.6)	(0.1)	(0.6)	(−0.6)	(0.2)	(−0.1)
C	8.8	7.6	12.4	9.4	11.4	8.4	11.6	8.4	10.4	11.6
	(−0.9)	(−2.0)	(1.6)	(−0.4)	(1.0)	(−1.2)	(1.1)	(−1.3)	(0.3)	(1.1)
CL	7.2	9.6	10.7	8.8	11.8	9.3	8.9	13.9	11.8	5.7
	(−1.8)	(−0.2)	(0.4)	(−0.7)	(0.9)	(−0.4)	(−0.6)	(1.8)	(0.9)	(−3.0)
DWD	11.0	12.7	9.6	7.1	8.7	7.2	11.0	11.4	9.7	11.6
	(0.5)	(1.3)	(−0.2)	(−1.9)	(−0.7)	(−1.8)	(0.5)	(0.7)	(−0.2)	(0.8)
GE	9.0	10.0	10.1	10.7	10.7	11.7	9.8	10.0	9.6	8.4
	(−0.8)	(−0.0)	(0.0)	(0.6)	(0.6)	(1.3)	(−0.1)	(0.0)	(−0.3)	(−1.5)
GM	7.7	9.0	9.5	10.4	9.9	10.1	9.0	10.8	11.9	11.8
	(−2.1)	(−0.9)	(−0.4)	(0.4)	(−0.1)	(0.1)	(−0.9)	(0.6)	(1.5)	(1.4)
IBM	9.4	10.0	11.1	11.6	7.9	10.6	10.6	10.6	8.8	9.4
	(−0.6)	(0.0)	(1.0)	(1.4)	(−2.1)	(0.6)	(0.5)	(0.5)	(−1.1)	(−0.6)
JPM	8.4	9.0	8.5	12.9	9.7	10.4	8.9	10.2	11.1	10.8
	(−1.6)	(−1.0)	(−1.5)	(2.4)	(−0.2)	(0.4)	(−1.0)	(0.2)	(1.0)	(0.7)
MOB	12.0	12.2	10.6	9.6	9.9	9.5	10.6	7.4	9.5	8.7
	(1.4)	(1.5)	(0.4)	(−0.3)	(−0.1)	(−0.4)	(0.5)	(−2.2)	(−0.4)	(−1.1)
PAC	8.0	7.5	8.4	11.3	12.2	9.4	11.0	9.5	11.8	10.9
	(−1.5)	(−1.9)	(−1.1)	(0.8)	(1.3)	(−0.4)	(0.6)	(−0.3)	(1.1)	(0.6)
PG	12.0	13.5	12.3	10.2	11.1	7.6	10.7	8.0	7.5	7.0
	(1.3)	(2.2)	(1.4)	(0.1)	(0.7)	(−1.8)	(0.5)	(−1.6)	(−2.0)	(−2.4)
SLE	12.0	9.7	9.4	9.7	12.5	9.1	10.0	10.7	9.8	5.5
	(1.0)	(−0.2)	(−0.3)	(−0.2)	(1.2)	(−0.5)	(−0.0)	(0.3)	(−0.1)	(−3.1)
VO	9.5	11.2	13.1	6.8	13.4	7.2	11.2	10.8	7.3	9.6
	(−0.2)	(0.4)	(1.1)	(−1.5)	(1.1)	(−1.2)	(0.4)	(0.3)	(−1.2)	(−0.2)
XRX	12.0	12.1	9.2	6.9	11.7	12.3	7.3	10.5	7.4	10.5
	(0.8)	(0.9)	(−0.4)	(−1.6)	(0.7)	(1.0)	(−1.4)	(0.2)	(−1.4)	(0.2)

higher is the probability of execution over any given time interval. Moreover, the plots show that the survival time is quite sensitive to the limit price, with survival-time probabilities doubling or tripling with just a one- or two-tick change in the limit

Fig. 5. Sensitivity to market conditions. Estimated survival functions for the time-to-completion model for buy limit orders. For each of four stocks (ABT, AXP, BUD, and C), four functions are plotted: the estimated survival function evaluated at three randomly selected limit orders and at the averages of the explanatory variables.

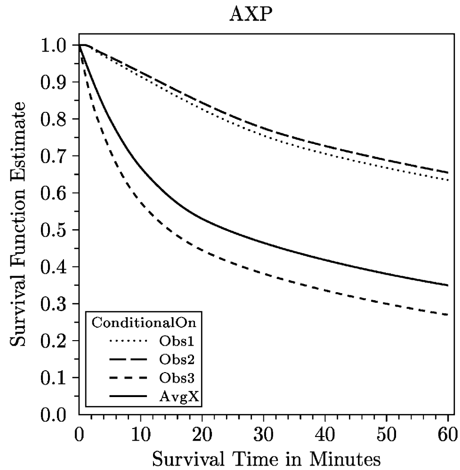


Obs1: LmtP=27.750 LmtS=19 OfrP=27.875 OfrS=19
BidP=27.625 BidS=42 MktP=27.750

Obs2: LmtP=31.625 LmtS=9 OfrP=31.875 OfrS=5
BidP=31.500 BidS=5 MktP=31.625

Obs3: LmtP=36.875 LmtS=1 OfrP=37.000 OfrS=115
BidP=36.875 BidS=350 MktP=36.875

AvgX: LmtP=35.875 LmtS=5 OfrP=36.000 OfrS=12
BidP=35.750 BidS=10 MktP=35.875

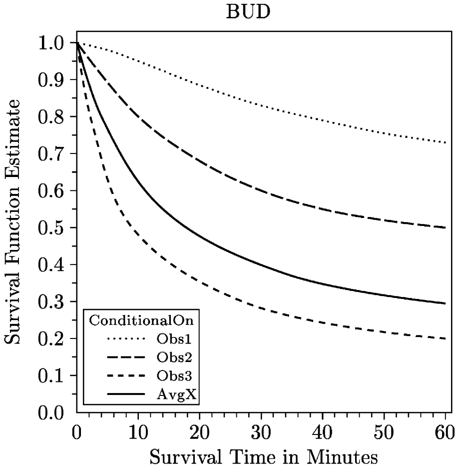


Obs1: LmtP=27.250 LmtS=1 OfrP=27.750 OfrS=1
BidP=27.250 BidS=1 MktP=27.375

Obs2: LmtP=30.125 LmtS=10 OfrP=30.375 OfrS=5
BidP=30.125 BidS=5 MktP=30.250

Obs3: LmtP=34.500 LmtS=4 OfrP=34.625 OfrS=5
BidP=34.375 BidS=5 MktP=34.500

AvgX: LmtP=34.875 LmtS=4 OfrP=35.000 OfrS=9
BidP=34.750 BidS=10 MktP=34.875

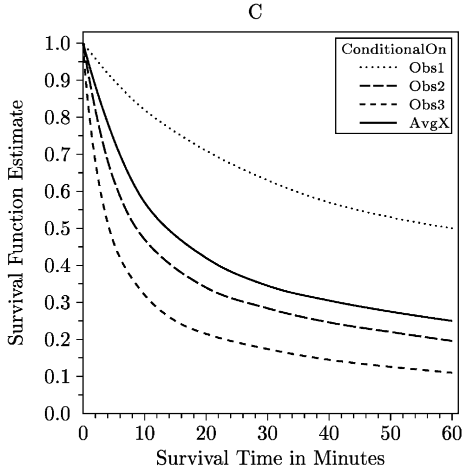


Obs1: LmtP=52.250 LmtS=1 OfrP=52.750 OfrS=3
BidP=52.250 BidS=3 MktP=52.500

Obs2: LmtP=56.625 LmtS=8 OfrP=56.750 OfrS=90
BidP=56.625 BidS=30 MktP=56.625

Obs3: LmtP=56.000 LmtS=2 OfrP=56.125 OfrS=5
BidP=55.750 BidS=5 MktP=56.000

AvgX: LmtP=55.625 LmtS=2 OfrP=55.750 OfrS=7
BidP=55.500 BidS=6 MktP=55.625



Obs1: LmtP=45.625 LmtS=3 OfrP=45.875 OfrS=5
BidP=45.500 BidS=3 MktP=45.875

Obs2: LmtP=48.000 LmtS=4 OfrP=48.125 OfrS=15
BidP=47.875 BidS=17 MktP=48.000

Obs3: LmtP=45.000 LmtS=4 OfrP=45.000 OfrS=50
BidP=44.875 BidS=250 MktP=45.000

AvgX: LmtP=47.500 LmtS=7 OfrP=47.750 OfrS=17
BidP=47.500 BidS=15 MktP=47.750

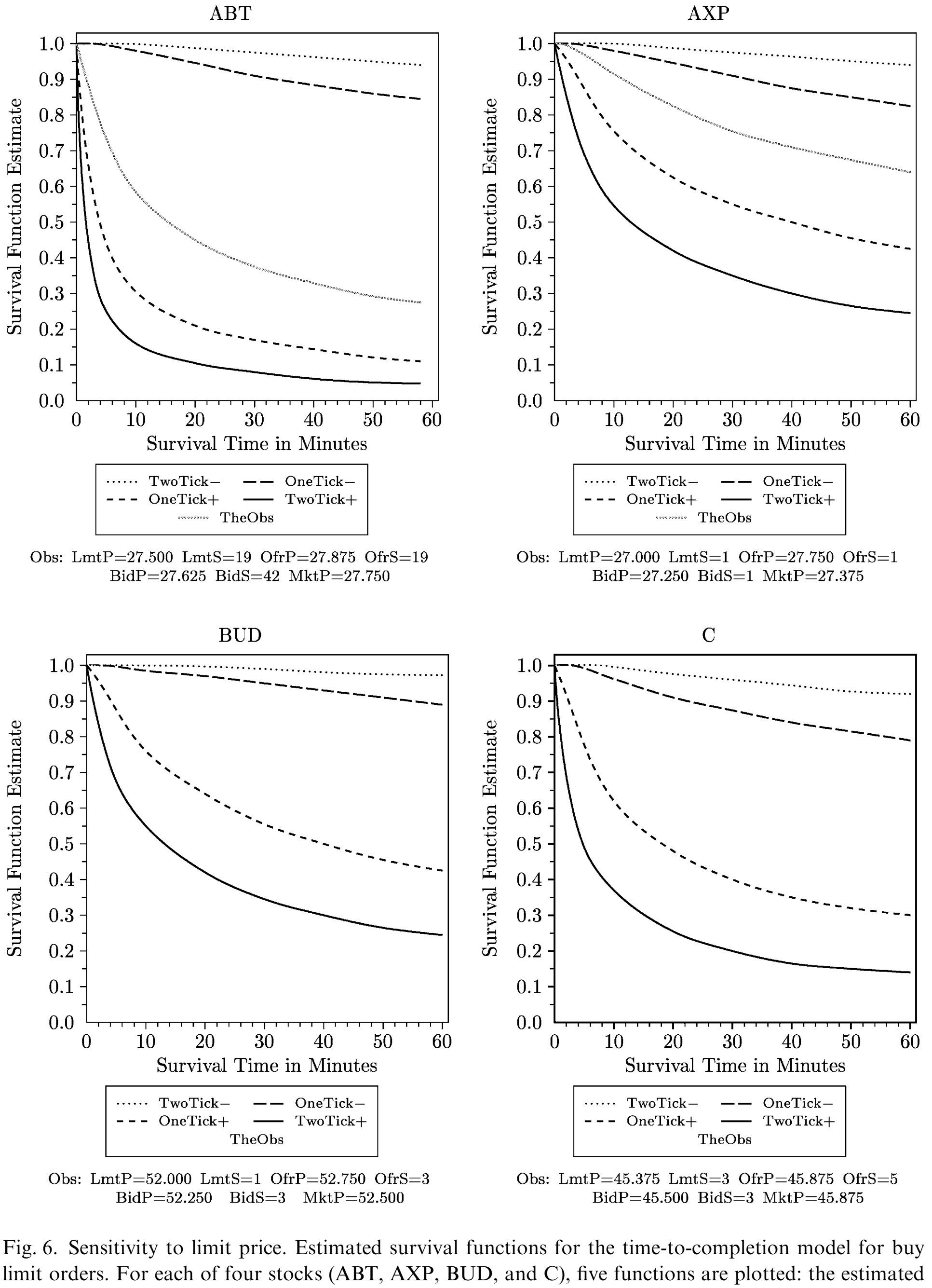


Fig. 6. Sensitivity to limit price. Estimated survival functions for the time-to-completion model for buy limit orders. For each of four stocks (ABT, AXP, BUD, and C), five functions are plotted: the estimated survival function evaluated at a randomly selected limit order and at four different limit-order prices, where the four prices are the actual limit-order price plus and minus one and two ticks, respectively.

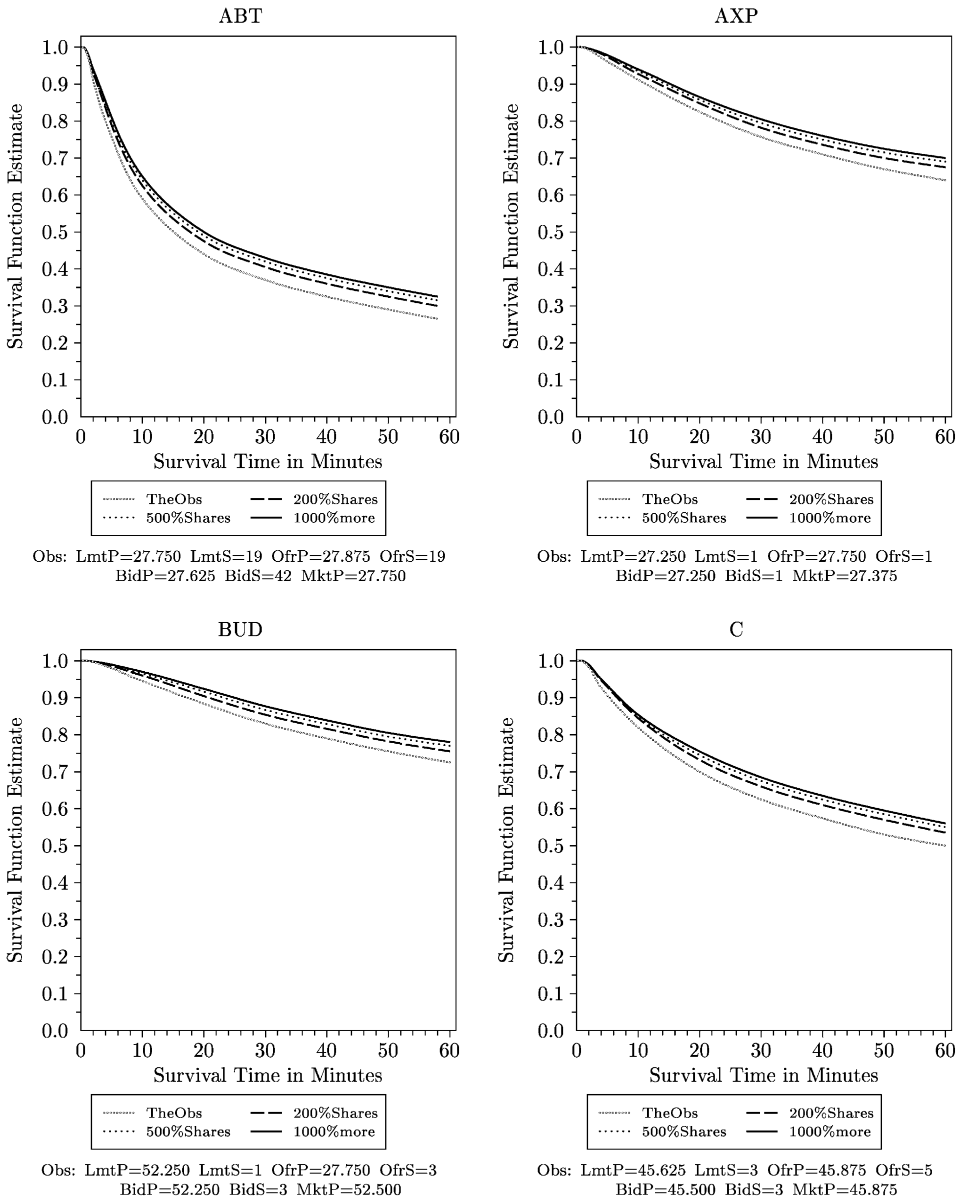


Fig. 7. Sensitivity to limit shares. Estimated survival functions for the time-to-completion model for buy limit orders. For each of four stocks (ABT, AXP, BUD, and C), four functions are plotted: the estimated survival function evaluated at a randomly selected limit order and at three different limit-order sizes, where the three sizes are chosen to be two, five, and ten times the actual limit-order size, respectively.

Table 8

Sensitivity of forecasts of median time-to-completion to limit-order price, using one randomly selected buy limit order from the sample of completed limit orders for each stock. Forecasts are obtained from the accelerated failure time specification of the limit-buy time-to-completion model using the generalized gamma distribution for a pooled sample of 100 stocks for the sample period from August 1994 to August 1995. Forecasts are reported for the limit-order price actually submitted (0 Ticks), for the limit-order price minus one and two ticks, and for the limit-order price plus one and two ticks. The completion time is the actual time-to-completion for the limit order. The order size is in units of round lots (100 shares) and the market conditions used are those at the time the order was submitted.

Stock	Buy limit order		Market conditions			Conditional median execution time					Completion time
	Price	Shares	Bid	Offer	Market	−2 Ticks	−1 Tick	0 Ticks	+1 Tick	+2 Ticks	
ABT	31 + 3/8	18	31 + 1/4	31 + 3/8	31 + 1/4	100.557	23.144	0.128	0.066	0.013	0.167
AXP	30 + 1/2	2	30 + 1/2	30 + 5/8	30 + 1/2	347.512	111.846	30.434	0.351	0.071	18.150
BUD	50 + 1/4	5	50 + 1/4	50 + 3/4	50 + 1/2	3031.245	582.587	111.970	21.520	4.136	30.250
C	47 + 7/8	1	47 + 5/8	48 + 1/2	48	221.718	43.855	8.074	1.716	0.339	8.133
CL	55 + 7/8	8	55 + 3/4	56 + 1/8	55 + 7/8	141.439	26.897	5.115	0.973	0.176	11.133
DWD	47 + 1/8	9	47	47 + 3/8	47 + 1/8	119.146	22.595	4.285	0.813	0.147	2.983
GE	49 + 3/8	4	49 + 3/8	49 + 1/2	49 + 1/2	224.268	62.424	15.586	0.602	0.121	13.967
GM	46 + 3/8	2	46 + 5/8	46 + 7/8	46 + 3/4	369.071	101.893	69.428	47.406	15.670	30.650
IBM	67 + 3/4	50	67 + 5/8	68 + 1/8	67 + 7/8	548.279	99.641	18.108	3.291	0.598	10.350
JPM	60 + 5/8	2	60 + 3/4	60 + 7/8	60 + 7/8	903.572	242.474	60.485	14.002	0.913	7.867
MOB	80 + 1/8	19	79 + 3/4	80 + 1/4	80	21.162	3.861	0.701	0.118	0.024	0.183
PAC	28 + 3/8	2	28 + 3/8	28 + 1/2	28 + 1/2	256.649	77.549	20.356	0.624	0.126	22.217
PG	58 + 1/8	40	58 + 1/8	58 + 1/4	58 + 1/4	626.419	195.124	51.723	0.661	0.133	26.883
SLE	25 + 1/4	119	25 + 1/8	25 + 1/4	25 + 1/4	470.946	139.102	1.472	0.296	0.060	3.483
VO	29 + 7/8	14	29 + 5/8	29 + 7/8	29 + 7/8	100.931	12.559	2.053	0.413	0.083	1.317
XRX	107 + 1/4	1	107 + 1/4	107 + 5/8	107 + 1/2	750.138	181.122	41.588	5.480	1.057	60.467

price. For example, the probability of an ABT buy limit order surviving 20 minutes drops from about 95% to just over 20% when the limit price changes from one tick below to one tick above the original limit price. This limit-price sensitivity is common to most of the limit orders we have examined.

Table 8 contains related results, reporting the sensitivity of the forecast median execution time to the limit price. This table is based on an actual limit order for each stock. The median time is reported for the actual limit-order price and for prices within two ticks in each direction. There can be substantial price sensitivity. For example, the median time for a buy limit order for ABT submitted at the offer price of $31\frac{3}{8}$ is 0.128 minutes. In contrast, if the buy order is submitted with a limit price of $30\frac{1}{8}$, the median time is 100.557 minutes, dramatically longer. Similar sensitivities exist across the other orders.

A similar experiment is conducted with limit shares in Fig. 7: the estimated survival function is plotted for a single randomly selected limit order for each stock, with the limit shares varying from its original value to ten times the original value, holding all other explanatory variables fixed. In contrast to the limit-price graphs of Fig. 6, Fig. 7 shows that the estimated survival functions are much less sensitive to the limit-shares variable. This somewhat surprising finding is even more striking in view of the fact that Fig. 7 is based on the time-to-completion model. Common intuition suggests (and the empirical evidence confirms) that the time-to-first-fill model is even less sensitive to the magnitude of limit shares. This may have important practical implications, for it implies that the size of a limit order has relatively little impact on its time-to-completion (holding other explanatory variables constant). Therefore, adjusting the size of a limit order is a relatively inefficient means of controlling execution times.

Alternatively, the insensitivity of execution times to limit size could be a symptom of a selection bias in our sample: traders might avoid submitting very large limit orders that they judge to be difficult to complete in a timely manner, choosing instead to break up large blocks into smaller orders to be submitted sequentially. Since we are conditioning on limit shares as a regressor, we have no simple way of accounting for this type of censoring in our dataset. We hope to obtain more refined data in the near future to be able to distinguish this possible explanation from others.

7. Conclusion

The behavior of limit-order execution times is critical to the price-discovery process of most market microstructure models, and we have shown that it can be quantified to a large extent by an econometric model based on survival analysis and estimated with actual limit-order data using the ITG limit-order dataset. Survival analysis is designed to model lifetime data and incorporates many of the subtleties that characterize such data, such as skewness and censoring. We find that the generalized gamma model with an accelerated failure time specification fits the data remarkably well, and that execution times are quite sensitive to some explanatory variables (e.g., limit price) but not to others (e.g., limit shares). Despite the fact that we pool the limit orders of 100 stocks to estimate an aggregate model of execution times, our diagnostics show that such aggregate models fit reasonably well stock by stock.

We also explore the properties of hypothetical limit-order executions, constructed theoretically from the first-passage times of geometric Brownian motion and empirically from transactions data. Although such models have a certain elegance due to their parsimony, and can be estimated using transactions data alone, they perform very poorly when confronted with actual limit-order data.

Our findings support the practical feasibility of sophisticated dynamic order-submission strategies, strategies that trade off the price impact of market orders against the opportunity costs inherent in limit orders. We hope to explore such strategies in future research.

References

Angel, J., 1994. Limit versus market orders. Working Paper No. FINC-1377-01-293, School of Business Administration, Georgetown University, Washington, DC.

Ball, C., 1988. Estimation bias induced by discrete security prices. Journal of Finance 43, 841–865.

Battalio, R., Greene, J., Hatch, B., Jennings, R., 1999. Does the order routing decision matter? Unpublished working paper, School of Business, Indiana University, Bloomington, IN.

Belonsky, G., 1996. Contributions, characteristics and rewards of limit orders at the NYSE. In: An analysis of market organization, order strategy and market quality. Ph.D. Dissertation, University of California, Berkeley, CA.

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

Chakravarty, S., Holden, C., 1995. An integrated model of market and limit orders. Journal of Financial Intermediation 4, 213–241.

Cho, D., Frees, E., 1988. Estimating the volatility of discrete stock prices. Journal of Finance 43, 451–466.

Chung, K., Van Ness, B., Van Ness, R., 1997. Limit orders and the bid–ask spread. Unpublished working paper, Fogelman College of Business and Economics, University of Memphis, Memphis, TN.

Cohen, K., Maier, S., Schwartz, R., Whitcomb, D., 1981. Transaction costs, order placement strategy and existence of the bid–ask spread. Journal of Political Economy 89, 287–305.

Cox, D., Oakes, D., 1984. Analysis of Survival Data. Chapman & Hall, New York.

Easley, D., O'Hara, M., 1991. Order form and information in securities markets. Journal of Finance 46, 905–927.

Foucault, T., 1993. Price formation in a dynamic limit order market. Working paper, L'École des Hautes Études Commerciales de Montréal (HEC).

Foucault, T., 1996. Price formation and order placement strategies in a dynamic order driven market. Working paper, Universitat Pompeu Fabra, Barcelona.

Glosten, L., 1989. Insider trading, liquidity and the role of a monopolist specialist. Journal of Business 62, 211–235.

Glosten, L., 1994. Is the electronic open limit order book inevitable? Journal of Finance 49, 1127–1161.

Gottlieb, G., Kalay, A., 1985. Implications of the discreteness of observed stock prices. Journal of Finance 40, 135–154.

Handa, P., Schwartz, R., 1996. Limit order execution. Journal of Finance 51, 1835–1861.

Harris, L., 1990. Estimation of stock variances and serial covariances from discrete observations. Journal of Financial and Quantitative Analysis 25, 291–306.

Harris, L., 1994. Optimal dynamic order submission strategies in some stylized trading problems. Working Paper No. 94-8, School of Business Administration, University of Southern California, Los Angeles, CA.

Harris, L., Hasbrouck, J., 1996. Market vs. limit orders: the SuperDot evidence on order submission strategy. Journal of Financial and Quantitative Analysis 31, 213–231.

Harrison, M., 1990. Brownian Motion and Stochastic Flow Systems. Wiley, New York.

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

Hollifield, B., Miller, R., Sandås, P., 1999. An empirical analysis of a pure limit order market. Working paper, University of Pennsylvania, Philadelphia, PA.

Kalbfleisch, J., Prentice, R., 1980. The Statistical Analysis of Failure Time Data. Wiley, New York.

Kaplan, E., Meier, P., 1958. Nonparametric estimation from incomplete observations. Journal of the American Statistical Association 53, 457–481.

Kavajecz, K., 1998. A specialist's quoted depth as a strategic choice variable. Working paper, Rodney White Center for Financial Research, Wharton School, University of Pennsylvania, Philadelphia, PA.

Keim, D., Madhavan, A., 1995. Anatomy of the trading process: empirical evidence on the behavior of institutional traders. Journal of Financial Economics 37, 371–398.

Kumar, P., Seppi, D., 1993. Limit and market orders with optimizing traders. Working paper, Graduate School of Industrial Administration, Carnegie-Mellon University, Pittsburgh, PA.

Lo, A., MacKinlay, C., Zhang, J., 1999. Econometric models of limit-order executions. Working Paper 12-99, Rodney White Center for Financial Research, Wharton School, University of Pennsylvania, Philadelphia, PA.

Lo, A., Wang, J., 1995. Implementing option pricing models when asset returns are predictable. Journal of Finance 50, 87–129.

Marsh, T., Rosenfeld, E., 1986. Non-trading, market making, and estimates of stock price volatility. Journal of Financial Economics 15, 359–372.

McInish, T., Wood, R., 1995. Hidden limit orders on the NYSE. Journal of Portfolio Management 21, 19–26.

Miller, R., 1981. Survival Analysis. Wiley, New York.

O'Hara, M., Oldfield, G., 1986. The microeconomics of market making. Journal of Financial and Quantitative Analysis 21, 361–376.

Parlour, C., 1998. Price dynamics in limit order markets. Review of Financial Studies 11, 789–816.

Petersen, M., Fialkowski, D., 1994. Posted versus effective spreads. Journal of Financial Economics 35, 269–292.

Rock, K., 1996. The specialist's order book and price anomalies. Working paper, Harvard University, Cambridge, MA.

Sandås, P., 1999. Adverse selection and competitive market making: empirical evidence from a pure limit order market. Working Paper 14-99, Rodney White Center for Financial Research, Wharton School, University of Pennsylvania, Philadelphia, PA.

Seppi, D., 1997. Liquidity provision with limit orders and a strategic specialist. Review of Financial Studies 10, 103–150.

U.S. Securities and Exchange Commission (SEC), 1997. Report on the practice of preferencing.