Abstract

Semiparametric Modeling of Competing Risks in a Limit Order Market

Konstantin Tyurin

2003

This dissertation is based on three papers that have come out of the research conducted at Yale University during 1999–2001 and finished in Indiana University during 2002.

Chapter two introduces the competing risks methodology as an empirical tool for modeling high-frequency financial data in continuous time. The competing risks are applied to the analysis of the timing and interaction between the Deutsche Mark/U.S. dollar quotes and transactions in the Reuters D2000-2 electronic brokerage system. Estimation of the model mostly supports the empirical evidence from previous research on electronic limit order markets. In particular, the composition of order flow is found to be sensitive to the state of the limit order book and the trading history. The direction of past trade is found to have strong predictive power for the future market activity. The model detects an adverse information effect due to non-trading as the traders submit and cancel their orders most aggressively immediately after the limit order book events.

Chapter three studies the problem of semiparametric hazard rate estimation in the competing risks environment. Special attention is paid to the situation where the sample of observed durations is highly skewed, which is fairly common for high-frequency financial data. The chapter provides a review of large sample properties of alternative k-nearest

neighbor estimators and local linear smoothers. The asymptotic theory is applied to the problem of baseline hazard rate estimation for a large number of limit order book events.

Chapter four extends the results of the previous chapters. The set of covariates is expanded to include a broad range of limit order trading and liquidity characteristics. The cross-sectional and serial correlation of Cox regression residuals is captured by the past order flow and the counts of recent transactions. The principal component analysis applied to the covariate indices identifies five pervasive factors that explain a major portion of trading activity. The multifactor modification leads to substantial data compression, improves the goodness-of-fit, and boosts the short-term predictive power of the model relative to popular moving average-type forecasting rules. The competing risks methodology provides a valuable framework for understanding and forecasting the behavior of heterogeneous agents in a competitive market environment.

Contents

1 Introduction 1

1.1 Microstructure Approach to Foreign Exchange Determination 2

1.2 The Challenge of Modeling Market Liquidity 8

1.3 Econometric Approach to Modeling Market Liquidity in Continuous Time . 12

1.4 Overview of the Dissertation 16

2 Competing Risks and the Order Flow Dynamics in FOREX Electronic Brokerage 23

2.1 The Reuters D2000-2 Electronic Dealing System 24

2.1.1 Original Data 26

2.1.2 Reconstruction of Trading History from the Data 28

2.2 Review of Survival Analysis 29

2.2.1 Basic Definitions 30

2.2.2 Estimation of the Cox Proportional Hazard Model for a Single Risk . 30

2.3 Semiparametric Markov Model of Competing Risks 32

2.3.1 Cumulative Baseline Hazard Functions 36

2.3.2 The Baseline Hazard Functions 37

2.4 Estimation Results 39

2.4.1 The Fully Parametric Model 42

2.4.2 Covariates in the Semiparametric Model 44
2.4.3 The Estimates of Baseline Hazard Functions 46
2.5 Conclusion 50
2.6 Appendix: Tables and Graphs 52

3 Statistical Foundations of the Competing Risks Theory: A Counting Process Approach **67**

3.1 Introduction 67
3.2 An Econometric Model of the Limit Order Book 70
3.2.1 Events, Epochs, Sub-Epochs, and the Internal Clock 71
3.2.2 A Counting Process Representation of the Event History 72
3.2.3 Example 1: Two Competing Risks 75
3.2.4 Example 2: Multiple Risks in a Dynamic Limit Order Market Environment 76
3.3 Semiparametric Estimation of the Competing Risks Model 77
3.3.1 The Nelson–Aalen estimators for the Cumulative Hazards and the Incidence Rates 79
3.3.2 Invariance Property of the Maximum Partial Likelihood Estimator . . 81
3.3.3 Alternative Kernel Estimators of the Baseline Hazard Rates 82
3.3.4 Local Linear Smoothing and Bias Reduction 84
3.3.5 Adaptive Nearest Neighbor Estimation of the Baseline Hazard Functions 86
3.3.6 Spline-Based Technique for Hazard Rate Estimation 89
3.4 Asymptotic Theory of Competing Risks Estimators 90
3.4.1 The Partial Likelihood Score Statistic 91
3.4.2 Covariate Effects 93
3.4.3 Cumulative Hazard Functions 93

3.4.4 Local Constant External Estimator of Baseline Hazard 95
3.5 Estimation of Baseline Hazard Rates for Sell Events 96
3.6 Simulation of Competing Risks 97
3.6.1 Alternative Simulation Procedures 98
3.6.2 Performance of Bootstrap Procedures 100
3.7 Conclusion 101
3.8 Appendix: Figures and Table 101

4 Endogenous Determination of Liquidity in Continuous Time: An Index-Based Model of Activity in the Electronic Limit Order Book 111

4.1 The Anatomy of a Pure Limit Order Market 115
4.1.1 Stylized Description of Electronic Limit Order Book 115
4.1.2 Motivation and Limits of the Econometric Approach 118
4.1.3 Empirical Questions 121
4.2 Empirical Analysis of Liquidity and Order Flow in the Electronic Forex Limit Order Market 123
4.2.1 Statistical Methodology 123
4.2.2 Description of Event Types and Covariates 124
4.2.3 Estimation Results 130
4.3 Discussion and Extensions of the Model 137
4.3.1 Implications of Partial Observability of the Limit Order Book 137
4.3.2 An Outline of the Actuarial Approach to Event History 139
4.3.3 Discrete Time VAR Approach to the Joint Dynamics of Price, Volume, and Market Depth 141
4.3.4 Summary of Results 142
4.4 Principal Components of the Competing Risk Indices 143

4.4.1 Relative Contribution of PCA Factors 145
4.4.2 Interpretation of PCA Factors 147
4.5 Simulation Experiments 151
4.5.1 Simulation of Quote and Transaction Histories 152
4.5.2 Simulation of Time to Event 155
4.6 Order Flow Forecasts Based on the Limit Order Book Information 157
4.6.1 Forecasting the Probability of Next Event 160
4.6.2 Evaluating the Performance of Probability Forecasting Models 163
4.6.3 Evaluation of Forecasts by the Method of Relative Cost Analysis . . . 167
4.6.4 Forecasting the Probabilities of Buyer- and Seller-Initiated Transactions 170
4.7 Appendix: Graphs and Tables 175

5 Conclusion and Future Research **208**

List of Tables

2.1 Arrivals and cancellations of market and limit orders (offers) to sell US dollars 52

2.2 Arrivals and cancellations of market and limit orders (bids) to purchase US dollars 53

2.3 Covariates describing the state of the limit order book and the trading history 54

2.4 Covariates describing the recent history of the limit order book 55

2.5 Fully parametric maximum likelihood estimates for competing risks of seller-initiated events in the limit order book (Markov state coefficients omitted) . 56

2.6 Fully parametric maximum likelihood estimates for competing risks of buyer-initiated events in the limit order book (Markov state coefficients omitted) . 57

2.7 Cox regressions for competing risks of seller-initiated events in the limit order book (Markov state coefficients omitted) 58

2.8 Cox regressions for competing risks of buyer-initiated events in the limit order book (Markov state coefficients omitted) 59

3.1 Selected seller-initiated events and graphs of their baseline hazard functions . 104

4.1 Classification of arrival and cancellation events on the sell side of limit order book 195

4.2 Covariates in the Cox regressions of competing risks 196

4.3 Estimated price and quantity coefficients for competing risks of seller-initiated events 197
4.4 Estimated lagged signed order flow coefficients for competing risks of seller-initiated events 198
4.5 Estimated lagged trading activity coefficients for competing risks of seller-initiated events 199
4.6 Estimated price and quantity coefficients for competing risks of buyer-initiated events 200
4.7 Estimated lagged signed order flow coefficients for competing risks of buyer-initiated events 201
4.8 Estimated lagged trading activity coefficients for competing risks of buyer-initiated events 202
4.9 Representation of PCA factor indices in terms of observable characteristics of the limit order book and recent trading history 203
4.10 Cox five-factor regressions for competing risks of sell order arrivals 204
4.11 Cox five-factor regressions for competing risks of buy order arrivals 205
4.12 Cox five-factor regressions for competing risks of limit order cancellations . . 206
4.13 Eigenvalues and cumulative contribution of principal components to the competing risk indices 207
4.14 Modified classification of observable events in the forecasting model 207

List of Figures

1.1 Evolution of transaction price and the cumulative signed order flow in Reuters D2000-2 18

1.2 Phase diagram of transaction price vs. cum trade flow on 10/06/1997 19

1.3 Phase diagram of transaction price vs. cum trade flow on 10/07/1997 20

1.4 Phase diagram of transaction price vs. cum trade flow on 10/08/1997 20

1.5 Phase diagram of transaction price vs. cum trade flow on 10/09/1997 21

1.6 Phase diagram of transaction price vs. cum trade flow on 10/10/1997 21

1.7 Evolution of the limit order book imbalance and the cumulative order flow in D2000-2 22

2.1 The Reuters D2000-2 automated brokerage terminal 60

2.2 Kernel estimates of baseline hazards for seller-initiated observable events . . 61

2.3 Kernel estimates of baseline hazards for buyer-initiated observable events . . 62

2.4 Estimated baseline hazard functions implied by Singh–Maddala parametric form for seller-initiated events 63

2.5 Estimated baseline hazard functions implied by Singh–Maddala parametric form for buyer-initiated events 64

2.6 Kernel estimates of risks for seller-initiated events in a hypothetical scenario 65

2.7 Kernel estimates of risks for buyer-initiated events in a hypothetical scenario 66

3.1 Hypothetical event history in the case of two competing risks 101
3.2 Stylized graphical presentation of the competition between two types of risk 102
3.3 Stylized graphical presentation of the competition between two observable and one unobservable types of risk 103
3.4 Baseline hazard rates for sell market order arrival events (A2) 105
3.5 Baseline hazard rates for sell limit order arrival events at $P^* = P_{\text{bid}}$ (A3) . . 106
3.6 Baseline hazard rates for sell limit order arrival events at $P^* = P_{\text{ask}} - 1$ (A5) 106
3.7 Baseline hazard rates for sell limit order arrival events at $P^* = P_{\text{ask}}$ (A6) . . 107
3.8 Baseline hazard rates for sell limit order arrival events at $P^* = P_{\text{ask}} + 1$ (A7) 107
3.9 Baseline hazard rates for sell limit order arrival events at $P^* = P_{\text{ask}} + 3$ (A9) 108
3.10 Baseline hazard rates for sell limit order arrivals between $P_{\text{ask}}+6$ and $P_{\text{ask}}+10$ (A12) 108
3.11 Baseline hazard rates for limit order cancellation events at $P^* = P_{\text{ask}}$ (AC6) . 109
3.12 Baseline hazard rates for limit order cancellation events at $P^* = P_{\text{ask}} + 1$ (AC7) 109
3.13 Baseline hazard rates for limit order cancellation events at $P^* = P_{\text{ask}} + 3$ (AC9) 110
3.14 Baseline hazard rates for limit order cancellations between $P_{\text{ask}}+6$ and $P_{\text{ask}}+10$ (AC12) 110

4.1 Example of supply and demand curves from Reuters D2000-2 dealing system 175
4.2 Supply and demand curves after the price improvement on demand side . . . 176
4.3 Supply and demand curves after the depth improvement on demand side . . 177
4.4 Supply and demand curves after the bid arrival one tick below the touch . . 178
4.5 Supply and demand curves after the bid arrival two ticks below the touch . . 179
4.6 Effect of subsidiary limit order cancellation on the demand curve 180
4.7 Effect of the large market buy order on the supply and demand curves . . . 181
4.8 Effect of the large limit bid on the demand and supply curves 182

4.9 A small subsample of market bid and ask quotes and transactions in Reuters D2000-2 trading system 183
4.10 Eigenvalues of principal components for the competing risk indices 184
4.11 Log-eigenvalue (LEV) diagram for principal components of the competing risk indices 185
4.12 Cross-correlograms of PCA factors 1 to 5 driving the limit order book dynamics 186
4.13 Autocorrelograms of the five PCA factors driving the limit order book dynamics 187
4.14 Cross-correlograms of activity imbalance, buyer pressure, and bull market momentum (PCA factors 2, 3 and 5) 188
4.15 A representative simulated history of transaction prices, signed order flow, and imbalance of the limit order book 189
4.16 Cumulative distribution functions of time to first seller-initiated transaction as a function of initial state and transaction price 190
4.17 A sample of one-step-ahead forecast probabilities of buyer- and seller-initiated transactions 191
4.18 Reliability plots for one-step-ahead probability forecasts of buyer- and seller-initiated transactions 192
4.19 In-sample reliability plots for probability forecasts of the direction of next trade in 30 sec. 193
4.20 Out-of-sample reliability plots for probability forecasts of the direction of next trade in 30 sec. 194

Semiparametric Modeling of Competing Risks in a Limit Order Market

A Dissertation

Presented to the Faculty of the Graduate School

of

Yale University

in Candidacy for the Degree of

Doctor of Philosophy

by

Konstantin Tyurin

Dissertation Director: Professor Peter C.B. Phillips

December 2003

Acknowledgements

I am deeply grateful to my advisor Professor Peter C.B. Phillips for his excellent guidance and support during the writing of this dissertation. Without his wisdom and encouragement the work contained in this thesis may not have been completed.

My special thanks go to the members of my dissertation committee, Professors Donald W. Andrews and Matthew Spiegel, and to Professor Oliver Linton for their advice and helpful comments on this dissertation.

My passion and inspiration for econometrics and empirical finance arise from the profound influence of the courses that I took at the Department of Economics and Yale School of Management. I appreciate the support of all faculty, staff, and graduate students at the Department of Economics. I am especially thankful to Pam O'Donnell and Mary Moulder, who are always very helpful and supportive.

I gratefully acknowledge the financial support from Cowles Foundation under the Anderson Fellowship and from the Department of Economics under the Dissertation Scholarship. The data set used in this dissertation was provided by Reuters and distributed by the Financial Market Group at London School of Economics whose help is acknowledged.

I also extend appreciation to my colleagues at the Department of Economics of Indiana University for their encouragement during the final stages of my work on this dissertation.

I thank my parents, who were in Russia during my years at Yale, for their support that I felt during all these years.

Finally, I thank Oksana Harbuzyuk for her patience and love.

Chapter 1

Introduction

This dissertation studies various aspects of the microstructural dynamics of an order-driven electronic financial market in the statistical framework of competing risks. In a nutshell, the approach identifies several discrete types of market and limit orders, and models the risks of their arrival, execution, and cancellation in continuous time. The hazard rates are allowed to depend semiparametrically on time since the last observable event and on a linear index of covariates characterizing the past history and current market conditions. The model is sufficiently flexible to incorporate various patterns of unobserved heterogeneity due to time-varying market conditions. To the extent that the structure of competing risks closely replicates the pervasive dynamic factors behind the trading activity, the approach can help researchers of market microstructure and practitioners better understand and interpret the behavior of market participants. Moreover, some empirical and simulation results might trigger interest among traders and their sponsors, who are intrinsically interested in developing dynamic strategies to square existing positions at the minimal cost.

1.1 Microstructure Approach to Foreign Exchange Determination

A large and rapidly growing body of theoretical and empirical literature explores the liquidity, price formation, and order flows in financial markets around the world. The enormous variety of trade organization patterns can be divided in two broad categories determined by their mode of liquidity provision. In *quote driven* markets dealers (market makers or specialists) continuously announce their two-way prices signalling their willingness to trade at or near these price levels. The dealers are approached by interested customers who may wish to trade their desired quantities at the announced prices, while the quotes are continuously updated taking into account the acquired information on the true value of the security and the imbalances in the dealers' own inventories. In *order driven* markets the limit orders submitted by some traders into a limit order book establish the prices at which other participants can trade. While most markets combine certain features of both organizations, an increasing number of exchanges such as the Paris Bourse, the Toronto Stock Exchange, the Tokyo Stock Exchange, the Island ECN, and the Eurex, rely exclusively or primarily on electronically supported limit order books in their provision of the liquidity.

Since the spot foreign exchange market historically developed as a decentralized 24-hour market, its microstructure has several distinct characteristics that are substantially different from those in the exchange-based trading typical for stock markets and in the electronic trading in futures and derivatives. The main segment of the foreign exchange market has been traditionally dominated by voice-based trading of large brokers and interbank dealers. Several electronic systems, most prominent of them Reuters D2000-2 and the Electronic Brokerage Service (EBS), that were actively promoted in the early 1990s to provide competition to traditional brokers and direct interbank dealers, have grown enormously in the

late 1990s.[1] The organization of foreign exchange trading on electronic brokerage systems is similar to the electronic limit order books existing on a number of stock exchanges. Both Reuters and EBS brokerage services are designed as essentially closed order driven systems. The liquidity is supplied by limit orders and consumed by market orders and limit orders directly crossing earlier submitted limit orders on the opposite side of the limit order book. The trading activity on the top of a limit order book (i.e., at or near the current best market buy and sell prices) is observable to any subscriber via the trading screens. However, all information on the identity of traders remains anonymous to other users of the system. The information on the counterparties involved in any transaction only becomes available from an electronic confirmation message sent directly to the agents finalizing the deal immediately after the match in the system occurs.

The microstructure approach to exchange rate determination postulates that the order flow, represented either by signed number of transactions or signed volume of trade, is one of the most important explanatory variables behind the short- and medium-run dynamics of exchange rates. The justification of importance of the signed trade volume as a price discovery channel comes from the observation that the large values of this variable may indicate the presence of informed buyers or sellers trading aggressively in order to exploit their informational advantage. The substantial body of theoretical and empirical literature analyzing the role of order flow as a channel of price formation in the foreign exchange market is nicely summarized by Lyons [96]. In the more narrow context of electronic brokerage, which will be the focus of this thesis, the informativeness of order flow can also be extended

[1]Trading in the D2000-2 system, which is the main subject of the present paper, does not involve direct communication between agents. Therefore it is likely to be a direct competitor primarily to the interbroker trading in the traditional voice-based segment of the spot foreign exchange market. Another Reuters system, D2000-1, that allows a direct electronic contact among the counterparties negotiating the deals can rather be considered a competitor to traditional direct interdealer voice-based trading. Lyons [94] has more on this distinction and its implications for the foreign exchange market microstructure research agenda.

on the ground that the dealers can observe continuously the direction of all trades on their trading screens and easily infer the information on the cumulative signed transaction volume. Among the growing body of research conducted on the Reuters D2000-2 database, we only mention Daniellson and Payne [31], which applied the VAR model of Hasbrouck [66] to study the long-run effect of transaction activity on exchange rates in the D2000-2 electronic brokerage system, Bjønnes and Rime [14], which conducted a comparative study of inventory management by foreign exchange dealers using alternative trading systems, and Hillman and Salmon [72], which applies variogram and kriging techniques to study spatial dependence between the returns and few other variables in the Reuters D2000-2 data set. All these, and many other authors find ample empirical evidence in support of the important role of order flow for exchange rate determination.

For our data, the relationship between the cumulative trade flow and transaction prices can be easily detected on the time series plot (Figure 1.1). Additional insights can be drawn from the scatter plots (phase diagrams) of spot transaction DEM/USD exchange rates against the cumulative signed number of transactions on each of the five consecutive days October 6–10, 1997, which are shown on Figures 1.2, 1.3, 1.4, 1.5, and 1.6.[2] Note that the relationship between these variables was found in the foreign exchange microstructure literature at the relatively low frequencies (c.f. Figure 1.2 in Lyons [96] at the daily frequency, and the results in Daníelsson and Payne [32] at the 15-minute frequency) and confirmed by Luo [92] for a broad spectrum of high and low frequencies.

The basic premise of the empirical market microstructure approach to foreign exchange determination squares nicely with the reported graphs, even though the relationship between

[2] We do not show the scatter plots of transaction prices against the cumulative signed trade volume which is frequently accepted as the definition of order flow. The qualitative relationship between the price and the cumulative signed volume is almost identical, with the scale on the horizontal axis magnified by the size of an average transaction, and the remaining characteristics of the graph virtually unchanged.

the cumulative flow and transaction price never appears strict, nor is it linear. In fact, it would be overly simplistic to expect a strict cointegrating relationship between prices and order flow at the intradaily frequencies, such as the one found at daily frequencies over a period of four months in Killeen, Lyons, and Moore [83]. As the order flow is almost never a uniform entity, additional characteristics of trading counterparties should provide explanatory power and therefore must be incorporated in the model. For example, in most microstructure models, distinction is made between trade and non-trade components of returns, with the former component thought to be driven by information and liquidity motivated trades, and the second component driven by public news announcements. Lyons [96] argues informally that the order flow should be more informative when trade activity is high and less informative when quote activity is high, and presents an empirical evidence supporting the view that the price impact of foreign exchange orders from financial institutions is significantly higher than the price impact of orders from non-financial institutions. In fact, it is always important to keep in mind that the bid-ask spread, volume, and volatility of foreign exchange returns are all determined simultaneously through interaction of numerous traders, who have heterogeneous and potentially quite complex preferences, beliefs, and risk profiles. Even when the data do not allow identification of distinct groups of market participants, the impact of unobserved heterogeneity will be reduced dramatically if it is incorporated into the modeling strategy.

The link between trade flow and nominal levels of exchange rate can be strengthened if one goes beyond the impact of trades initiated by large financial institutions and considers additional observable and unobservable explanatory variables. For example, Luo [92] recently confirmed that the order flow in foreign exchange markets tends to be more informative during the periods of large bid-ask spreads, low trading volumes, and high volatility. This empirical fact was earlier discovered by Dufour and Engle [36] in the stock market, where the periods of intense trading also were found to be associated with stronger effect of trades on the prices. Figure 1.7 provides an example of one such variable, which is represented by

the difference between the number of active buy and sell limit orders in Reuters D2000-2 brokerage system at any moment of time. This variable (called the *imbalance* of the limit order book) represents the willingness of market participants to trade at the prices far above and below the current competitive ask and bid prices, as well as their readiness to hedge catastrophic losses associated with sudden adverse movement of the foreign exchange level.[3] Since the major portion of the demand and supply curves contributing to the limit order book imbalance is usually unobserved by market participants (even though dealers might "feel" the relative strength of the market from private commications and from the order activity of their customers), the relative strength represented by the book imbalance can be an important piece of information about the future evolution of exchange rates.

Despite the heavy fragmentation and high heterogeneity of the spot foreign exchange order flow and its nontrivial role in the price discovery that should be apparent from the previous discussion, the markets in major currencies continuously provide very high liquidity as they are open around the clock and generally characterized by low transaction costs and low margin requirements. But even though the average daily foreign exchange turnover exceeds several times the combined trading volume in major stock markets,[4] the empirical research on foreign exchange market microstructure is growing relatively slowly. In particular, most empirical applications of high-frequency foreign exchange data dealt until recently with "indicative" quotes, primarily those collected by Olsen and Associates from the Reuters EFX informational pages. While these quotes offer a fair representation of the overall trading activity and remain an excellent source of high-frequency data for many econometric applications, they do not contain volumes and do not present binding commitments to trade from

[3]Osler [111] analyzes the role played by stop-loss orders in propagation of price cascades and provides empirical evidence of their importance for the fat tails in foreign exchange returns.

[4]According to the world Central Banks combined statistics, the total foreign exchange market was estimated at 1.4 trillion US dollars a day in 1995.

the institutions submitting the quotes. Therefore they are likely to provide only indirect and suggestive evidence on agents' behavior at the micro-level and hardly can be used to obtain reliable empirical results in foreign exchange market microstructure studies.

On the other hand, the electronic trading systems like Reuters D2000 and EBS remain the primary source of the data on binding quotes and real intradaily transaction activity in the spot foreign exchange market, no matter how limited the segments of the market covered by these data used to be. Among the pioneers of this research we mention Goodhart *et al.* [54], who processed and extensively analyzed a short snapshot of prices and quantities that appeared on a seven-hour videotape of the Reuters D2000-2 trading screen on one day in June 1993, at the early stages in the development of this brokerage system. In a subsequent paper, Goodhart and Payne [53] used the same sample to examine the determinants of quote revisions and spreads as well as some microstructural hypotheses in the foreign exchange market dynamics. The second, and perhaps most interesting data until recently came from the samples of the time-stamped quotes, deals, and positions for single representative marketmakers, or the time-stamped prices and quantities for transactions mediated by one or several major brokers. Such data sets covering a whole week in August 1992 were first used by Lyons [93], [95], to test inventory and asymmetric information hypotheses in the foreign exchange market.

Finally, various aspects of foreign exchange trading are analyzed in the studies by Lequeux [85], Daníelsson and Payne [31], [32], and Acar [1], primarily on the descriptive level. Even though some of these authors based their studies on larger and more representative samples of quotes and transactions and gave very interesting accounts of trading activity on both EBS and Reuters, for a variety of reasons they refrained from conducting their analyses on a tick-by-tick level, which may be crucial for understanding microstructure effects. This dissertation contains the detailed analysis of limit and market order activity in Reuters D2000-2 for the Deutsche Mark/US dollar exchange rate, which was the main bilateral

exchange rate in 1997, when the data was collected. This data set provides a rare opportunity to expand our knowledge about the role of order flow and other microstructure variables in the price formation on foreign exchange markets, and better understand how these markets work.

1.2 The Challenge of Modeling Market Liquidity

Historically, the large body of research studying frequency of quote revisions and transactions was inspired and motivated by the central question of modern finance: How long does it take for information to be fully incorporated into prices? Since private information in modern financial markets, according to proponents of the microstructure approach, is ultimately transmitted via continuous interaction of quotes and transactions, the frequency and complexity of arrival patterns for these bits and pieces of information should be one of the key ingredients to any transition mechanism toward the efficient price. Even though the information processing lags, random delays, and occasional congestions in communication networks can provide a partial explanation for the high level of noise and irregularity of intertrade and interquote durations, their persistence and high correlation with economically relevant variables such as the bid-ask spread and price volatility warrant the closer look at the duration processes. At least, it would be fair to say that discounting intertrade durations as pure noise cannot be justified in many microstructure applications, whereas the attempts to fully explain the duration dynamics by purely technical factors are likely to tell only part of the story.

It was long recognized in theoretical microstructure literature (Diamond and Verrecchia [34], Admati and Pfleiderer [3], [4], Easley and O'Hara [37], O'Hara [109], Chapter 6) that some of non-trading and quote delays can be purposeful and informationally motivated. For example, a Bayesian market maker in Easley and O'Hara [37] infers about the presence

of informed traders on the market from the intertrade durations, with shorter durations signalling the informed trader activity. However, in statistical applications, a purely statistical model for duration process is almost always augmented by some proxy characterizing the amount of private information on the market. The choice of such proxy is justified either theoretically by market microstructure models of price determination or empirically by availability of appropriate data at high frequencies.

The choice of the trading mode in early theoretical models was usually restricted to market orders that could be filled at the prices provided continuously by the market maker, or, alternatively, to limit orders for a fixed quantity that remained valid for a single time period and expired automatically if the order was unfilled. This obvious simplification was a price to pay for analytical tractability and crisp implications of the models. For example, in static models of the automatic limit order book by Glosten [50], and Chakravarty and Holden [20], the focus was on the optimal bidding strategies of limit order traders who were unwilling, or unable, to use market orders. A more recent strand of dynamic trade execution models (Parlour [112], Foucault [45]) emphasize the importance of the risk of non-execution and the risk of being picked off by informed traders for the order placement strategies employed by uninformed market participants. The first analytical results on the non-trivial interplay between the limit order price and the time-to-execution in complex dynamic environments began to appear only recently. One such example is the paper by Foucault, Kadan and Kandel [46], where the interactions between the trading decisions of patient and impatient traders play central role in the determination of dynamic equilibrium quotes and the bid-ask spread.

It should be clear from the above discussion that the heterogeneity of investors emerges as one of the driving forces behind the nontrivial properties of duration dynamics. At the same time it must be emphasized that the timing of transaction needs is not always synchronized across traders, even in the absence of traders with superior information. Traders

always have a choice between the submission of a market order that will be filled immediately at the best price available at the moment of submission, and the submission of a limit order that can improve the execution price for the submitting counterparty at the cost of delay and uncertainty of execution. As a result, the variations in traders' demand for immediacy of execution affects simultaneously the bid-ask spread and the depth of the market, driving the dynamics of the limit order book and the market liquidity. Investigation of the non-trivial dynamic relationship between three dimensions of liquidity, such as *immediacy* (the ability to trade a given quantity at a given cost), *breadth* (the cost of doing a trade of a given size quickly represented by the bid-ask spread), and *depth* (the size of trade that can be dealt quickly at a given cost) then naturally becomes the central object of research.[5]

The complete theoretical analysis of market liquidity in a realistic general equilibrium framework remains the major challenge for the market microstructure literature. The problem of solving and analyzing the multiple equilibria can be extremely complicated since traders' choices are not restricted to one parameter such as the price or quantity, but also include the decisions between the limit and market orders, sell and buy orders, as well as about the timing of order execution. Moreover, in real markets the traders can cancel and resubmit strategically their orders at any moment of time. Since full analysis of such a dynamic game remains excessively complex and just impractical to implement, the theoretical literature so far has been focusing on one or two dimensions of traders' decision, holding all other variables fixed or making other simplifying assumptions about traders' behavior.

In an attempt to better understand the decisions made by traders in the real world, many researchers concentrated their efforts on the search for stylized facts and empirical regularities with regard to different aspects of supply and demand for liquidity in the real markets. This effort was facilitated by broader availability of transaction and quote data

[5]See the discussion of different aspects of liquidity in Chapter 19 of Harris [64].

at intradaily frequencies. Since this search was originally restricted to the markets that make such data readily available, most of empirical papers in this area focused on the stock markets across the globe where systematic collection of data on quote and transaction activity was part of the institutional design enforced by financial regulators.[6] Sometimes the data were also collected by private enterprises and provided to researchers on proprietary basis. Examples of stock market analyses include Biais, Hilllion and Spatt [12], Handa and Schwartz [61], Harris and Hasbrouck [65], Lo, MacKinlay and Zhang [89], Griffith *et al.* [58], and Hollifield, Miller and Sandås [74]. For example, Biais, Hillion and Spatt [12] found that market orders in the Paris Bourse consume the major portion of liquidity available on the opposite side of the bid-ask spread, which then reverts to its original level as the limit order traders place new orders within the best bid and ask quotes. Even though most of transactions in the Paris Bourse occur at the small values of spread, the authors observed high-frequency negative autocorrelation between the quotes, as the spread showed a tendency to alternate between small and large values.

All of these studies either avoid the analysis of cancellation events altogether, or make very simplistic assumptions about traders' cancellation policy for the existing limit orders. The results of empirical studies taking this problem more seriously suggest that going beyond the trivial assumption might be a challenge. For example, no systematic studies presently available explain rigorously the empirical finding of Hasbrouck and Saar [69] who find that the majority of limit orders submitted on Island ECN that are cancelled, get cancelled within the first couple of seconds after submission. Needless to say, incorporating nontrivial order cancellation strategies, even though crucial for understanding the dynamics of liquidity, is

[6]However it must be emphasized that empirical research of liquidity patterns reveals a surprising degree of similarity across alternative market instruments, trading organizations, and locations. In particular, we believe that implications of our research will be mostly valid for a broad range of markets organized as electronic limit order books.

likely to make the theoretical analysis far more complex, if not infeasible. At present the most promising approach appears to be the accumulation of additional stylized facts about alternative aspects of liquidity in financial markets and interpretation of their properties and interaction within an appropriate econometric framework. The toolkit of available econometric methods is briefly discussed in the next section.

1.3 Econometric Approach to Modeling Market Liquidity in Continuous Time

The econometric approach provides a statistical framework for the rigorous empirical analysis of financial market data. Even though the tradition of using the theory of marked point processes to study heterogeneous events occurring randomly in time has had a long history in applied sciences, applications of univariate and bivariate marked point processes to the analysis of high-frequency data began to appear only relatively recently, first in a series of papers by Engle and Russell [41], [42], followed by Engle and Lunde [40], Gouriéroux *et al.* [55], Engle [39], and Russell *et al.* [119], among many others. In all these papers the irregularly spaced arrivals of transactions or quotes are modeled as a self-exciting point process with memory. The autoregressive conditional duration model (ACD, Engle and Russell [41], [42]) and its reincarnations have been the most popular specification of this point process so far. The markers usually represented by quotes or transaction prices and volumes (when available) are assumed weakly exogenous and modeled conditionally on their arrival times.

Development of numerous modifications of the ACD model largely parallels the history of the ARCH-GARCH literature in the early 1990s. Extensions of the ACD framework include the exponential ACD model by Bauwens and Giot [10], which precludes the occur-

rence of negative durations that cannot be ruled out by the original ACD specification, the fractionally integrated ACD (FIACD) model by Jasiak [77], which captures the long-range dependence in the empirical duration process, various versions of the threshold ACD model where the model specification hinges on two or more regimes identified from the data (for example, Russell, Tsay and Zhang [119]), and the stochastic duration model (Bauwens and Veredas [11]) applying the structure of stochastic model to log-durations. Other researchers (e.g., Ghysels and Jasiak [49], Engle [39], Grammig and Wellner [56]) combined the ACD-style analysis of durations with GARCH-type models for the financial returns.

One of the major impediments to application of this approach to multivariate financial data has been the controversy over the appropriate way to approach the markers of qualitatively distinct nature appearing randomly in time. To illustrate the idea, we use as an example the problem of extending the ACD framework to bivariate tick-by-tick data on Deutsche Mark/U.S. dollar and Japanese yen/U.S. dollar exchange rates. Prior to practical implementation of such an extension researcher should decide on the main object of his interest and make some exogeneity assumptions along the way. For the object of interest, a decision should be made whether the primary goal is to study durations between consecutive quotes for both currencies, only one of the two currencies, or perhaps only the intervals when Deutsche Mark ticks are followed by Japanese yen ticks. In principle, one may hope that this issue can always be resolved on the *ad hoc* basis, provided by institutional environment, or motivated by the goals of specific research projects. But even in this case, a systematic way of dealing with the markers of qualitatively distinct nature when they do not arrive in a natural order provides a clear benchmark for comparison, clarifying the issue of exogeneity, and increasing the discipline and rigor of future research. The similar problems and alternative approaches to their solution are also discussed by Engle and Lunde [40] in their model of trade and quote timing. Creative developments of the alternative approaches can be found in Spierdijk *et al.* [122], and in the application of the multivariate doubly stochastic Poisson

process theory by Davis *et al.* [33].

Unlike the rest of the literature on point processes in finance, which ignores the inquiries such as what could have happen if some particular marker did not arrive (or arrived several seconds later), the competing risks methodology put this type of retrospective thinking at the centerfold of the analysis. The competing risks model is developed on the premise that the sample of empirical durations may be dominated by a large number of latent durations leading to a limited observability of cause-specific durations, most of which happen to be censored by other cause-specific durations. Such an approach may be more relevant in the analysis of liquidity than traditional methods, as the self-censoring mechanism implied by the competition of traders for time priority is directly implied by the market design. The identification of competing risks is usually achieved with the assumption of conditional independence of the risks, given a sufficiently broad set of covariates.

A close counterpart to the competing risks model developed in this dissertation appears in Bisière and Kamionka [13] that applies a fully parametric competing risks model to the analysis of dynamics and sequencing of orders to trade the Alcatel shares at the Paris Bourse. The competing risks model by Bisière and Kamionka provides the joint explanation of the durations between consecutive order arrivals and their aggressiveness, emphasizing the role of information about the limit order book in the price discovery process. Unlike Bisière and Kamionka [13], in the present work we treat the hazard functions of competing risks semiparametrically, allowing freedom in the specification of baseline hazards but reducing the computational requirements at the expense of a relatively rigid structure of the covariate index. The competing risks approach thus views the limit order flow as a sequence of independent or quasi-independent realizations of a multivariate marked point process (Snyder and Miller, [121]), the number of its components being equal to the number of notional risks that must be identified in advance. Depending on the main focus of research and institutional details of the market, the history of a limit order book in the competing risks environment

can also be viewed as a mixture of birth-and-death processes (McLachlan and Peel, [100]) or as a nonstationary Markov random field in a discrete or continuous marker space.

From a broader perspective, the competing risks model with Cox specification of the hazard rates of notional risks developed in this dissertation can be thought of as a continuous time analogue of the multinomial logit model reformulated for the instantaneous conditional probabilities of individual risks (Greene, [57], Chapter 21). The baseline hazard rates of the notional competing risks are proportional to exponentials of the intercepts in the multinomial logit risk indices, whereas the conditional probability of survival over infinitesimally small continuous time intervals are proportional to the instantaneous risk of the base category in the multinomial logit model. Of course, the standard procedure of choosing some fixed time interval followed by aggregating the information within each interval, that must be performed prior to application of multinomial logit to the data, inevitably results in a loss of some information in the aggregates. While this loss may be dismissed as insignificant for some econometric applications, it appears to be critical for the studies of price formation and sequencing of orders when literally every quote does count and the order of their appeared can make a big difference for estimation and interpretation of results.

The basic assumption behind the specification of competing risks is that the state of a subject of study (in our case, it is predominantly the publicly observable part of the electronic limit order book) can be changed by a finite number of causes (sometimes called *notional risks*), which can be potentially of entirely different nature and usually are assumed independent. At any moment of time the relative importance of notional risks is determined by the odds of their instantaneous realization which can be characterized by the risk-specific hazard functions. After a single risk is realized and the limit order book is updated, the remaining risks that were competing with each other under the prior market conditions become irrelevant in the new state. Then the "internal clock" of the "competition" is initialized, and a new "race" begins immediately among the participating notional risks under the new market

conditions, proceeding up to the moment when the next "winner" triggers another change in the limit order book, and so forth. A number of alternative ways to initialize and stop the "internal clock", or, more generally, set pace for multiple "internal clocks" associated with different limit orders, allow for a substantial richness of the competing risks methodology and its general flexibility as a modelling device.

1.4 Overview of the Dissertation

This dissertation applies the competing risks model to the analysis of the order flow and price formation in the Reuters D2000-2 electronic brokerage system.

Chapter 2 introduces the competing risks methodology as an empirical tool for modeling high-frequency financial data in continuous time. After a brief review of key concepts and ideas of survival analysis, the competing risks model is applied to the analysis of the timing and interaction between the Deutsche Mark/U.S. dollar quotes and transactions in the Reuters D2000-2 electronic brokerage system. Estimation of this model generally supports empirical evidence from previous research on electronic limit order books. In particular, the composition of the order flow is found to be very sensitive to the state of the limit order book and the trading history. The direction of past trade has a strong predictive power for the future activity of buyers and sellers in the market. There is some evidence of an adverse information effect due to non-trading that manifests itself in the negative dependence of aggressive order arrival rates on time since the last observed event. It is found that traders tend to submit and cancel their orders most aggressively immediately after changes in the electronic order book.

Chapter 3 studies the problem of semiparametric hazard rate estimation of competing risks, with special attention paid to the case when the sample of observed durations is highly skewed. In this situation, which is fairly common for high-frequency financial data, the stan-

dard kernel estimators often fail to detect fine spikes of the hazard functions on the left and lead to extremely volatile hazard estimates in the right tail of the duration range. The chapter provides a review large sample properties of alternative k-nearest neighbor estimators and local linear smoothers. The asymptotic theory of k-nearest neighbor estimators, which has been developed in the covariate-free case, is also discussed in the context of semiparametric competing risk estimation problems and applied to the problem of baseline hazard rate estimation for a diverse range of limit order book events.

Chapter 4 extends the modeling techniques of the previous chapters in several directions. First, the set of covariates is expanded to include a broad range of limit order book activity characteristics in addition to those traditionally used in the empirical market microstructure literature to quantify the liquidity. Second, the cross-sectional and serial correlation of the residuals in the Cox regression is captured by the lagged activity measures such as the past order flow and the cumulative counts of recent transactions in the hazard function specifications for individual risks. Application of principal component analysis (PCA) to the covariate indices of the competing risks identifies five pervasive factors that capture 85% of the limit order book trading activity. The modified competing risks model incorporating the PCA factors leads to substantial data compression and improves the predictive performance of the model. By most accounts, the short-term forecasting power of the multifactor competing risks model is good relative to popular simple moving average-type forecasting rules. Even though directional signals generated by the model do not allow easy identification of systematic profit opportunities, the competing risks methodology is found to be a valuable tool for short-term forecasting of market activity and for understanding the behavior of heterogeneous agents in a competitive market environment.

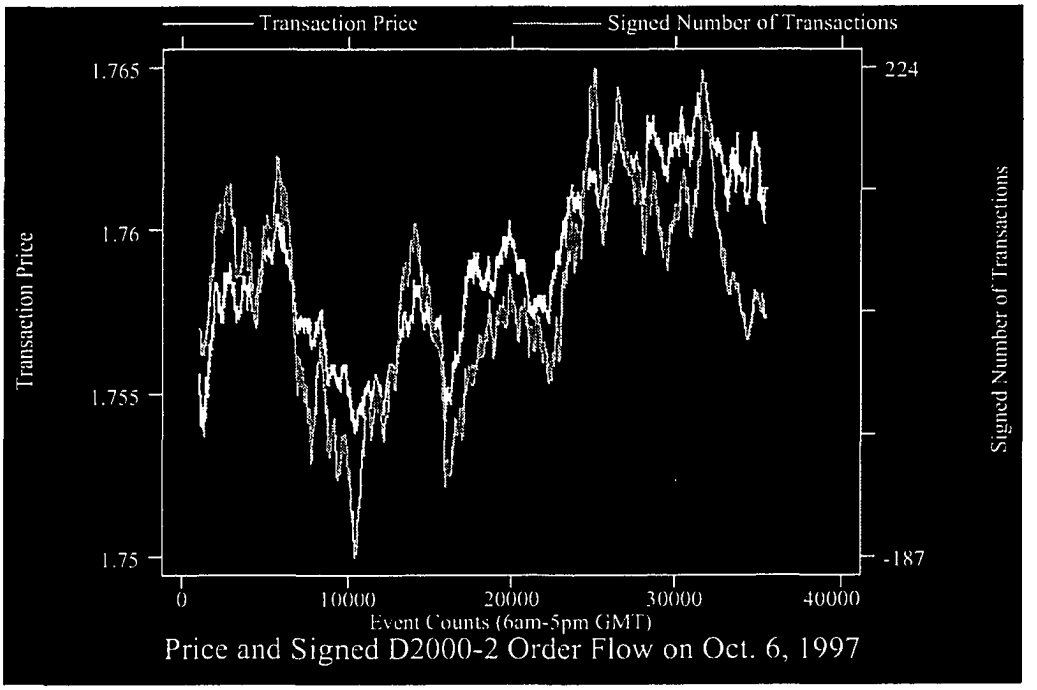


Figure 1.1: Evolution of transaction price and the cumulative signed order flow in Reuters D2000-2

Figure 1.1 displays the evolution of transaction price (exchange rate) and the signed cumulative flow of transactions in the Reuters D2000-2 trading system on Monday, October 6, 1997. The time series plots are shown in the activity scale, where one unit of time corresponds to one tick (submission of market or limit order, or cancellation of a limit order) in the system.

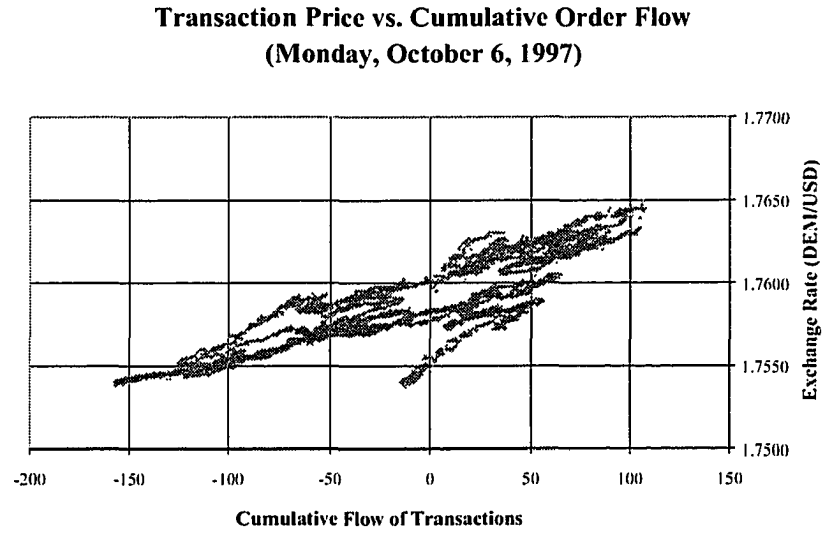


Figure 1.2: Phase diagram of transaction price vs. cum trade flow on 10/06/1997

This phase diagram on Figure 1.2. along with the similar diagrams on Figures 1.3–1.6 below, shows the scatter plots of transaction price (exchange rate) versus the signed cumulative flow of transactions in the Reuters D2000-2 trading system on five consecutive days, October 6–10, 1997. The cumulative flow of transactions is defined as the sum of signed indicators of transactions, where trades initiated by buyers of US dollars are counted as $+1$ and trades initiated by sellers of US dollars are counted as -1. All five diagrams show a strong positive relationship between the order flow and exchange rates, even though it never appears to be strict, or precise.

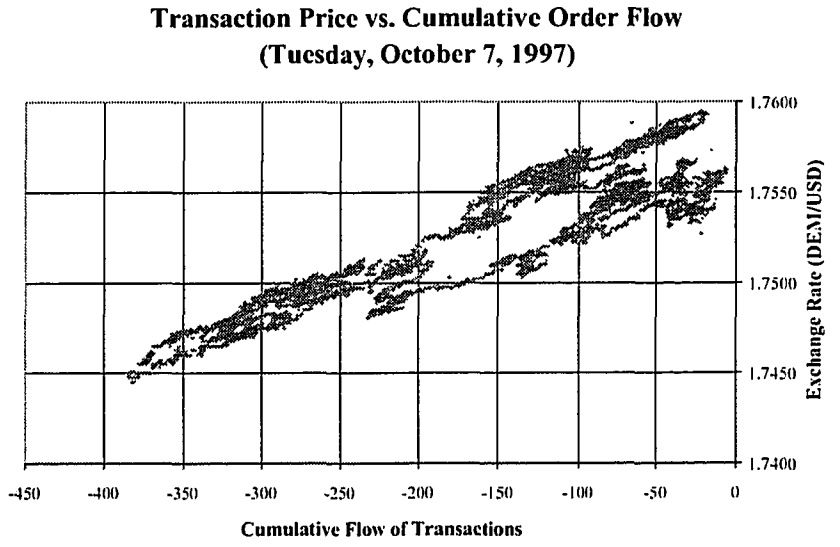


Figure 1.3: Phase diagram of transaction price vs. cum trade flow on 10/07/1997

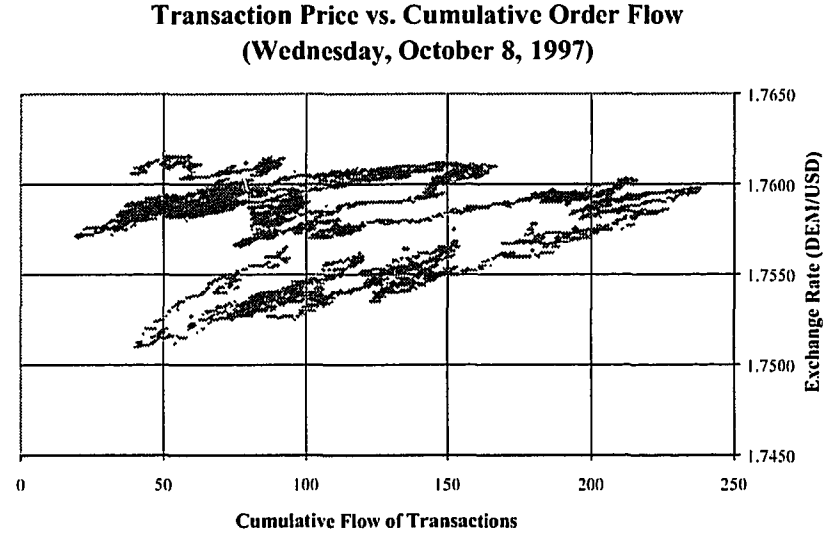


Figure 1.4: Phase diagram of transaction price vs. cum trade flow on 10/08/1997

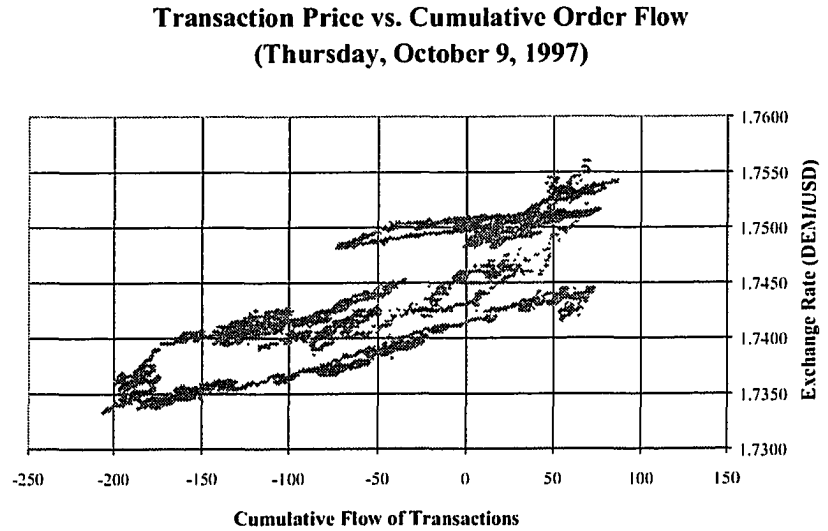


Figure 1.5: Phase diagram of transaction price vs. cum trade flow on 10/09/1997

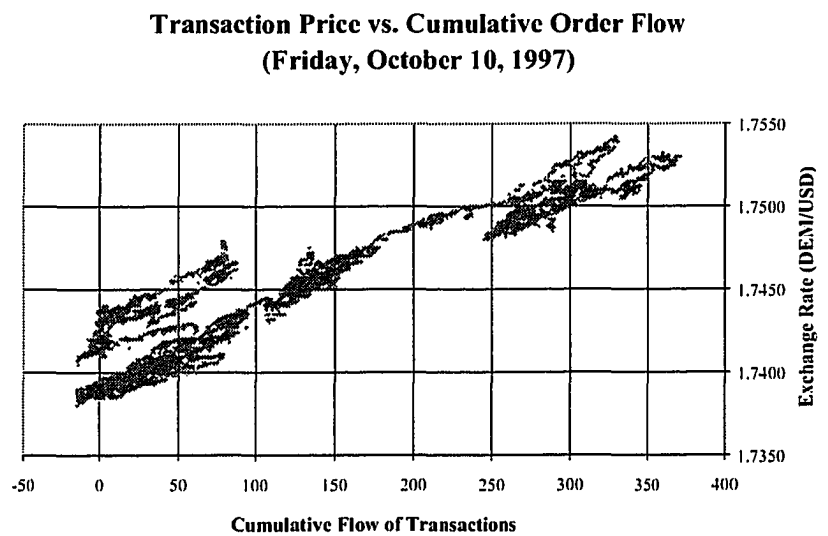


Figure 1.6: Phase diagram of transaction price vs. cum trade flow on 10/10/1997

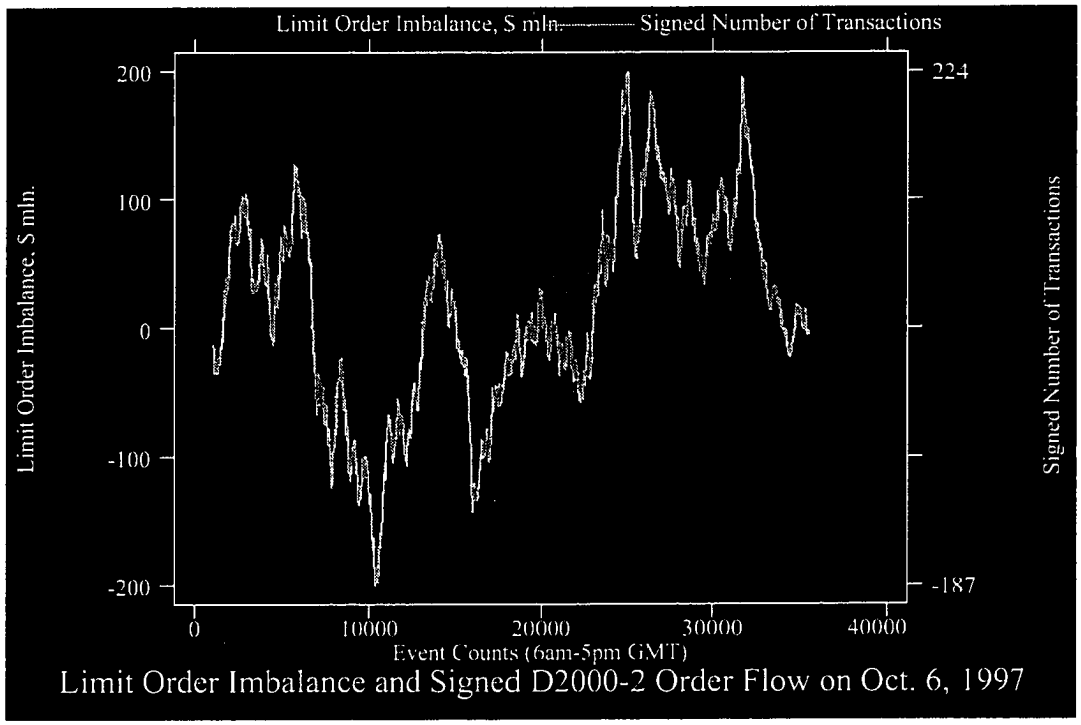


Figure 1.7: Evolution of the limit order book imbalance and the cumulative order flow in D2000-2

Figure 1.7 displays the evolution of the limit order book imbalance (measured as the difference between the total dollar value available for sale and the total dollar value available for purchase) and the signed cumulative flow of transactions in the Reuters D2000-2 trading system on Monday, October 6, 1997. The time series plots are shown in the activity scale, where one unit of time corresponds to one tick (submission of market or limit order, or cancellation of a limit order) in the system.

Chapter 2

Competing Risks and the Order Flow Dynamics in FOREX Electronic Brokerage

This chapter approaches the microstructural dynamics in an order-driven electronic financial market within the statistical framework of competing risks. The competing risks are identified as the events defined by arrivals and cancellations of alternative discrete types of market and limit orders. Then the hazard rates of these events are modeled in continuous time. The hazard rates are allowed to depend semiparametrically on the time since the last observable market event and on a linear index of covariates characterizing the past history and current market conditions. The model can also incorporate various patterns of unobserved heterogeneity due to the time-varying market conditions.

The competing risks methodology is applied to analyze the timing and interaction between Deutsche Mark/U.S. dollar quotes and trades in the Reuters D2000-2 electronic brokerage system. The market microstructure effects observed in this segment of the foreign exchange market are compared to empirical characteristics of similarly designed electronic limit order markets for stocks.

Estimation of this model generally supports the empirical evidence from previous studies of financial markets organized as electronic limit order books. In particular, the composition of the order flow is found to be very sensitive to the state of the limit order book and the trading history. There is some evidence of an adverse information effect due to non-trading that manifests itself in the negative dependence of hazards of aggressive order arrival on the time since the last observed event. Traders tend to submit and cancel their orders more aggressively immediately after changes in the electronic order book. Finally, the direction of past trade has strong predictive power for future activity of buyers and sellers in the market.

This chapter organized as follows. Section 2.1 describes in detail the Reuters D2000-2 dealing system focusing on the public and private information on the order flow available to the subscribers from trading screens. It also describes the original data source and the method of reconstructing the trading history from the data. Section 2.2 gives an overview of the basic concepts of survival analysis and illustrates them in the context of the Cox proportional hazard model. Section 2.3 introduces the Markov semiparametric model of competing risks which is the central model of this chapter. It describes the maximum likelihood estimators of the covariate effects, specifies the kernel estimators of baseline hazards, and briefly discusses their properties. Section 2.4 presents and interprets the empirical results and draws some links to the existing market microstructure literature. Technical details on the counting process interpretation of competing risks and asymptotic properties of the estimators will be reviewed and analyzed in Chapter 3 of this dissertation. Section 2.5 concludes and briefly discusses extensions. One of such extensions will appear in Chapter 4 of this thesis.

2.1 The Reuters D2000-2 Electronic Dealing System

The Reuters D2000-2 dealing system, which is the source of our data, is one of the two largest providers of electronic brokerage services on the spot foreign exchange market in

major currencies.[1] It operates as an anonymous matching service for trading in major currencies.[2] The D2000-2 dealing system forms the queues of buy and sell limit orders on the precise criteria of price and then time priority. The system incorporates a pre-screened pricing facility to ensure that two anonymous counterparties in a potential trade have mutual credit limits before a match can happen. Once a match between the two orders occurs, the parties are notified on the resulting transaction and provided with the information needed to settle the deal.[3]

The Reuters clients for D2000-2 are the dealers and traders in foreign exchange trading rooms of major financial institutions around the globe. As D2000-2 was designed to offer clear benefits to its subscribers over traditional brokerage or contacting a counterparty over the phone, it became quite popular among clients and had a network of active counterparties as large as 1,124 sites globally in the middle of 1997.

The D2000-2 trading screen has the following sections (Figure 2.1):

1. The multiple currency display allows to trade up to six currency pairs simultaneously and highlight the active pair.

2. The market quote area shows the anonymous best bid and offer prices of limit orders entered into the system.

3. The market depth indicator shows the accumulated quantities available at the market bid and offer.

[1]For the description of the EBS, which is the second major electronic brokerage system, see Lequeux [85]. Both systems, Reuters D2000-2 and EBS, have a similar trading mechanism and architecture, but the focus will be on the D2000-2 service which is the source of the data analyzed in the present paper.

[2]Clients must subscribe to the voice-based dealing system of D2000-1 to access D2000-2. Users can move from one service to the other on the same keystation. The D2000-2 service has two components of its own, D2000-2 Spot and D2000-2 Forwards, the latter designed to facilitate foreign exchange forward trading.

[3]Only the counterparty initiating the deal is charged for the transaction. The transaction fee as small as $25 is paid by the aggressor to Reuters Transaction Services Ltd.

4. The best quote area shows the best bid and offer prices available to this trader based on mutual credit.

5. The best depth indicates the quantities available at the best pre-screened prices.

6. The trader quote and trader depth section indicates the best prices and quantities of active limit orders entered onto D2000-2 at this keystation.

7. The last price indicator shows the last transaction price and direction of trade that occured on D2000-2 for each currency pair.

8. The trader mailbox section contains immediate confirmation messages of the trader's quoting activity and the deal tickets showing all details of her trades.

The trading screen can be customized to highlight various aspects of the subscriber's activity in single or multiple currencies.

2.1.1 Original Data

The data set made available by Reuters and provided by the Financial Markets Group at LSE covers the trading days from October 6 to October 10, 1997 and also contains a few orders originating late on October 5. As mentioned in the introduction, the only similar data previously available to academics is a short compilation of quotes from a seven-hour videotape of the D2000-2 screen dated by June 16, 1993 (Goodhart *et al.* [54]). The data contain information about 130535 limit and market orders made on the bid or ask side of the market. Each line of the file represents one limit or market order and contains exact entry and exit times, price and quantity ordered, quantity dealt, and information on whether the order came on the bid or offer side of the market. In addition there are a few other entries on each line, most of them redundant, which are used to validate the information on the reported order characteristics.

The data set does not include the confidential information on the identity of dealers submitting orders and completing transactions. Another highly relevant piece of information which is observed by individual subscribers on the D2000-2 screens but cannot be inferred from the data set are the best bid and ask quotes and the quantities of foreign exchange available at these prices to any of individual traders. At every point in time these best quotes and quantities are based on the existence of mutual credit lines between any given subscriber and her potential trade partners submitting limit orders at this time. Although the quotes and quantities available to individual subscribers would coincide with the best quotes and quantities of the market much of the time, on average the effective bid-ask spreads encountered by individual traders are slightly larger than the market bid-ask spread. For the same reason, the quantities available for trade to individual subscribers might be different from the quantities on the market at any given point in time.

Since the Reuters D2000-2 data set reports the exact entry and exit times for every limit and market order submitted into the system, it was expected initially to provide sufficient information for reconstruction of the market demand and supply curves at any point in time. These supply and demand curves could be used to keep track of the evolution of the best market bid and ask quotes and the quantities of foreign exchange available at those prices and to reconstruct all transactions in real time. Then it could be studied how the quotes and transactions move the equilibrium market price and how the arrival and cancellation events for limit and market orders interact with each other and affect price volatility, transaction volumes, market spread, and elasticities of market demand and supply. However the complete reconstruction of the history turned out to be infeasible without making occasional *ad hoc* assumptions that compensate for the lack of knowledge of traders' identity. Details of the implementation of the procedure are described in the next subsection.

2.1.2 Reconstruction of Trading History from the Data

Preliminary analysis of the data shows that the discrepancies between the best market quotes and quantities and those available for trade to individual customers do happen regularly in the system and impede the precise recovery of trading history from the data. As there was no perfect way to get around this problem, a computational model of the limit order market very similar to the one described by Daníelsson and Payne [32] and Hillman and Salmon [72] was created to build and maintain the entire limit order book over the trading week covered by the data. The object was to obtain the exact timestamps and characteristics of trades that can be either the matches between market and limit orders or the crosses between limit orders submitted on the opposite sides of the limit order book. The entry and exit timestamps for arrivals and cancellations of the limit orders not leading to immediate transactions were used to validate the matching procedure.

The limit orders were sorted by their arrival times and entered into the modeled limit order book by the price and time priority rules determined by the trading protocol. Whenever a cross between the arriving order and top priority limit orders on the opposite side of the book occured, a transaction was attempted and validated against subsequent arrival and cancellation events. The validation of trading history against available data was conducted on several levels, including verification of exit timestamps and quantities for all expiring entries. In the event of contradiction between the data and the reconstructed version of trading history, the computational model created another version of trading history under alternative assumptions about matching in the system (the orders having lower priority were matched ahead of higher priority orders whenever possible, higher priority orders were allowed to stay in the book, leading to negative market bid-ask spreads for a number of periods, etc.)

In a few instances when the computational model failed to resolve apparent contradictions, a supplementary manual search for anomalous entries was conducted to correct the errors. With the exception of two cases when limit orders leading to the errors had to be removed from the file, the multistage validation procedure described above successfully

recovered a non-contradictory version of the entire trading history. The estimation results of this chapter are based on the data for bid-ask spread, market depth, and other variables implied by the recreated version of trading history.

2.2 Review of Survival Analysis

This section presents an overview of survival analysis in the context of limit order market applications. For many years the methods of survival analysis were successfully applied in biostatistics, biomedicine, engineering, and actuarial science, and relatively recently they found numerous applications in labor, development, and environmental economics, as well as empirical finance.[4] Survival analysis, which focuses on evaluation of empirical risks in various environments, provides natural tools to quantify and analyze the risks faced by traders in their decision making. Yet the empirical work on limit orders that would extensively use the methods of survival analysis is rare, and many theoretical and methodological issues remain wide open.[5]

To illustrate the statistical techniques that could be consistently applied in the analysis of limit order markets, the Cox proportional hazard model (Cox [26]) appears to be a natural starting point. The Cox proportional hazard (CPH) model has its origins in the biostatistics literature and provides a simple framework to introduce the basic concepts and definitions of survival analysis. Section 2.3 covers a flexible and parsimonious extension of the CPH model for interarrival times of limit and market orders. Conceptually the model of section 2.3 can be considered a broad generalization of the single risk case studied in this section. The model also incorporates competing risks of cancellation for the existing limit orders and thus exploits the unique information contained in the data.

[4] See the monographs by Lancaster [84] and Cox and Oakes [28], and the papers by Han and Hausman [60], Meyer [102], Lunde, Timmermann, and Blake [91], Ridder and Tunali [116], and many others.

[5] Lo, MacKinlay, and Zhang [89], Bisière and Kamionka [13], Hautsch [71] are just the few recent papers consistently applying methods of survival analysis to the empirical analysis of limit order markets.

2.2.1 Basic Definitions

Let T denote a random variable that represents the duration between two consecutive publicly observed events in the history of the limit order book. Let $f(t)$ and $F(t)$ denote the probability density function (pdf) and cumulative distribution function (cdf) of T. The *hazard function* of T, denoted by $h(t)$, is defined as

$$h(t) = \frac{f(t)}{S(t-)}.$$

where the *survivor function* $S(t) = 1 - F(t)$ is the probability that the duration of a period between two consecutive events will be at least t. The hazard rate is interpreted as the instantaneous failure rate of T at time t, conditional on the event that no failure has been observed through time t. The corresponding *cumulative hazard function* is defined as

$$H(t) = \int_0^t h(u)du = -\log(S(t-)).$$

Note that the survivor function can be explicitly recovered from the cumulative hazard function as the product integral of $-H$,

$$S(t-) = \prod_{u\in[0;t)} (1 - \Delta H^{\mathrm{d}}\{u\}) \cdot \exp(-H^{\mathrm{c}}(u)), \tag{2.1}$$

where $\Delta H^{\mathrm{d}}\{u\}$ is the jump of H at u, and $H^{\mathrm{c}}(u)$ is the continuous part of H.

2.2.2 Estimation of the Cox Proportional Hazard Model for a Single Risk

According to the Cox proportional hazard (CPH) model (Cox and Oakes [28]), the hazard function $h(t|\mathbf{z})$ satisfies the relationship

$$h(t|\mathbf{z}) = h_0(t)\exp(\mathbf{z}'\boldsymbol{\beta}), \tag{2.2}$$

where $\mathbf{z}$ is a fixed vector of explanatory variables (covariates), $\boldsymbol{\beta}$ is a parameter vector, and $h_0(t)$ is called the baseline hazard function. The functional form of $h_0(t)$ is usually estimated nonparametrically.

Consider estimating the CPH model based on right-censored data. Suppose that data are available on durations $t_1, ..., t_N$ indexed by $n = 1, ..., N$, $\mathbf{z}_n$ are the values of covariates corresponding to duration t_n, and $\delta_{(j)}$ are indicator variables equal to zero if $t_{(j)}$ is a censored duration and unity if it is not. Assume that $t_{(1)} < ... < t_{(j)} < ... < t_{(J)}$ are the observed ordered *uncensored* duration values for a given event.[6] Denote by $\mathcal{R}_{(j)}$ the *risk set* associated with duration value $t_{(j)}$, i.e. the set of all durations equal to or larger than $t_{(j)}$. Then the partial likelihood (Cox [27]) is given by

$$\mathcal{L}_{\text{part}}(\boldsymbol{\beta}) = \prod_{j=1}^{J} \frac{\exp(\mathbf{z}'_{(j)}\boldsymbol{\beta})}{\sum_{n \in \mathcal{R}_{(j)}} \exp(\mathbf{z}'_n \boldsymbol{\beta})}, \tag{2.3}$$

where the product is taken over the set of uncensored duration values indexed by j with $\delta_{(j)} = 1$, $j = 1, ..., J$.

The partial likelihood approach to estimation of the CPH model takes two steps. At the first stage, vector $\boldsymbol{\beta}$ is estimated by maximization of the partial likelihood. At the second stage, the estimate of $\boldsymbol{\beta}$ is used to construct a nonparametric Breslow estimator of the cumulative baseline hazard function $H_0(t)$, which is

$$\widehat{H}_0(t, \hat{\boldsymbol{\beta}}) = \sum_{j:\, t_{(j)} \le t} \left[\sum_{n \in \mathcal{R}_{(j)}} \exp(\mathbf{z}'_n \boldsymbol{\beta}) \right]^{-1}, \tag{2.4}$$

where the external sum is again over the uncensored durations j characterized by $\delta_{(j)} = 1$. The baseline survivor function

$$S_0(t) = \exp(-H_0(t)) \tag{2.5}$$

and the survivor function

$$\log(S(t|\mathbf{z})) = \log(S_0(t)) \exp(\mathbf{z}'\boldsymbol{\beta})$$

are estimated by plugging in the estimator (2.4).

The assumption of no ties is not always satisfied in practice, even in the case considered in this chapter, where the durations are measured virtually on a continuous time scale.

[6] For clarity of exposition, the treatment of tied observations is delayed to the end of this subsection.

Suppose there are m_j uncensored failure events occuring at time $t_{(j)}$. Then the partial likelihood (2.3) can be approximated as follows (Breslow [15])

$$\mathcal{L}_{\text{part}}(\boldsymbol{\beta}) = \prod_{j=1}^{J} \prod_{\substack{l=1 \\ t_l = t_{(j)}}}^{m_j} \left[\frac{\exp(\mathbf{z}_l'\boldsymbol{\beta})}{\sum_{n \in \mathcal{R}_{(j)}} \exp(\mathbf{z}_n'\boldsymbol{\beta})} \right],$$

while

$$\widehat{H}_0(t, \widehat{\boldsymbol{\beta}}) = \sum_{t_{(j)} \leq t} \left[\sum_{n \in \mathcal{R}_{(j)}} \exp(\mathbf{z}_n'\widehat{\boldsymbol{\beta}}) \right]^{-m_j}.$$

and expressions for the remaining estimators follow automatically.

2.3 Semiparametric Markov Model of Competing Risks

Following the general approach of Kalbfleisch and Prentice [82] and the more recent application by Bisière and Kamionka [13] to Paris Bourse trading, the duration between consecutive events (arrivals and cancellations of limit orders and arrivals of market orders) and the type of the next event can be jointly analyzed conditional on the history of the process. The R types of events (notional risks) shown in Tables 2.1 and 2.2 are identified as arrivals or cancellations of different types of orders (market or limit orders, bids or offers) and their aggressiveness. The state space includes $S \leq R$ discrete Markov states associated with a subset of events that can be identified from the public information available on the Reuters D2000-2 trading screens. The remaining $R - S$ types of events can also affect the future state of the market but may not be immediately observed on the trading screens.[7]

The additional public information closely monitored by market participants includes the set of covariates observed on the screens that may include most recent transaction prices, the trade direction, the best bid and ask quotes currently available on the market, and the foreign exchange demand and supply at these prices. This information is pooled into a vector

[7]No doubt, traders have access to much larger information sets than publicly displayed on the trading screens. The pre-screened prices and quantities based on mutual credit (section 2.1) constitute one of the major pieces of private information used by agents in their decision making.

of random covariates $\mathbf{x}$ which will be referred in the future as the *state of the limit order book*.

The competing risks model belongs to the wide class of finite-state Markov models with covariates. In the beginning of every period the market is in a transient state $\mathbf{z}$ characterized by the type of the last observed event $s \in S$ and the current state of the limit order book described by covariates $\mathbf{x}$. If $\mathbf{d}$ is the random vector of indicators of the S observable events, then the transient state can be described as $\mathbf{z} = (\mathbf{x}', \mathbf{d}')'$. There are R absorbing states; state $r = 1, ..., R$ corresponds to the type of the next event (change in the limit order book), not necessarily publicly observed. Conditional on the state of the vector $\mathbf{z}$, the durations associated with R risks are simultaneously and independently drawn. Only the smallest of these R durations is observed while all the other durations are right-censored. The competing risks are characterized by transition intensities $h_r(t|\mathbf{z})$, $r = 1, ..., R$, which are termed "cause specific", or "notional" hazard functions, and interpreted as the arrival rates of type r events given that the current state of the market is $\mathbf{z}$ and no observable event occured for t units of time.[8]

It is easy to see how the competing risks model can be immersed in the framework of multivariate counting processes (Andersen *et al.* [8]). Consider N independent random vectors of durations $(T_{n1}, ..., T_{nR})$, $n = 1, ..., N$, and the associated hazard functions $h_1(t|\mathbf{z}), ..., h_R(t|\mathbf{z})$. Assume that random variables $T_{n1}, ..., T_{nR}$ are conditionally independent given the past history and the current covariates $\mathbf{z}_n$. Then consider N realizations of multivariate single-jump counting processes

$$\mathbf{N}_n(t) = (N_{n1}(t), ..., N_{nR}(t)), \quad n = 1, ..., N,$$

with

$$N_{nr}(t) = \mathbf{1}\{T_{nr} = \min_{r'} T_{nr'} \text{ and } T_{nr} \leq t\}.$$

[8] An adjustment of the model to the more realistic situation when the hazards of almost contemporaneous events depend on the state of the market before the earliest of those events occurs is relatively straightforward and leads to similar results.

Every individual counting process N_{nr} satisfies the multiplicative intensity model

$$\lambda_{nr}(t|\mathbf{z}) = Y_{nr}h_r(t|\mathbf{z}), \quad r = 1, ...R; \; n = 1, ..., N.$$

where Y_{nr} is an observable indicator that contains information whether or not the market in period n is at risk of experiencing an event of type r.

The pdf of duration T conditional on the next event being of type r and the current covariate vector (partially determined by the previous event as explained above) being equal to $\mathbf{z}$, is

$$\begin{aligned} f_r(t|\mathbf{z}) &= h_r(t|\mathbf{z})S(t-|\mathbf{z}) = h_r(t|\mathbf{z}) \prod_{r'=1}^{R} S_{r'}(t-|\mathbf{z}) \\ &= h_r(t|\mathbf{z}) \cdot \exp\left[-\int_0^t \sum_{r'=1}^{R} h_{r'}(u|\mathbf{z})du\right]. \end{aligned} \tag{2.6}$$

On the other hand, by the definition of survivor function,

$$S(t-|\mathbf{z}) = \frac{f(t|\mathbf{z})}{h(t|\mathbf{z})}.$$

where $h(t|\mathbf{z})$ and $f(t|\mathbf{z})$ are the conditional hazard function and pdf of duration t, whatever is the type of event associated with it. Therefore formula (2.6) can be rewritten as follows

$$f_r(t|\mathbf{z}) = \frac{h_r(t|\mathbf{z})}{h(t|\mathbf{z})} f(t|\mathbf{z}) = \pi_r(t|\mathbf{z})f(t|\mathbf{z}),$$

where $\pi_r(t|\mathbf{z})$ is the probability that an event of type r occurs exactly t units of time since the last event was observed, conditional on covariates $\mathbf{z}$ and given the information that some event happens at time t at all. Thus, the probability density function of duration between the previous observable event in transient state $\mathbf{z}$ and the next event (not necessarily observable) is

$$f(t|\mathbf{z}) = \sum_{r=1}^{R} h_r(t|\mathbf{z}) \cdot \exp\left[-\int_0^t \sum_{r'=1}^{R} h_{r'}(u|\mathbf{z})du\right].$$

and the associated hazard rate is

$$h(t|\mathbf{z}) = \sum_{r=1}^{R} h_{r'}(t|\mathbf{z}).$$

Bisière and Kamionka [13] assumed a flexible parametric Singh–Maddala form for hazard functions $h_r(t|\mathbf{z})$ in their model of market and limit order arrivals (they did not consider limit order cancellations). The Singh–Maddala form is defined as

$$h_r(t|\mathbf{z};\boldsymbol{\theta}_r) = \exp(\mathbf{z}'\boldsymbol{\beta}_r) \cdot \frac{\gamma_r t^{\gamma_r - 1}}{1 + \alpha_r t^{\gamma_r}}, \tag{2.7}$$

with estimated parameters $\boldsymbol{\theta}_r = (\boldsymbol{\beta}_r', \gamma_r, \alpha_r)'$ and covariates $\mathbf{z}$. The model considered in this chapter is semiparametric in the sense that the parametric form of baseline hazard function $h_{0r}(t)$ is left unspecified. The baseline hazard function is treated instead as an arbitrary smooth function of duration, and the infinite dimensional "parameter" to be estimated is $\boldsymbol{\theta}_r = (\boldsymbol{\beta}_r', h_{0r}(\cdot))'$.

Assuming there are N distinct arrival and cancellation events during the trading day, the likelihood function is

$$\begin{aligned}
\mathcal{L}(\boldsymbol{\theta}) &= \mathcal{L}(\boldsymbol{\theta}_1, ..., \boldsymbol{\theta}_R) = \prod_{n=1}^{N} f_{r_n}(t_n|\mathbf{z}_n;\boldsymbol{\theta}) \\
&= \prod_{n=1}^{N} h_{r_n}(t_n|\mathbf{z}_n;\boldsymbol{\theta}_r) S(t_n - |\mathbf{z}_n;\boldsymbol{\theta}) \\
&= \prod_{r=1}^{R} \prod_{n=1}^{N} h_r(t_n|\mathbf{z}_n;\boldsymbol{\theta}_r)^{\delta_{nr}} S_r(t_n - |\mathbf{z}_n;\boldsymbol{\theta}_r) \\
&= \prod_{r=1}^{R} \mathcal{L}_r(\boldsymbol{\theta}_r),
\end{aligned}$$

where

$$\mathcal{L}_r(\boldsymbol{\theta}_r) = \prod_{n=1}^{N} h_r(t_n|\mathbf{z}_n;\boldsymbol{\theta}_r)^{\delta_{nr}} S_r(t_n - |\mathbf{z}_n;\boldsymbol{\theta}_r),$$

t_n is the duration measured from the last observable event s_n associated with the present covariate vector $\mathbf{z}_n = (\mathbf{x}_n', \mathbf{d}_n')'$, and

$$\delta_{nr} = \begin{cases} 1 & \text{if the } n\text{th event is of type } r, \\ 0 & \text{otherwise.} \end{cases}$$

The maximum likelihood estimates of parameters $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, ..., \boldsymbol{\theta}_R)$ are obtained by independent maximization of functions $\mathcal{L}_r(\boldsymbol{\theta}_r)$ with respect to $\boldsymbol{\theta}_r$.

2.3.1 Cumulative Baseline Hazard Functions

To validate the Singh–Maddala parametric form of duration dependence for the baseline hazard functions (Bisière and Kamionka [13]), this dependence can be estimated nonparametrically. The notional hazard rate of type r as a function of time since inception (i.e., the time measured since the last publicly observed event) is assumed to be of the CPH form

$$h_r(t|\mathbf{z};\boldsymbol{\theta}_r) = h_{0r}(t)\exp(\mathbf{z}'\boldsymbol{\beta}_r)$$

with the unknown parameter $\boldsymbol{\theta}_r = (\boldsymbol{\beta}_r', h_{0r}(\cdot))'$, and the dependence of notional hazards on time being determined by functions $h_{0r}(\cdot)$ which are permitted to vary arbitrarily over the R types of risk.

Let $t_{(1)r} < ... < t_{(N_r)r}$ denote the N_r distinct ordered durations of type r $(r = 1, ..., R)$, and let $\mathbf{z}_{(j)r}$ characterize the covariates for the observed duration $t_{(j)r}$. The partial likelihood function

$$\mathcal{L}_{\text{part}}(\boldsymbol{\beta}) = \mathcal{L}_{\text{part}}(\boldsymbol{\beta}_1, ..., \boldsymbol{\beta}_R) = \prod_{r=1}^{R}\prod_{j=1}^{N_r}\left(\frac{\exp(\mathbf{z}_{(j)r}'\boldsymbol{\beta}_r)}{\sum_{n\in\mathcal{R}(t_{(j)r})}\exp(\mathbf{z}_n'\boldsymbol{\beta}_r)}\right)$$

depends on a finite dimensional parameter $\boldsymbol{\beta}$, and the risk set $\mathcal{R}(t_{(j)r})$ is defined as the set of observed durations that are equal to or larger than $t_{(j)r}$ (c.f. the general definition of risk set from section 3.2). Insertion of maximum partial likelihood estimators $\widehat{\boldsymbol{\beta}}_1, ..., \widehat{\boldsymbol{\beta}}_R$ into the expressions for hazard functions yields the Breslow estimators of cumulative hazard functions

$$\widehat{H}_{0r}(u, \widehat{\boldsymbol{\beta}}_r) = \sum_{j:\, t_{(j)r}\le t}\left[\sum_{n\in\mathcal{R}(t_{(j)r})}\exp(\mathbf{z}_n'\widehat{\boldsymbol{\beta}}_r)\right]^{-m_{jr}},$$

where m_{jr} is the number of tied durations of type r at $t_{(j)r}$. The estimates of survivor functions $S_r(t|\mathbf{z};\widehat{\boldsymbol{\beta}}_r)$ are obtained exactly as shown above for the case of a single risk (section 3.2). The estimators of cumulative incidence

$$\mathbf{I}_r(t|\mathbf{z}) = \mathbf{P}(T_r \le t \text{ and } T_r = \min_{r'=1,...,R} T_{r'}|\mathbf{z}) \tag{2.8}$$

$$= \int_0^t h_r(u|\mathbf{z};\boldsymbol{\theta}_r) \prod_{r'=1}^{R} S_{r'}(u|\mathbf{z};\boldsymbol{\theta}_{r'})du$$
$$= \int_0^t h_r(u|\mathbf{z};\boldsymbol{\theta}_r) S(u|\mathbf{z};\boldsymbol{\theta})du$$

can be obtained by inserting into (2.8) the appropriate estimators for S and h_r.

2.3.2 The Baseline Hazard Functions

The baseline hazard functions $h_{0r}(t)$ for competing risks $r = 1, ..., R$ are estimated from the cumulative baseline hazard estimators $\widehat{H}_{0r}$ by smoothing their increments. A natural estimator of $h_{0r}(t)$ can be defined as

$$\widehat{h}_{0r}(t) = \frac{1}{b_r} \int \mathcal{K}_r \left(\frac{t-u}{b_r} \right) d\widehat{H}_{0r}(u, \widehat{\boldsymbol{\beta}}_r), \tag{2.9}$$

where $\mathcal{K}_r$ is a kernel function and b_r is a bandwidth parameter that both may depend on the type of risk. If $t_{(1)r} < ... < t_{(j)r} < ...$ denote the successive jump times of counting process $N_{\cdot r} = N_{1r} + ... + N_{Nr}$, then the equivalent definition of $\widehat{h}_{0r}(t)$ would be

$$\widehat{h}_{0r}(t) = \frac{1}{b_r} \sum_j \mathcal{K}_r \left(\frac{t-t_j}{b_r} \right) \Delta\widehat{H}_{0r}(t_j, \widehat{\boldsymbol{\beta}}_r), \tag{2.10}$$

that is, $\widehat{h}_{0r}(t)$ is a weighted mean of the increments of the Breslow estimator $\widehat{H}_{0r}$ (see Chapter 3 for details).

Andersen *et al.* ([8], section IV.2) show that the kernel function estimator (2.10) of $h_{0r}(t)$ is consistent provided the bandwidth b_r tends to zero sufficiently slowly as the total number of observations increases. In order to pick an optimal bandwidth, some measure of global performance of the kernel function estimator has to be selected. A popular criterion is the cross-validation function (CV) which is defined by

$$\mathrm{CV}(b_r) = \sum_j \left(\Delta\widehat{H}_{0r}(t_j, \widehat{\boldsymbol{\beta}}_r) - \widehat{h}_{0r}^{-j}(t_j) \right)^2, \tag{2.11}$$

where $\widehat{h}_{0r}^{-j}(t_j)$ is the leave-one-out estimator evaluated at the jump point t_j,

$$\widehat{h}_{0r}^{-j}(t_j) = \frac{1}{b_r} \sum_{i \neq j} \mathcal{K}_r \left(\frac{t_j - t_i}{b_r} \right) \Delta\widehat{H}_{0r}(t_i, \widehat{\boldsymbol{\beta}}_r).$$

Minimization of $CV(b_r)$ over a grid of bandwidths usually gives reasonable results and is asymptotically optimal in terms of a quadratic error measure under additional assumptions (Härdle [63], section 5.1).

Unfortunately, the distribution of the jump durations t_j is highly positively skewed, and further modifications of the method are needed. One practical way to proceed would be an automatic local bandwidth selection such as Friedman's "supersmoother" (Härdle [63], section 5.3) which essentially performs a "local cross-validation" procedure for k-NN estimators combined with consecutive spline smoothing. Another option to be considered is a modification of the approach of Wand *et al.* [125] based on the deformation of duration scale. The main idea of this alternative approach can be outlined as follows. First, the problem is reformulated in terms of log-transformed data. Then the standard (global) cross-validation criterion (2.11) is applied to the new data, and finally the inverse deformation of duration back to the original scale is performed. The log-transform of duration scale removes the excessive skewness in the sample distribution of jump points and makes the sample design more balanced and closer to normal. A family of transforms alternative to the log such as the Box–Cox transformation might be considered in this approach. The cross-validation can be performed jointly over the kernel bandwidth parameter and the parameter of the transformation family.

An alternative to the kernel estimator (2.9) is

$$h^*_{0r}(t) = \frac{\frac{1}{b_r}\sum_{j=1}^{J_r} \mathcal{K}_r\left(\frac{t-t_{(j)r}}{b_r}\right) m_{jr}}{\frac{1}{b_r}\sum_{j=1}^{J_r} \mathcal{K}_r\left(\frac{t-t_{(j)r}}{b_r}\right) \sum_{n\in\mathcal{R}(t_{(j)r})} \exp(\mathbf{z}_n'\widehat{\boldsymbol{\beta}}_r)\Delta t_{(j)r}}, \tag{2.12}$$

where $t_{(1)r} < \ldots < t_{(j)r} < \ldots < t_{(J_k)r}$ are the ordered jump times of process $N_{\cdot r}(\cdot)$, and $N_{\cdot r}(t)$ is the number of events of type r that occured at or prior to time t,

$$m_{jr} = \Delta N_{\cdot r}(t_{(j)r}) = N_{\cdot r}(t_{(j)r}) - N_{\cdot r}(t_{(j-1)r})$$

is the multiplicity of type r events at point $t_{(j)r}$, and $\Delta t_{(j)r} = t_{(j)r} - t_{(j-1)r}$ is the interval between consecutive jumps of process $N_{\cdot r}$. The motivation for (2.12) comes from a similar

estimator in a marker dependent hazard model investigated by Nielsen and Linton [108]. Yet another option is to take the kernel and the bandwidth parameter independent of the type of risk, and estimate the baseline hazards by

$$h_{0r}^{**}(t) = \frac{\frac{1}{b}\sum_{j=1}^{J} \mathcal{K}\left(\frac{t-t_{(j)}}{b}\right) m_{jr}}{\frac{1}{b}\sum_{j=1}^{J} \mathcal{K}\left(\frac{t-t_{(j)}}{b}\right) \sum_{n \in \mathcal{R}(t_{(j)})} \exp(\mathbf{z}_n' \widehat{\boldsymbol{\beta}}_r) \Delta t_{(j)}}, \tag{2.13}$$

which is analogous to (2.12) but the summation in the numerator and denominator is over the jumps $t_{(1)} < ... < t_{(j)} < ... < t_{(J)}$ that occur in the sample for all types of risk,

$$m_{jr} = \Delta N_{\cdot r}(t_{(j)}) = N_{\cdot r}(t_{(j)}) - N_{\cdot r}(t_{(j-1)})$$

is the multiplicity of type r events at point $t_{(j)}$, and $\Delta t_{(j)} = t_{(j)} - t_{(j-1)}$ is the interval between the consecutive jump points of vector process $\mathbf{N}_{\cdot}(t) = (N_{\cdot 1}(t), ..., N_{\cdot R}(t))$.

The asymptotic properties of estimators (2.12) and (2.13) will be investigated in the next chapter of this dissertation. A nearest neighbor modification of these estimators is applied in the next section to estimate the baseline hazard functions of competing risks.

2.4 Estimation Results

For a given type of event r, the components of vector $\mathbf{z}_n = (\mathbf{x}_n', \mathbf{d}_n')'$ $(n = 1, ..., N_r)$ are associated with the following characteristics of the limit order book and major trading events:

1. indicator of the direction of the most recent transaction (1 if seller initiated, -1 if buyer initiated);

2. market spread (the difference between the best ask and bid quotes), in ticks;

3. *ex post* profitability of the last trade, or positioning bias, defined as the distance between the price of the last transaction and the current bid-ask midquote, in ticks, for buyer-initiated events, and the negative of the same quantity, for seller-initiated events;

4. natural log of the depth at the offered touch (lowest ask price of the market), in $ mln.;

5. natural log of the depth at the bid touch (highest bid price of the market), in $ mln.;

6. vector of indicator variables $\mathbf{d}_n = (d_{n1}, ..., d_{nS})'$ for the present state of the market s ($d_{ns'} = 1$ for $s' = s$, $d_{ns'} = 0$ otherwise).

The components of $\mathbf{x}_n$ incorporate the information available on the screen which is closely monitored by traders. Although the Markov state variable $s \in S$ can be more difficult to follow on the trading screens, it can be usually derived from the movements of market quotes and quantities as explained in the next paragraph. The definitions of covariates $\mathbf{x}_n$ describing the state of the limit order book are given in Tables 2.3 and 2.4. The direction of prior trade indicator is used to identify the asymmetry in the impact of *completed transactions* on hazard rates as opposed to the asymmetry in the impact of aggressive quotes captured by indicator variables of vector $\mathbf{d}_n$, as the latter only indicate the *intention to trade.* There is a strong evidence that the buy-sell indicator has a high predictive power for the direction of future transactions on the foreign exchange market (Goodhart *et al.* [54], Acar [1]) and on the stock markets (Hausman *et al.* [70], Lo *et al.* [89], Huang and Stoll [75]). The size of market bid-ask spread is often identified with an intuitive notion of illiquidity in the market microstructure literature, and is expected to have a strong impact on the types of submitted orders. The positioning bias, or *ex post profitability* of the prior trade, can be interpreted as the profit accrued to the trader who was an aggressor in the last transaction (the counterparty initiating the trade) if she liquidates her position at the mid-point of the current bid-ask spread.[9] Finally, the two market depth variables represent another dimension of liquidity, viz. how many units of foreign exchange can be bought (or sold) at the current ask (or bid) market prices. The depth variables are also expected to be significant for the notional risks of cancellation events as the risk of cancellation should be related to the total

[9]This interpretation of positioning bias disregards the transaction cost, which is always incurred by aggressor according to the trading protocol of the D2000-2 trading system.

number of active limit orders, and the latter number is expected to be correlated with the quoted depth at the best market price.[10]

The taxonomy of limit orders in this chapter is very similar to the one introduced by Biais *et al.* [12], which appeared in subsequent papers on limit orders with minor modifications.[11] We specify $R = 22$ event types and select $S = 14$ state variables corresponding to publicly observable types of events. The definitions of events are given in Tables 2.1 and 2.2.

Stars and double stars next to the type of event indicate the publicly observable changes in the limit order book. Double stars indicate the events that can potentially trigger an immediate trade execution (note that transactions will not occur automatically following an aggressive order arrival because of potential credit constraints, as explained in section 2.1). Whenever an event associated with a change in quoted bid or ask market prices or quantities occurs, the identity of discrete Markov state $s = 1, ..., S$ changes.[12] However, the events associated with changes of the suboptimal quotes and quantities available at those prices are not included in the public information about the limit order book, even though those events could be observed by some traders. Those events do not change the state of the market and do not reset the "internal clock" of the "race" among the risks, according to the version of the model considered in this chapter. Thus all types of events except A6, A7, A10, A11, B6, B7, B10, and B11 correspond to distinct observable Markov states.

[10]The determinants of limit order submision strategies are briefly discussed in O'Hara [109] and Goodhart and O'Hara [52]. The main theoretical contributions on the topic are Chakravarty and Holden [20], and Parlour [112]. Biais et al. [12], Handa and Schwartz [61]. Harris and Hasbrouck [65], and Brown et al. [17] contain interesting empirical results.

[11]Griffiths et al. [58]. Bisière and Kamionka [13], and Al-Suhaibani and Kryzanowski [6] are just a few examples from the recent literature. Minor variations in the classification of orders depend primarily on the data availability as well as the main object of paper.

[12]Note that the state of the same typeis also assumed to have a distinct identity, so the new period can start after transition $s' \to s$.

2.4.1 The Fully Parametric Model

First, the fully parametric model (2.7) by Bisière and Kamionka [13] is estimated by the maximum likelihood procedure. The component of log-likelihood corresponding to the events of type r,

$$\log \mathcal{L}_r(\boldsymbol{\theta}_r) = \sum_{n=1}^{N} [\delta_{nr} \cdot \log(h_r(t_n|\mathbf{z}_n;\boldsymbol{\theta}_r)) + \log(S_r(t_n|\mathbf{z}_n;\boldsymbol{\theta}_r))] \tag{2.14}$$

$$\begin{aligned} = & \sum_{n=1}^{N} \delta_{nr}(\mathbf{z}_n'\boldsymbol{\beta}_r + \log\gamma_r + (\gamma_r - 1)\log t_n - \log(1+\alpha_r t_n^{\gamma_r})) \\ & - \sum_{n=1}^{N} \exp(\mathbf{z}_n'\boldsymbol{\beta}_r) \cdot \frac{1}{\alpha_r} \log(1+\alpha_r t_n^{\gamma_r}), \end{aligned}$$

is maximized with respect to $\boldsymbol{\theta}_r = (\boldsymbol{\beta}_r', \gamma_r, \alpha_r)'$ given the vectors of Markov state indicators $\mathbf{d}_n$ and covariates $\mathbf{x}_n$, the durations t_n from prior observed events leading to the present state s_n, and indicators δ_{nr} showing that the nth realized risk is of type r ($n = 1, ..., N$). The fully parametric maximum likelihood (2.14) estimates for the events recorded between 6 a.m. and 5 p.m. GMT on the week of October 6–10, 1997, are reported in Tables 2.5 and 2.6.[13] The estimates are accompanied by the t-statistics calculated from the "robust" variance-covariance matrix of coefficients.

Several patterns can be observed from this empirical exercise. First, parameters γ_r are always substantially larger than unity. The estimates of parameters α_r (not shown in Tables 2.5 and 2.6) are always positive and usually quite large, leading to overwhelming rejection of the Weibull distribution hypothesis $\alpha_r = 0$.

[13]The overnight trading hours (5 p.m. to 6 a.m. GMT) characterized by low market liquidity and high and unstable values of the bid-ask spread cover less than 5% of all events and are excluded from the analysis. Apparently, the standard competing risks methodology can be applied to the analysis of such an off-hours market with minor modifications. However, since the quotes and trades on the routinely illiquid Reuters D2000-2 system in the overnight hours **de facto** represent a separate trading regime, the estimates of competing risks covering this period are likely to be very different from those reported in the paper for the high-activity hours.

Second, the combination of the sign and *ex post* profitability of the prior transaction is always statistically significant, which indicates the important role of the information about the last trade reported on the D2000-2 trading screens. The influence of the *Side* variable is almost uniformly positive for seller-initiated activity and negative for buyer-initiated activity. The coefficients of the *Slippage* variable are large and positive for the risk of aggressive limit order arrivals (types A1, B1, A3, and B3, according to the classification of Tables 2.1 and 2.2), slightly smaller for market order arrivals (types A2 and B2), and much smaller but still significant for the arrivals of limit orders improving the market price within the spread (types A4 and B4). The *Slippage* coefficients become significantly negative for cautious limit order arrivals and all cancellation events. This offers strong support for clustering of quoting and trading activity on the ask and bid sides of the limit order book, and suggests the important role of the information transmitted by the last transaction event for the direction of this clustering.

Third, the *Spread* variable is always statistically significant and has the strongest influence on the hazard rates of most aggressive orders. Almost all activity in the limit order book is negatively related to the size of the market bid-ask spread. Quite naturally, the only exception is the odds of limit order arrivals improving the bid-ask spread without the risk of immediate transaction, i.e., events of types A4 and B4.

Finally, the dependence of hazard functions on the quoted depth variables is statistically significant for most types of risks. In some cases the significance of these coefficients can be partially explained by the *event classification biases.* Most obvious of those biases involve the definition of aggressive vs. small limit orders hitting on or taking out the liquidity available on the opposite side of the book (type A1 vs. A3 and B1 vs. B3) and the definition of full vs. partial cancellation of liquidity at the touch (type A8 vs. A9 and B8 vs. B9). One obvious pattern apparently unrelated to the event classification scheme of this chapter is the positive influence of market depth on the odds of price improvement on the same side of the limit order book. The mechanism behind this pattern is the competition for time priority among traders. Large quoted depth implies larger execution lags and forces the

submission of orders for immediate execution or orders for delayed execution at the prices which are more attractive to potential counterparties on the opposite side of the market. On the other hand, small quoted depth at the best market price gives stronger incentive to add liquidity at the same tick and get the price advantage if the limit order is hit (or taken) by an aggressor from the opposite side of the market. Another observed empirical regularity is some reluctance of traders to submit market orders (that is, the orders almost guaranteed to be traded, events of types A2 and B2) when the quoted depth on the opposite side of the market is high. A potential link between this effect and the information about the future adverse price movements that may be transmitted by the market depth quoted on the opposite side requires further investigation.

Note that the reported covariate coefficients of the competing risks should be treated with caution as the Singh–Maddala parametric form (2.7) might be a poor approximation for the hazard functions of notional risks. Some suggestive evidence of such a misspecification reported in the next subsection will be followed by a more careful analysis in subsection 2.4.3.

2.4.2 Covariates in the Semiparametric Model

Table 2.7 reports the estimates of the semiparametric CPH competing risks of seller-initiated events. Table 2.8 reports similar estimates for buyer-initiated events. Entries of the table contain the estimated coefficients at the covariates accompanied by their robust t-statistics (Lin and Wei [87]) for the covariates characterizing the state of the limit order book. Significant covariate coefficients are marked by stars.

Most qualitative results are similar to those obtained from the fully parametric model and can be related to empirical facts on limit order dynamics reported in the literature. In particular, the hazards of arrival for market and aggressive limit orders on both sides of the book are much higher if the previous transaction was initiated by an order of similar type (*Side* effect). Independently, these hazard rates increase even more if the price of the

last transaction relative to the current bid-ask midquote is high for buyer-initiated and low for seller-initiated transactions (*Slippage* effect). Similar, but less pronounced effects are observed for arrivals of quotes that improve the current market prices without immediate execution (events of types A4 and B4).

Large quantities available at the best market price substantially increase the risk that the next order will either improve this quote without transaction or trigger an immediate execution. This is a direct consequence of competition for the priority in the limit order book discussed in the previous subsection; the price improvement for a potential counterparty on the opposite side of the limit order book is most likely to occur when the competition among the limit order traders providing the liquidity is high. On the other hand, traders appear to act more cautiously if they consider the decision to submit a market order during the periods when substantial liquidity is available from the opposite side of the book.

The significance of *Spread* variable for all notional risks confirms the important role played by the market bid-ask spread for price discovery on the electronically brokered segment of the market. In fact, large bid-ask spreads have a strong negative effect for the general level of market activity. The only type of risk reacting positively to high bid-ask spreads is the hazard of price improvement on one or the other side of the market. Still, this increase is not as dramatic as it should have been to compensate for the reduction of the other notional risks. The liquidity is restored after a relatively long period characterized by relatively wide bid-ask spreads and slightly higher chances of arrival for price improving offers and bids.

The only substantial difference between the estimates of the two models is observed for the coefficients characterizing the market depth effect on the competing risks. Most depth coefficients in the fully parametric model are substantially lower than the corresponding coefficients in the semiparametric CPH model. In particular, almost all coefficients characterizing the influence of quantities bid or offered on the opposite side of the market (i.e., Q_{bid} for seller-initiated events and Q_{ask} for buyer-initiated events) are insignificant in the semiparametric model, although they are significantly negative in the fully parametric model.

The same side market depth effects in the two models are also drastically different for the odds of quantity improvements at the touch that occur without price improvement (events of types A5 and B5). While the fully parametric model reports a significant negative effect of the same side depth, which might be interpreted as the evidence of switching to more aggressive limit orders when the competition among limit order traders for time priority is high, the same coefficients in the semiparametric CPH model are positive and very significant. If traders interpret the large liquidity observed on the same side of the market as an encouraging sign of price stabilization at the given quote, they would be willing to supply more liquidity and wait longer to execute orders at the same price. Such a sharp difference between the estimates of depth coefficients in the two models appears puzzling, especially considering the remarkable similarity of all other coefficients. The evidence reported in the next subsection suggests one possible explanation for the systematic negative bias of the depth coefficients in the parametric model – misspecification of the baseline hazard functions for competing risks.

2.4.3 The Estimates of Baseline Hazard Functions

A version of the k-nearest neighbor (k-NN) estimator (2.13) is employed for the estimation of baseline hazards.[14] Although no asymptotic theory has been developed for k-NN estimators of hazard functions, the nearest neighbor design is preferred as it offers a simple solution to the problem of bandwidth selection for every duration.[15] The only parameter that has to be selected in conjunction with the weighting scheme is the number of neighbors k as the percentage of the total number of jumps J. The values of k/J in the range between 5% and

[14]The alternative hazard estimators (2.10) and (2.12) have been calculated but are not included in the present paper. The qualitative properties of estimates based on (2.10) and (2.12) are similar to those reported for estimator (2.13).

[15]The k-NN estimation was also the preferred approach in Engle [39] to the nonparametric hazard function estimation for interarrival times of stock quotes. There was no discussion of confidence intervals for the estimate of hazard function reported in that paper.

50% have been tried, but only results for the subjective choice of $k/J = 0.2$ are reported. To reduce the kernel truncation biases that arise near the boundaries of the duration range, the number of neighbors covered by the support of the kernel progressively decreases from $k \approx 0.2J$ to $k = 1$ as the duration value at which the hazard function is estimated moves from the 10% quantile to the minimum of the duration sample and from the 90% quantile to the maximum of the sample.

As for the weighting schemes, the standard choices are analogous to the case of constant bandwidth and include uniform, triangular, Epanechnikov, and quartic kernels. The uniform weights tend to produce extremely rough curves, even for reasonably large values of k/J. The Epanechnikov weights substantially reduce the degree of "wiggliness" and simplify integration, at the same time putting more weight on the duration values close to the duration of interest. The *quartic* kernel weights

$$W_k(j, j') = \begin{cases} \left(1 - \left(\frac{j-j'}{[k/2]+1/2}\right)^2\right)^2 & \text{for } j - j' = -[k/2], ..., [k/2], \\ 0 & \text{otherwise,} \end{cases} \tag{2.15}$$

used in this chapter produce the estimates that are visually indistinguishable from smooth curves over a wide range of durations excluding the "thin sample design" area in the right tail of the sample distribution. The normalizing constant in the weighting scheme is not important, since the k-NN analogue of estimator (2.13) is defined as a ratio of two weighted averages,

$$\sum_{j'=1}^{J} W_k(j, j') \Delta N_{\cdot r}(t_{(j')})$$

and

$$\sum_{j'=1}^{J} W_k(j, j') \sum_{n \in \mathcal{R}(t_{(j')})} \exp(\mathbf{z}_n' \widehat{\boldsymbol{\beta}}_r) \Delta t_{(j')}.$$

The estimated baseline hazard functions of publicly observable seller- and buyer-initiated events are plotted for quartic weights and the nearest neighbor "bandwidth" ratio $k/J = 0.2$, respectively, in Figures 2.2 and 2.3.[16] To facilitate the comparison of different

[16]The baseline hazard estimates for the events which are not publicly observable according to the classifi-

curves, all baseline hazard functions are normalized at the origin so that

$$\lim_{t \to 0} h^{**}_{0r}(t) = 1.$$

The legends of the plots show the types of risks in the order of their appearance on the graphs, from top to bottom.

The graphs demonstrate the considerable variation in the shapes of the baseline hazard functions across the types of notional risks. This variation is unlikely to be reduced to a single form, an implicit assumption made about the baseline hazards in more conventional models of high-frequency dynamics. The hazard rates of most aggressive orders (A1, A2, B1, and B2) tend to decline faster over the first few seconds after the previous market event. The hazard functions for the risk of complete removal of liquidity at the current touch (types A8 and B8) have a double-humped shape, with the first local maximum attained 0.25 seconds after the previous event and the second (global) maximum attained around 3.5 seconds after the previous event. For the remaining hazard functions, there is also a tendency to decline with the duration, but not as pronounced as for the baseline hazards of most aggressive limit orders.

The plots of hazard rates in Figures 2.2 and 2.3 do not offer support to the hypothesis of inverse U-shaped baseline hazards, the pattern commonly accepted when the baseline hazard functions of interarrival times and other risks are fitted parametrically, for example, in the parametric specification (2.7) by Bisière and Kamionka [13] studied above, or in a conceptually similar model by Hautsch [71] for the time intervals between LIFFE Bund futures transactions. Therefore, the duration dependence of hazard functions for individual risks is unlikely to be well explained by the gamma distributed individual heterogeneity multiplier, the model which is usually applied to justify the Singh–Maddala form (2.7) of the observed mixed hazard functions (Addison and Portugal [2]). On the other hand, despite the failure of the Singh–Maddala parametric form to capture the shapes of the baseline hazards

cation of this paper (types A6, A7, A10, A11, B6, B7, B10, and B11 in Tables 2.1 and 2.2) are omitted from the graphs. All those hazard rates demonstrate a tendency for slow decline with the duration time, and look very similar to the baseline hazards of the events of types A4 and B4.

implied by the semiparametric CPH model for notional risks, most parameter estimates in the two models appear consistent with each other.

To illustrate the misspecification of hazard functions imposed by form (2.7), Figures 2.4 and 2.5 report the estimated baseline hazard functions implied by the Singh–Maddala functional form for publicly observed seller- and buyer-initiated events and obtained by substitution of coefficients $\widehat{\alpha}_r$ and $\widehat{\gamma}_r$ in equation (2.7) evaluated at $\mathbf{z} = 0$. Apparently, the plots implied by the parametric model (2.7) are do not to capture the salient features of the plots reported in Figures 2.2 and 2.3 for the more flexible semiparametric CPH model specification. In particular, the peculiar bimodal structure of baseline hazard functions for cancellation events A8 and B8 is completely missing from the plots of Figures 2.4 and 2.5. Moreover, the parametric model is picking up the narrow left spikes which are less essential for close approximation of the general tendency, while missing completely the major maxima to the right. As for the other baseline hazards, their estimates from the parametric model correctly pick up the general decreasing trend in the nonparametric hazard rates but appear to decline too fast, at least for the first five seconds since the last observed event.

The negative biases of market depth coefficients in Tables 2.5 and 2.6 are likely to originate from the misspecification described above. High levels of market depth are generally associated with the periods of active trading, characterized by relatively high hazard rates and small observed intervals between consecutive events. Conversely, longer observed durations and lower hazard rates are occur more frequently when the market depth is below average. If the estimated hazards implied by the Singh–Maddala parametric form decline faster than the true hazard functions over the most informative range of durations (which is, presumably, the first few seconds after the observed events), the specification errors could be reduced if the smaller coefficients were assigned to the covariates that are positively correlated with the hazard rates. In particular, this intuition predicts the negative biases for most market depth coefficients in the fully parametric model, the kind of pattern confirmed by Tables 2.7–2.8.

Finally, we give a simple example of the evolution of hazard rates for notional risks

estimated from the CPH model by the k-NN kernel methods as described in the beginning of this subsection. Figures 2.6 and 2.7 display the estimates based on the quartic kernel weighting scheme. The nearest neighbor "bandwidth" ratio k/J is taken equal to 0.05 in both figures. The following hypothetical scenario has been selected. Suppose the limit order book is characterized by a relatively low but positive bid-ask spread ($Spread = 2$), a large quoted demand for USD at the current best bid ($Q_{\text{bid}} = 10$) and a small supply of USD at the current best offer price ($Q_{\text{ask}} = 2$). Moreover, the most recently recorded trade was a buyer-initiated transaction at the current midquote of the bid-ask spread ($Side = -1$, $Slippage = 0$) and the last event that occured on the market is complete cancellation of liquidity at the previous best market price on the bid side, i.e., the bid price deterioration event (B8).

The plots displayed in both figures are very similar and generic for the evolutions of hazard rates observed under several alternative scenarios. In the scenario described above, the odds that the next market event will be buyer-initiated are almost twice as high as the odds of a seller-initiated event; this asymmetry is primarily determined by the direction of the last trade. Even more disparity is observed between the odds of buyer- vs. seller-initiated *aggressive* limit and market orders; much of this imbalance is also explained by the sign of the last trade and the quoted market depth. Finally, the larger is the time elapsed after the last event, the smaller are the odds that the next event will be a transaction and the larger are the relative chances to observe a submission or cancellation of a limit order at the touch or just within the bid-ask spread.

2.5 Conclusion

The main contribution of this chapter can be summarized as follows. First, the unified competing risks framework is proposed for the analysis of multivariate marked point processes in continuous time. The model has an attractive behavioral interpretation and applies naturally to the irregularly spaced observations frequently encountered in applied finance and

economics. The asymptotic theory developed by Andersen *et al.* [8] for counting processes can be adopted with minimal adjustments to conduct the inference for competing risks. The model of this chapter represents a semiparametric alternative to the fully parametric model of limit order trading by Bisière and Kamionka [13] and can be applied more generally without imposing unjustified restrictions on the form of the baseline hazards, the practice leading to potential biases of the estimated covariate effects and incorrect inferences.

The model is applied to analyze the timing and interaction between quotes and trades in the Reuters D2000-2 electronic brokerage system, the segment of the foreign exchange market that has been rarely studied before, despite its rapid development in the 1990s. The major stylized facts about the high-frequency foreign exchange dynamics that have been confirmed in the application of the model to empirical data include:

- clustering of market activity on the directional characteristics of last trade ("buyer or seller pressure")
- considerable sensitivity of the order submission strategies employed by traders to the state of the limit order book and the quoting and trading history.

These conclusions will be reaffirmed and analyzed more thoroughly in Chapter 4 where the original model is further extended in several directions.

2.6 Appendix: Tables and Graphs

Table 2.1: Arrivals and cancellations of market and limit orders (offers) to sell US dollars

Type (1)	Event (ask side) (2)	Quoted price change (3)	Quoted depth change (4)
A1**	Offer at $P^* < P_b$	$\Delta P_b < 0$, $\Delta P_a \leq 0$	Q_b changes, Q_a may not
A2**	Market sell order arrival	$\Delta P_b \leq 0$	Q_b changes to 2nd best
A3**	Offer at $P^* = P_b$	No price effect	$\Delta Q_b < 0$
A4*	Offer at $P^* < P_a$	$\Delta P_a < 0$	$Q_a = Q^*$
A5*	Offer at $P^* = P_a$	No price effect	$\Delta Q_a > 0$
A6	Offer at $P_a < P^* \leq P_a + 2$	No price effect	No quantity effect
A7	Offer at $P^* \geq P_a + 3$	No price effect	No quantity effect
A8*	Cancel all at $P^* = P_a$	$\Delta P_a > 0$	Q_a changes to 2nd best
A9*	Cancel part at $P^* = P_a$	No price effect	$\Delta Q_a < 0$ at the touch
A10	Cancel at $P_a < P^* \leq P_a + 2$	No price effect	No quantity effect
A11	Cancel at $P^* \geq P_a + 3$	No price effect	No quantity effect

Table 2.1 summarizes the sell-side event classification scheme introduced in this chapter. Events are identified by the changes in the limit order book. The rows of Table 2.1 correspond to the eleven distinct types of events associated with sellers' activity and termed "ask-side events." The type of event determined by a combination of order type (market or limit order), character of activity (submission or cancellation of a limit order), and by the arriving (or cancelled) limit order price P^* relative to the prevailing best bid and ask quotes P_{bid} and P_{ask} (column 2). The consequences for the best bid and ask prices and for the liquidity at these prices are summarized in columns 3 and 4. The types of events marked in column 1 by single and double stars can always be detected on the trading screens. The event types marked by the double stars typically trigger transactions at prices. The unmarked event types associated with limit order arrivals or cancellations at suboptimal prices cannot be identified from the common part of the D2000-2 trading screens.

Table 2.2: Arrivals and cancellations of market and limit orders (bids) to purchase US dollars

Type (1)	Event (bid side) (2)	Quoted price change (3)	Quoted depth change (4)
B1**	Bid at $P^* > P_a$	$\Delta P_a > 0$, $\Delta P_b \geq 0$	Q_a changes, Q_b may not
B2**	Market buy order arrival	$\Delta P_a > 0$	Q_a changes to 2nd best
B3**	Bid at $P^* = P_a$	No price effect	$\Delta Q_a < 0$
B4*	Bid at $P^* > P_b$	$\Delta P_b > 0$	$Q_b = Q^*$
B5*	Bid at $P^* = P_b$	No price effect	$\Delta Q_b > 0$
B6	Bid at $P_b - 2 \leq P^* < P_b$	No price effect	No quantity effect
B7	Bid at $P^* \leq P_b - 3$	No price effect	No quantity effect
B8*	Cancel all at $P^* = P_b$	$\Delta P_b < 0$	Q_b changes to 2nd best
B9*	Cancel part at $P^* = P_b$	No price effect	$\Delta Q_b < 0$ at the touch
B10	Cancel at $P_b - 2 \leq P^* < P_b$	No price effect	No quantity effect
B11	Cancel at $P^* \leq P_b - 3$	No price effect	No quantity effect

Table 2.2 summarizes the buy-side event classification scheme introduced in this chapter. Events are identified by the changes in the limit order book. The rows of Table 2.2 correspond to the eleven distinct types of events associated with buyers' activity and termed "bid-side events." The type of event determined by a combination of order type (market or limit order), character of activity (submission or cancellation of a limit order), and by the arriving (or cancelled) limit order price P^* relative to the prevailing best bid and ask quotes P_{bid} and P_{ask} (column 2). The consequences for the best bid and ask prices and for the liquidity at these prices are summarized in columns 3 and 4. The types of events marked in column 1 by single and double stars can always be detected on the trading screens. The event types marked by the double stars typically trigger transactions at prices. The unmarked event types associated with limit order arrivals or cancellations at suboptimal prices cannot be identified from the common part of the D2000-2 trading screens.

Table 2.3: Covariates describing the state of the limit order book and the trading history

Covariate	Definition
Side	1 if prior trade seller-initiated, -1 if buyer-initiated
Spread	Difference between the best market ask and bid quotes
Profit ex post	Ex-post profitability of prior transaction (see the text)
$\log(Q_{ask})$	Natural log of the depth at the lowest ask quote
$\log(Q_{bid})$	Natural log of the depth at the lowest bid quote

Table 2.3 on this page describes the list of covariates $\mathbf{x}$, while Table 2.4 on the next page describes the list of covariates $\mathbf{d}$ in the fully parametric and semiparametric Cox proportional hazard (CPH) specifications $h_r(t|\mathbf{z}) = h_{0r}(t)\exp(\mathbf{z}'\boldsymbol{\beta}_r)$, $\mathbf{z} = (\mathbf{x}', \mathbf{d}')'$ for the competing risks $r = 1, ..., R$. The components of vector $\mathbf{x}$ incorporate the information about the limit order book and the past transactions, which is available from the screens prior to the events. Components of vector $\mathbf{d}$ are indicator variables taking values 0 or 1, depending on the type $s = 1, ..., S$ ($S = 14$) of event that has been last publicly observed on the market.

Table 2.4: Covariates describing the recent history of the limit order book

State dummy d_s	Definitions of Markov states $s = 1, ..., S$ $(S = 14)$
$d_1 = d_{[A9]}$	Last event = A9: partial cancellation of the lowest offer
$d_2 = d_{[A8]}$	Last event = A8: full cancellation of the lowest offer
$d_3 = d_{[A5]}$	Last event = A5: arrival of limit order to sell at the lowest offer
$d_4 = d_{[A4]}$	Last event = A4: arrival of limit order to sell within spread
$d_5 = d_{[A3]}$	Last event = A3: arrival of sell limit order at the highest bid, when size of the new limit order does not exceed size of bid
$d_6 = d_{[A2]}$	Last event = A2: arrival of market sell order
$d_7 = d_{[A1]}$	Last event = A1: arrival of sell limit order at the highest bid when size exceeds the bid size, or arrival of sell order below bid
$d_8 = d_{[B1]}$	Last event = B1: arrival of buy limit order at the lowest ask when size exceeds the ask size, or arrival of buy order above ask
$d_9 = d_{[B2]}$	Last event = B2: arrival of market buy order
$d_{10} = d_{[B3]}$	Last event = B3: arrival of buy limit order at the lowest ask, when size of the new limit order does not exceed size of ask
$d_{11} = d_{[B4]}$	Last event = B4: arrival of limit order to buy within spread
$d_{12} = d_{[B5]}$	Last event = B5: arrival of limit order to buy at the highest bid
$d_{13} = d_{[B8]}$	Last event = B8: full cancellation of the highest bid
$d_{14} = d_{[B9]}$	Last event = B9: partial cancellation of the highest bid

Table 2.5: Fully parametric maximum likelihood estimates for competing risks of seller-initiated events in the limit order book (Markov state coefficients omitted)

Risk Type (r)	Side	Spread	Slippage	$\log(Q_{ask})$	$\log(Q_{bid})$	$\widehat{\gamma}_r - 1$
A1: Hit by Aggr.Ask	−0.186* (−10.98)	−0.352* (−32.26)	0.070* (15.97)	0.172* (8.30)	−0.833* (−30.06)	0.83* (28.01)
A2: Hit by Market	0.242* (20.40)	−0.333* (−33.30)	0.059* (12.41)	0.065* (5.44)	−0.162* (−23.33)	0.76* (36.30)
A3: Hit by Small Ask	0.131* (9.28)	−0.350* (−22.60)	0.072* (14.04)	0.092* (5.07)	0.235* (12.29)	1.21* (37.50)
A4: Offer within Spread	0.066* (9.45)	0.048* (13.28)	0.023* (6.38)	0.156* (12.73)	−0.121* (−11.01)	1.44* (47.97)
A5: Offer at the Touch	0.051* (6.10)	−0.083* (−14.54)	−0.107* (−18.16)	−0.083* (−7.35)	−0.113* (−10.47)	0.96* (37.14)
A6: Offer above Touch	0.001 (0.08)	−0.274* (−27.32)	−0.169* (−19.11)	−0.359* (−27.02)	−0.083* (−7.08)	0.49* (19.02)
A7: Offer Far from Touch	−0.026* (−2.95)	−0.099* (−12.11)	−0.130* (−19.44)	−0.251* (−16.22)	−0.137* (−11.48)	0.70* (24.68)
A8: Full Canc.Ask Touch	0.218* (15.62)	−0.176* (−12.74)	−0.182* (−16.10)	−0.479* (−32.49)	−0.119* (−5.87)	1.38* (23.92)
A9: Part.Canc.Ask Touch	0.124* (9.92)	−0.045* (−7.43)	−0.097* (−9.73)	0.640* (38.36)	−0.090* (−6.61)	1.73* (49.69)
A10: Canc.above Touch	0.211* (19.18)	−0.231* (−31.26)	−0.169* (−20.32)	−0.176* (−12.86)	−0.130* (−10.07)	0.71* (23.05)
A11: Canc.Far from Touch	0.233* (19.94)	−0.103* (−12.87)	−0.089* (−10.95)	−0.097* (−7.71)	−0.140* (−11.17)	0.83* (31.88)

Table 2.5 reports the fully parametric maximum likelihood estimates $(\widehat{\boldsymbol{\beta}}_r', \widehat{\gamma}_r)'$ of competing risks $r = 1, ..., R$, based on the sample of ask-side events between 6 a.m. and 5 p.m. GMT on the week of October 6–10, 1997. The baseline hazard functions $h_{0r}(t)$ are specified using a biparametric Singh−Maddala form $h_{0r}(t) = \frac{\gamma_r t^{\gamma_r - 1}}{1 + \alpha_r t^{\gamma_r}}$. The estimated coefficients at the Markov state covariates $\mathbf{d}$, omitted for brevity, are available from the author upon request. The t-statistics reported in parentheses for vectors $(\widehat{\boldsymbol{\beta}}_r', \widehat{\gamma}_r - 1)'$ are based on the robust ("sandwich-type") estimates of the variance-covariance matrix of the coefficients. Entries of the table are marked by stars, whenever the estimated coefficients are statistically significant at the 95% level of confidence.

Table 2.6: Fully parametric maximum likelihood estimates for competing risks of buyer-initiated events in the limit order book (Markov state coefficients omitted)

Risk Type (r)	Side	Spread	Slippage	$\log(Q_{ask})$	$\log(Q_{bid})$	$\widehat{\gamma}_r - 1$
B1: Taken by Aggr.Bid	−0.111* (−5.53)	−0.368* (−21.15)	0.222* (19.67)	−0.866* (−33.89)	0.155* (7.13)	0.91* (30.51)
B2: Taken by Market	−0.181* (−14.84)	−0.367* (−32.80)	0.126* (13.90)	−0.160* (−13.86)	0.046* (4.17)	0.73* (35.23)
B3: Taken by Small Bid	−0.037* (−2.05)	−0.412* (−27.56)	0.181* (17.08)	0.192* (10.89)	0.107* (5.72)	1.29* (32.96)
B4: Bid within Spread	−0.044* (−4.59)	0.060* (20.74)	0.050* (14.04)	−0.124* (−10.37)	0.137* (11.70)	1.48* (61.16)
B5: Bid at the Touch	−0.012 (−1.50)	−0.069* (−12.60)	−0.048* (−10.10)	−0.105* (−10.03)	−0.090* (−7.77)	0.89* (36.16)
B6: Bid below Touch	0.046* (5.57)	−0.250* (−28.16)	−0.061* (−9.13)	−0.085* (−7.79)	−0.333* (−26.72)	0.51* (18.40)
B7: Bid Far from Touch	0.047* (5.05)	−0.102* (−15.45)	−0.052* (−9.27)	−0.134* (−11.55)	−0.286* (−21.54)	0.63* (26.39)
B8: Full Canc.Bid Touch	−0.192* (−11.53)	−0.117* (−8.07)	−0.064* (−13.11)	−0.114* (−7.24)	−0.488* (−27.90)	1.47* (21.22)
B9: Part.Canc.Bid Touch	−0.121* (−13.11)	−0.035* (−5.68)	−0.067* (−8.00)	−0.097* (−7.52)	0.636* (37.41)	1.70* (46.22)
B10: Canc.below Touch	−0.174* (−19.91)	−0.217* (−26.57)	−0.066* (−10.35)	−0.132* (−11.22)	−0.194* (−14.56)	0.69* (23.01)
B11: Canc.Far from Touch	−0.216* (−19.53)	−0.082* (−12.70)	−0.053* (−10.50)	−0.154* (−11.22)	−0.125* (−10.84)	0.77* (30.37)

Table 2.6 reports the fully parametric maximum likelihood estimates $(\widehat{\beta}_r', \widehat{\gamma}_r)'$ of competing risks $r = 1, ..., R$, based on the sample of bid-side events between 6 a.m. and 5 p.m. GMT on the week of October 6–10, 1997. The baseline hazard functions $h_{0r}(t)$ are specified using a biparametric Singh–Maddala form $h_{0r}(t) = \frac{\alpha_r t^{\gamma_r - 1}}{1 + \alpha_r t^{\gamma_r}}$. The estimated coefficients at the Markov state covariates $\mathbf{d}$, omitted for brevity, are available from the author upon request. The t-statistics reported in parentheses for vectors $(\widehat{\beta}_r', \widehat{\gamma}_r - 1)'$ are based on the robust ("sandwich-type") estimates of the variance-covariance matrix of the coefficients. Entries of the table are marked by stars, whenever the estimated coefficients are statistically significant at the 95% level of confidence.

Table 2.7: Cox regressions for competing risks of seller-initiated events in the limit order book (Markov state coefficients omitted)

Risk Type (r)	Side	Spread	Slippage	$\log(Q_{ask})$	$\log(Q_{bid})$
A1: Hit by Aggress.Ask	0.231* (9.40)	−0.379* (−22.91)	0.069* (12.28)	0.217* (8.69)	−0.832* (−27.14)
A2: Hit by Market	0.366* (27.66)	−0.314* (−29.71)	0.060* (10.36)	0.179* (13.42)	−0.085* (−6.21)
A3: Hit by Small Ask	0.191* (9.20)	−0.381* (−20.66)	0.075* (12.41)	0.151* (6.85)	0.310* (14.70)
A4: Offer within Spread	0.088* (6.89)	0.050* (11.25)	0.009 (1.38)	0.246* (17.24)	−0.012 (−0.82)
A5: Offer at the Touch	0.058* (5.67)	−0.068* (−12.50)	−0.121* (−18.28)	0.129* (10.85)	0.001 (0.05)
A6: Offer above Touch	−0.002 (−0.13)	−0.199* (−24.24)	−0.178* (−18.89)	−0.112* (−8.10)	0.061* (4.60)
A7: Offer Far from Touch	−0.028* (−2.26)	−0.063* (−9.67)	−0.139* (−15.34)	−0.024 (−1.68)	0.002 (0.18)
A8: Full Canc.Ask Touch	0.263* (13.84)	−0.118* (−12.16)	−0.166* (−14.98)	−0.359* (−17.37)	0.009 (0.41)
A9: Part.Canc.Ask Touch	0.146* (10.32)	−0.043* (−5.86)	−0.111* (−8.83)	0.866* (47.74)	−0.030 (−1.85)
A10: Canc.above Touch	0.243* (19.09)	−0.170* (−22.21)	−0.182* (−20.12)	−0.007 (−0.45)	0.016 (1.12)
A11: Canc.Far from Touch	0.262* (20.09)	−0.069* (−9.42)	−0.097* (−9.10)	0.047* (3.30)	−0.016 (−1.09)

Table 2.7 reports the estimates $\widehat{\beta}_r$ in the parametric part of the semiparametric CPH model for competing risks $r = 1, ..., R$, based on the sample of ask-side events recorded between 6 a.m. and 5 p.m. GMT on the week of October 6-10, 1997. The estimated coefficients at the Markov state covariates $\mathbf{d}$, omitted for brevity, are available from the author upon request. The t-statistics reported in parentheses are based on the robust estimates of the variance-covariance matrix of coefficients (Lin and Wei [87]). Statistically significant covariate coefficients (at the 95% level) are marked by stars.

Table 2.8: Cox regressions for competing risks of buyer-initiated events in the limit order book (Markov state coefficients omitted)

Risk Type (r)	Side	Spread	Slippage	$\log(Q_{ask})$	$\log(Q_{bid})$
B1: Taken by Aggress.Bid	−0.171* (−6.84)	−0.405* (−21.86)	0.219* (16.61)	−0.877* (−27.80)	0.209* (8.34)
B2: Taken by Market	−0.320* (−24.32)	−0.335* (−30.36)	0.140* (12.76)	−0.080* (−6.01)	0.176* (13.42)
B3: Taken by Small Bid	−0.097* (−4.71)	−0.441* (−28.05)	0.187* (12.99)	0.263* (12.54)	0.163* (7.22)
B4: Bid within Spread	−0.063* (−4.92)	0.058* (17.58)	0.048* (8.89)	−0.021 (−1.43)	0.217* (14.88)
B5: Bid at the Touch	−0.018 (−1.80)	−0.054* (−11.26)	−0.050* (−9.13)	0.005 (0.45)	0.129* (10.71)
B6: Bid below Touch	0.050* (4.40)	−0.176* (−22.19)	−0.060* (−9.72)	0.053* (4.02)	−0.086* (−6.30)
B7: Bid Far from Touch	0.050* (4.30)	−0.059* (−10.07)	−0.053* (−8.85)	0.014 (1.09)	−0.035* (−2.52)
B8: Full Canc.Bid Touch	−0.237* (−12.81)	−0.075* (−7.04)	−0.063* (−9.44)	−0.003 (−0.17)	−0.368* (−18.11)
B9: Part.Canc.Bid Touch	−0.147* (−10.62)	−0.033* (−5.11)	−0.065* (−6.74)	−0.039* (−2.45)	0.865* (47.33)
B10: Canc.below Touch	−0.205* (−16.69)	−0.152* (−19.54)	−0.065* (−10.41)	0.011 (0.79)	−0.022 (−1.55)
B11: Canc.Far from Touch	−0.247* (−20.14)	−0.050* (−8.51)	−0.054* (−9.30)	−0.024 (−1.75)	0.032* (2.28)

Table 2.8 reports the estimates $\widehat{\beta}_r$ in the parametric part of the semiparametric CPH model for competing risks $r = 1, ..., R$, based on the sample of bid-side events recorded between 6 a.m. and 5 p.m. GMT on the week of October 6-10, 1997. The estimated coefficients at the Markov state covariates $\mathbf{d}$, omitted for brevity, are available from the author upon request. The t-statistics reported in parentheses are based on the robust estimates of the variance-covariance matrix of coefficients (Lin and Wei [87]). Statistically significant covariate coefficients (at the 95% level) are marked by stars.

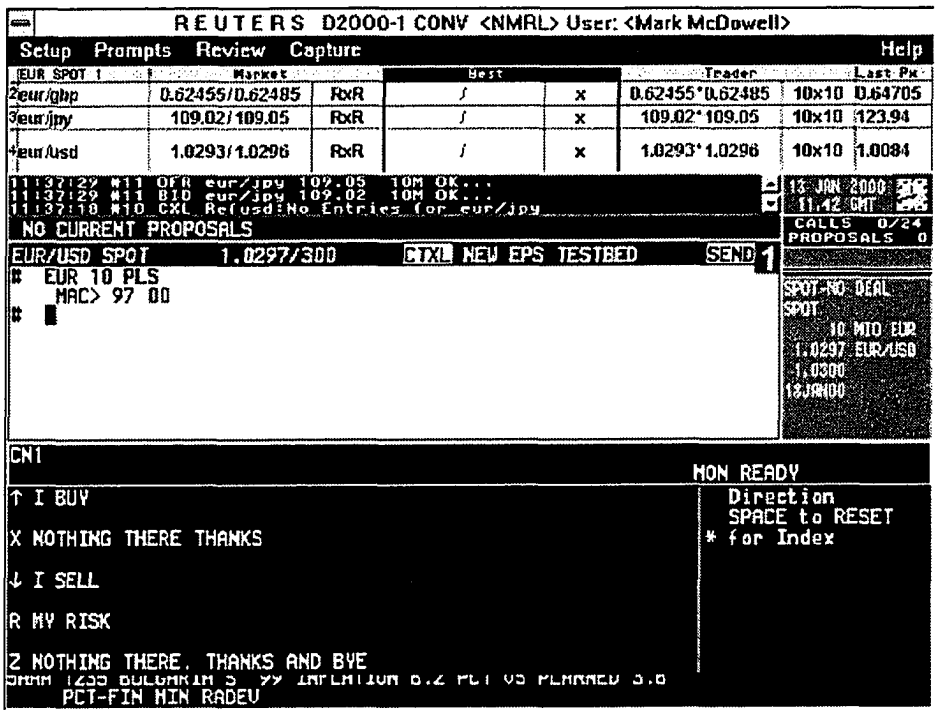


Figure 2.1: The Reuters D2000-2 automated brokerage terminal

Part of display in the center of the screen on Figure 2.1 contains electronic communication messages for direct bilateral trades transmitted via D2000-1 system (which is not covered in this dissertation). The upper part of the display provides information about the state of the Reuters D2000-2 electronic limit order book. The dealer can choose up to five exchange rates or select just one pair. The exchange rates are displayed in two alternative formats: in the upper left side of the screen the best market quotes and quantities available at these quotes are displayed; to the upper right, the dealer can see the best bid and ask quotes and quantities available to him. In the upper right corner of the screen, the dealer observes the direction and price of the last trade through the Reuters D2000-2 automated brokerage. The image of the Reuters D2000-2 electronic trading screen used in this insert was taken from the Reuters information webpage at http://about.reuters.com/transactions/d22s.htm which provides detailed information on the Reuters D2000 trading system.

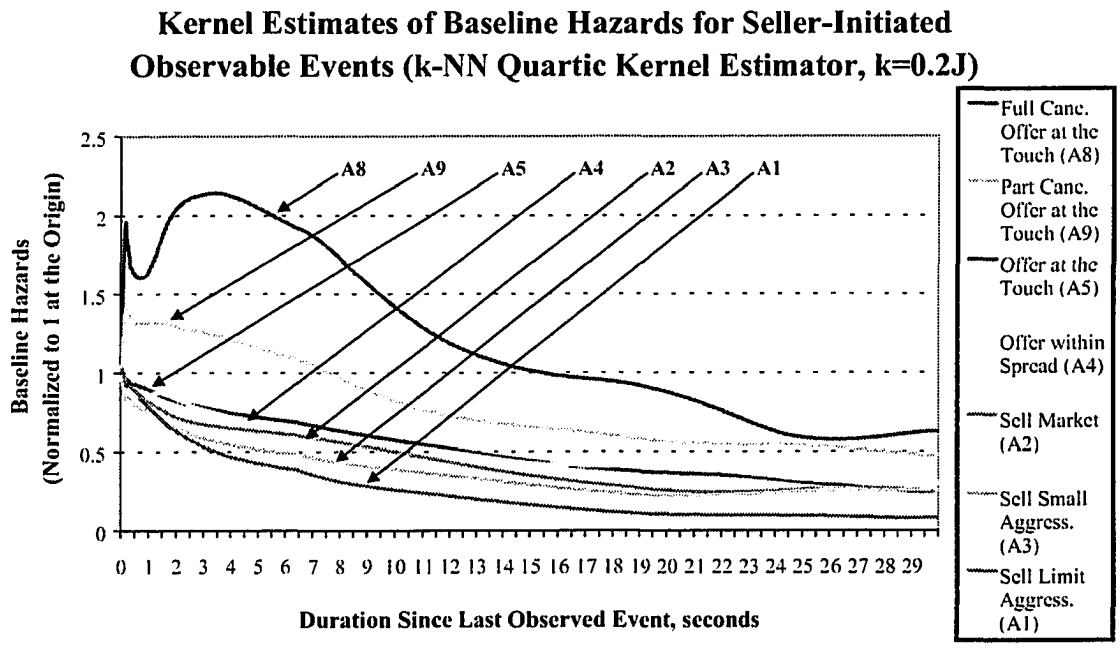


Figure 2.2: Kernel estimates of baseline hazards for seller-initiated observable events

Figure 2.2 shows the normalized k-nearest neighbor estimates $h_{0r}^{**}(t)$ of the baseline hazard functions for publicly observable seller-initiated events. The hazards of notional risks are shown on the graph in the following sequence, from top to bottom: A8, A9, A5, A4, A2, A3, A1. The weights assigned to the nearest neighbors are determined by the quartic function (2.15) with the "neighborhood" as large as 20% of distinct duration values in the sample. All estimated hazard functions are normalized to unity near the origin, to facilitate the comparison of their shapes.

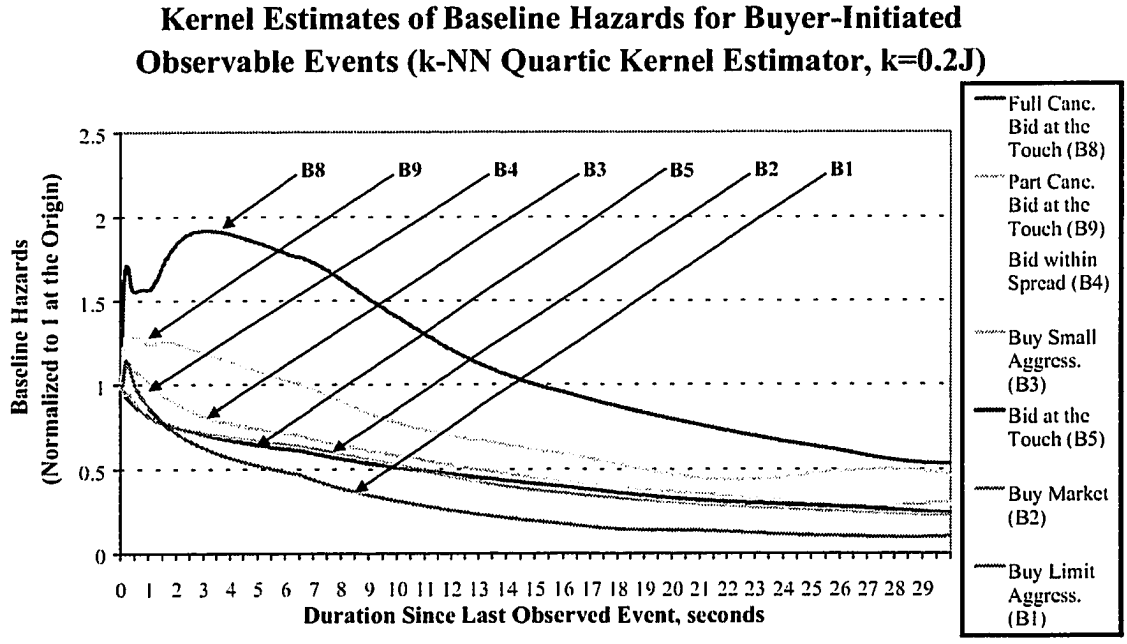


Figure 2.3: Kernel estimates of baseline hazards for buyer-initiated observable events

Figure 2.3 shows the normalized k-nearest neighbor estimates $h^{**}_{0r}(t)$ of the baseline hazard functions for publicly observable buyer-initiated events. The hazards of notional risks are shown on the graph in the following sequence, from top to bottom: B8, B9, B4, B3, B2, B5, B1. The weights assigned to the nearest neighbors are determined by the quartic function (2.15) with the "neighborhood" as large as 20% of distinct duration values in the sample. All estimated hazard functions are normalized to unity near the origin, to facilitate the comparison of their shapes.

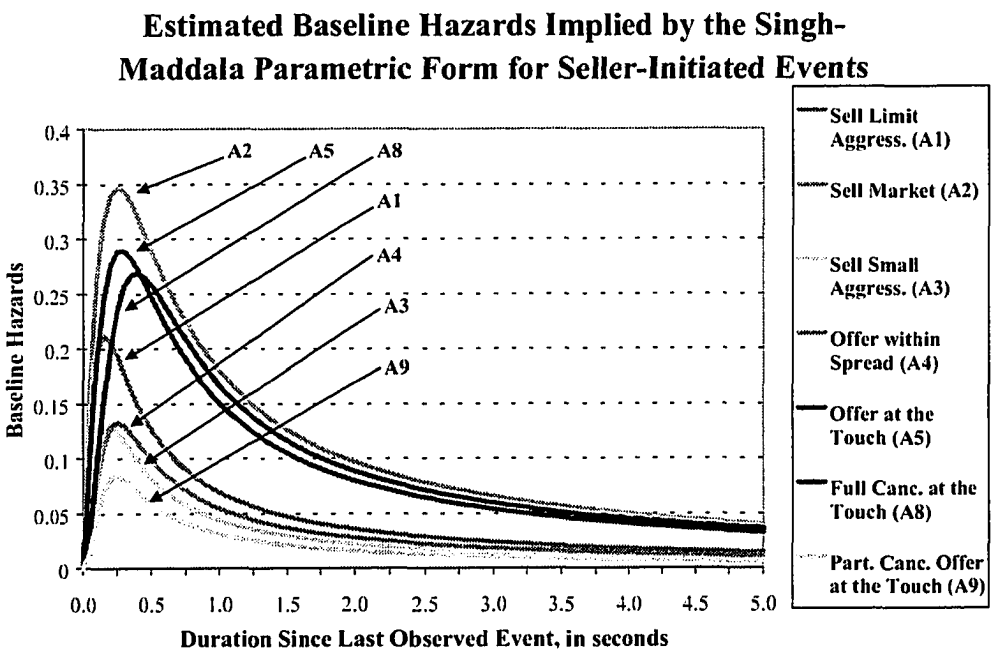


Figure 2.4: Estimated baseline hazard functions implied by Singh–Maddala parametric form for seller-initiated events

Figure 2.4 shows the estimated baseline hazards

$$h_{0r}(t) = \frac{\widehat{\gamma}_r t^{\widehat{\gamma}_r - 1}}{1 + \widehat{\alpha}_r t^{\widehat{\gamma}_r}}$$

implied by the Singh–Maddala functional form (2.7) for publicly observable seller-initiated events. The hazards of seller-initiated events are shown in the following sequence, from top to bottom: A2, A8, A5, A1, A4, A3, A9. The values of parameters $\widehat{\alpha}_r$ and $\widehat{\gamma}_r$ for all types of events are equal to the fully parametric maximum likelihood estimates reported in Table 2.5.

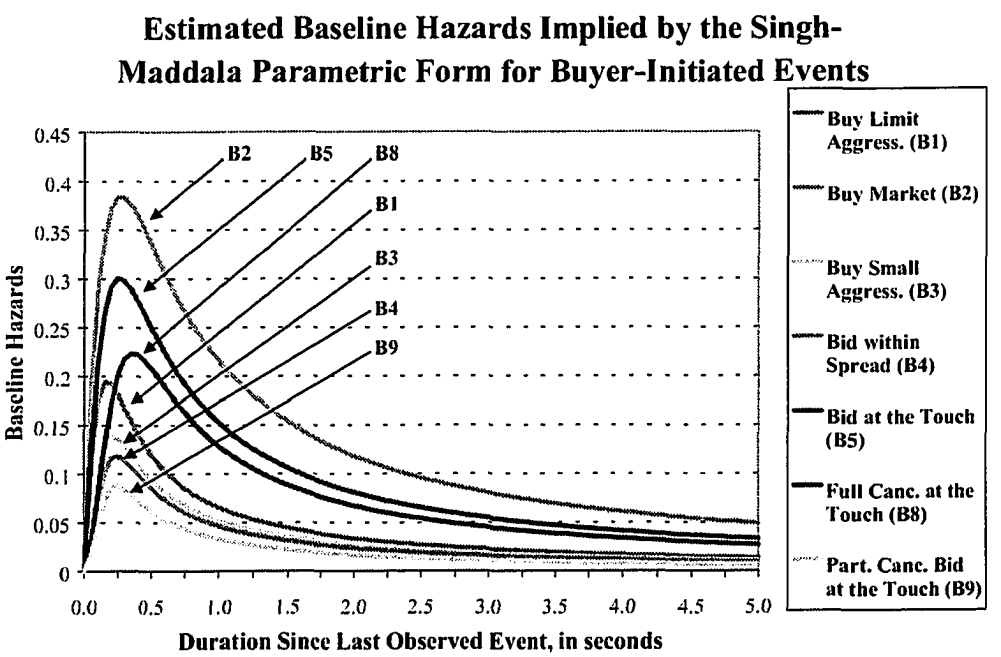


Figure 2.5: Estimated baseline hazard functions implied by Singh–Maddala parametric form for buyer-initiated events

Figure 2.5 shows the estimated baseline hazards

$$h_{0r}(t) = \frac{\widehat{\gamma}_r t^{\widehat{\gamma}_r - 1}}{1 + \widehat{\alpha}_r t^{\widehat{\gamma}_r}}$$

implied by the Singh–Maddala functional form (2.7) for publicly observable buyer-initiated events. The hazards of buyer-initiated events are shown in the following sequence, from top to bottom: B2, B5, B8, B1, B3. B4, B9. The values of parameters $\widehat{\alpha}_r$ and $\widehat{\gamma}_r$ for all types of events are equal to the fully parametric maximum likelihood estimates reported in Table 2.6.

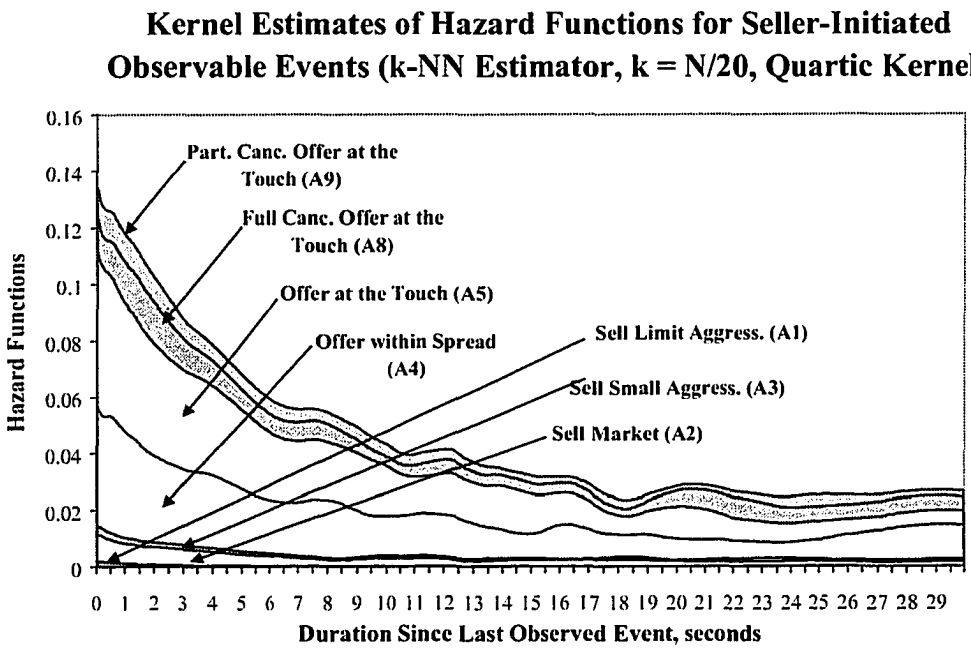


Figure 2.6: Kernel estimates of risks for seller-initiated events in a hypothetical scenario

Figure 2.6 displays the k-nearest neighbor estimates $h_{0r}^{**}(t)\exp(\mathbf{z}_n'\widehat{\boldsymbol{\beta}}_r)$ of the hazard functions for publicly observable seller-initiated events, stacked from bottom to top in the order A1–A5. A8. A9. The plots are constructed for the covariates summarized by vector $\mathbf{z} = (\mathbf{x}', \mathbf{d}')'$ implied by the following hypothetical trading history: (1) the current bid-ask spread equals two ticks (0.02 Pfennig); (2)the quoted depth equals two million dollars at the best offer price and ten million dollars at the best bid price; (3) the last prior trade was a buyer-initiated transaction that occured in the middle of the current bid-ask spread; (4) the last event observed on the market is complete cancellation of bid at the previous best bid quote (leading to price deterioration. event B8). The weights assigned to the nearest neighbors are determined by the quartic function (2.15). with the "window" equal to 20% of the number of distinct duration values in the sample.

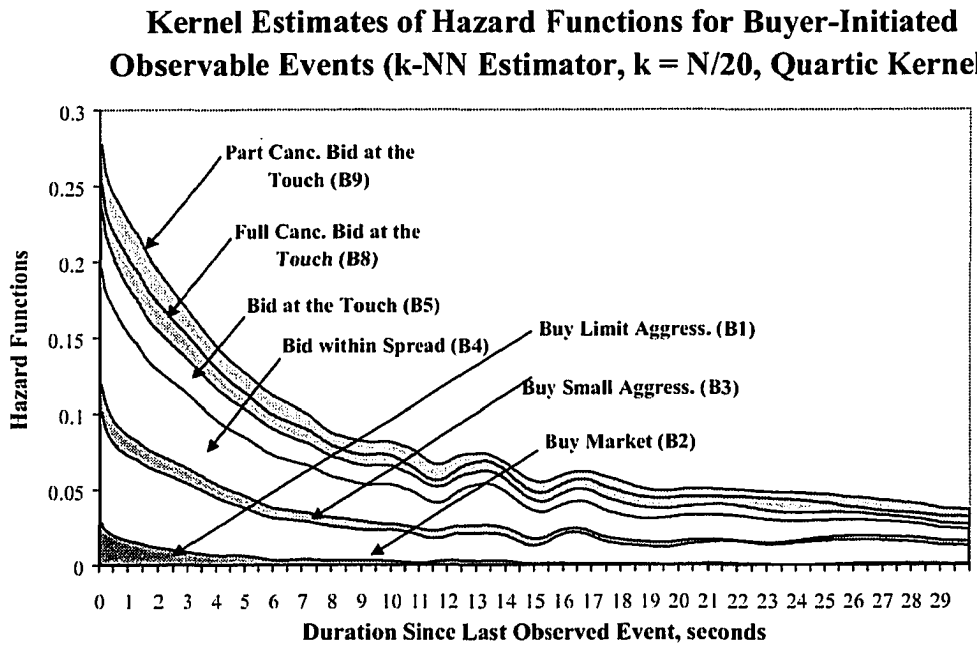


Figure 2.7: Kernel estimates of risks for buyer-initiated events in a hypothetical scenario

Figure 2.7 displays the k-nearest neighbor estimates $h_{0r}^{**}(t)\exp(\mathbf{z}_n'\widehat{\boldsymbol{\beta}}_r)$ of the hazard functions for publicly observable buyer-initiated events, stacked from bottom to top in the order B1–B5, B8, B9, respectively. The plots are constructed for the covariates summarized by vector $\mathbf{z} = (\mathbf{x}', \mathbf{d}')'$ implied by the following hypothetical trading history: (1) the current bid-ask spread equals two ticks (0.02 Pfennig); (2)the quoted depth equals two million dollars at the best offer price and ten million dollars at the best bid price; (3) the last prior trade was a buyer-initiated transaction that occured in the middle of the current bid-ask spread; (4) the last event observed on the market is complete cancellation of bid at the previous best bid quote (leading to price deterioration, event B8). The weights assigned to the nearest neighbors are determined by the quartic function (2.15), with the "window" equal to 20% of the number of distinct duration values in the sample.

Chapter 3

Statistical Foundations of the Competing Risks Theory: A Counting Process Approach

3.1 Introduction

This chapter introduces and develops statistical foundations for the econometric framework of competing risks. The model can be applied, with modifications, to any sequence of events that occur at random, rather than predetermined, time intervals. In the context of electronic limit order market, such events can naturally be associated with changes of the information publicly available from the trading screens. The events can also include public "news" announcements or any other information that arrives at irregular time intervals, and therefore can be modeled as a point process in continuous time.

The central example considered in this chapter involves application of the popular statistical methodology of competing risks to the analysis of the order flow and price formation in the Reuters D2000-2 brokerage system, which has been one of the main liquidity providers in the electronic segment of the foreign exchange market. The basic assumption behind the

general set-up of competing risks is that the state of a subject of study (in our case, it will be primarily the publicly observable part of the electronic limit order book) can be changed by a finite number of causes (sometimes called *notional risks*), which can be potentially of entirely different nature and usually are assumed independent. At any moment of time the relative importance of notional risks is determined by the odds of their instantaneous realization which can be numerically characterized by the risk-specific hazard functions. After a single risk is realized and the limit order book modifies, the remaining risks that competed with each other under the prior market conditions become irrelevant in the new state. Then the "internal clock" of the "race" is initialized, and a new "race" begins immediately among the participating notional risks under the new market conditions, proceeding up to the moment when the next "winner" triggers another change in the limit order book, and so forth.

From a broader perspective, the competing risks approach developed in this chapter considers the limit order flow as a sequence of independent or quasi-independent realizations of a multivariate marked point process (Snyder and Miller [121]), the number of its components being equal to the number of identified notional risks. The tradition of using univariate and bivariate marked point processes in the analysis of high-frequency financial data was started relatively recently with a series of papers by Engle and Russell [41], [42], Engle and Lunde [40], Gouriéroux *et al.* [55], Engle [39], and Russell *et al.* [119]. In this literature the irregularly spaced arrivals of transactions or quotes are modeled by a self-exciting point process with memory, the autoregressive conditional duration (ACD) being its most popular specification. The markers are usually represented by quotes or transaction prices and volumes (when available) and modeled conditional on their arrival times.

One major impediment to broad applications of this approach to multivariate financial data has been the unresolved issue of how the markers of qualitatively distinct nature should be treated. This problem is quite difficult since the intervals between irregularly spaced ticks (events) of one type can overlap with the intervals between irregularly spaced ticks of other types in a complicated way, and there is no natural ordering of such multiple duration intervals. For instance, the researcher extending the ACD framework to bivariate tick-by-tick

data on Deutsche Mark/U.S. dollar and Japanese yen/U.S. dollar exchange rates needs to decide whether he is interested primarily in the durations between all consecutive quotes, between the quotes coming from the same market, or, probably, the duration intervals initiated by a Deutsche Mark quote followed by a Japanese yen quote. A similar problem arises when the ACD model is applied to simultaneous modeling of quotes and trades that come from the same market. If the durations between transactions and subsequent quote revisions represent the primary object of investigation (Engle and Lunde [40]), then transaction ticks can be naturally assumed a forcing variable driving the quote revisions. Similarly, an extension of this approach to multivariate setting requires imposing a specific recursive structure describing interaction between different types of events. In this setting, modeling directly the distribution of durations is subject to *ad hoc* restrictions on the information arrival process which are usually hard to justify from the first principles.

The competing risks model developed in this chapter treats events of each type symmetrically and thus avoids the need to select the "driving process." It has at least two counterparts in the recent literature. Bisière and Kamionka [13] apply a fully parametric competing risks model to the analysis of dynamics and sequencing of orders to trade the Alcatel shares at the Paris Bourse, gives a joint explanation of the duration between consecutive orders and their aggressiveness, and stresses the important role of information about the limit order book in the price discovery process. Unlike Bisière and Kamionka [13], this chapter treats the hazard functions of competing risks semiparametrically, allowing more freedom in the specification of baseline hazards without increasing computational requirements.

Symmetric treatment of components of the multivariate point process also makes the model of this chapter similar in many respects to the multivariate version of the autoregressive conditional intensity model introduced by Russell [118]. The feature of the competing risks methodology that distinguishes it from the mainstream literature on point processes is the inherently limited observability of components of the underlying multivariate stochastic process due to a self-censoring mechanism directly built into the structure of the model. This

leads to a natural behavioral interpretation of the hazard functions of competing risks as the instantaneous likelihoods of observing the action of a particular type of trader in the market populated by agents possessing heterogeneous information, using different models, and exercising different trading strategies. This representation appears to be especially relevant in the foreign exchange market, where the true model is far from obvious and the information is widely dispersed across traders.

This chapter starts with a stylized presentation of trading history in the form of a sequence of marked random events. Section 3.2 casts this idea into the framework of counting processes and puts the approach outlined in the previous chapter on the solid theoretical ground. The emphasis is made on the results that prove useful for the development of the likelihood-based approach to semiparametric estimation of competing risks in section 3.3. Section 3.4 reviews some of the large sample properties for the popular estimators of competing risks in the Cox regression context. Section 3.5 applies the methods and theory developed in the previous sections to the problem of baseline hazard rate estimation for a range of limit order book events. Finally, the popular simulations procedures in the competing risks framework and their practical implementation are briefly discussed in section 3.6.

3.2 An Econometric Model of the Limit Order Book

In this section we start with the formalization of the trading history process as a sequence of marked random events. Then we recast the model in the counting process framework, give necessary definitions, and online the technical assumptions that will facilitate development of the asymptotic theory of competing risks in the later sections. We complete this section with a set of examples that will clarify the definitions.

3.2.1 Events, Epochs, Sub-Epochs, and the Internal Clock

Suppose we observe N time intervals $(T_{n-1}; T_n]$, $n = 1, 2, ..., N$, with the endpoints identified by consecutive occurences of publicly observable events. Such events are assumed to have one of the S types $r = 1, 2, ..., S$, and can be interpreted as observations generated by the underlying economic process. In the context of a pure limit order market, any such event marks the beginning of a new *epoch*. At the outset of an epoch the internal clock is re-started and run until the first arrival of a publicly available information represented by the components of the covariate vector $\mathbf{Z}_n$. The covariates may be representing the information on the electronic trading screens, time-of-the-day indicators, public news announcements, etc. By design, vector $\mathbf{Z}_n$ is predetermined by the previous history and by the initial conditions, and thus it remains constant throughout the entire epoch n.

In addition to publicly observable events, there are events of types $r = S + 1, ..., R$ that are assumed to be unobserved by the general public. More precisely, occurences of unobservable events of types $r = S + 1...., R$ remain private information of the agents initiating such events. Submissions and cancellations of buy and sell limit orders at suboptimal prices may serve as examples of such events in the electronic market environment. Any epoch n contains a non-negative number of such unobservable events, which naturally identify *sub-epochs* within a given time interval $(T_{n-1}; T_n]$. The internal clock of the system is not re-started at the beginning of sub-epochs until the beginning of the next epoch, and the covariate vector $\mathbf{Z}_n$ representing the publicly available information also remains unchanged until the occurrence of next observable event at the random time T_n.

The general structure of an epoch is shown on Figure 3.1. To fix the notation, we assume throughout this section that the sample contains the total number N of observable and $N^* - N$ unobservable events. We also assume $R \ll N$, which means that the number of event types must be much smaller than the number of interarrival epochs in the sample. The terms "event", "failure", and "risk" will be used interchangeably. The term "previous history" will refer to the records of durations and events prior to the beginning of a given epoch.

3.2.2 A Counting Process Representation of the Event History

Let the random process $N_{nr}(t)$ count the events of type r that occur between zero and t seconds after the beginning of the nth epoch $(T_{n-1}; T_n]$. Thus $N_{nr}(\cdot)$ is a non-decreasing univariate counting process taking values in $\{0; 1\}$ for the observable events of types $r = 1, ..., S$, and taking any non-negative integer values for unobservable events of types $r = S+1, ..., R$. Consider the random process

$$N_{\cdot r}(t) = N_{1r}(t) + N_{2r}(t) + ... + N_{Nr}(t),$$

which counts the total number of events of type r in the duration interval $(0; t]$ since the inception of the sample interval. Assume that all processes $N_{nr}(t)$ are well-defined with respect to a right-continuous, increasing, and complete filtration $(\mathcal{F}_n(t))_{t=0}^{\tau}$, where τ is a fixed positive number or infinity, which means that they obey the usual technical conditions imposed in the statistical literature on counting processes (see Andersen *et al.* [8], p.60).

The dynamic properties of the analyzed collection of counting processes

$$\mathbf{N} = (N_{nr},\ r = 1, ..., R;\ n = 1, ..., N) \tag{3.1}$$

are characterized by their compensators

$$\Lambda(\boldsymbol{\theta}) = (\Lambda_{nr}(\boldsymbol{\theta}),\ r = 1, ..., R;\ n = 1, ..., N). \tag{3.2}$$

The compensators (3.2) are absolutely continuous and given by

$$\Lambda_{nr}(t; \boldsymbol{\theta}) = \int_0^t \lambda_{nr}(u; \boldsymbol{\theta}) du, \quad r = 1, ..., R.\ n = 1, ..., N, \tag{3.3}$$

where $\boldsymbol{\theta}$ is the parameter of interest. The individual intensity processes $\lambda_{nr}(\cdot; \boldsymbol{\theta})$ in (3.3) have a multiplicative structure

$$\lambda_{nr}(t; \boldsymbol{\theta}) = Y_{nr}(t) h_{nr}(t|\mathbf{Z}_n; \boldsymbol{\theta}), \quad r = 1, ..., R,\ n = 1, ..., N.$$

The indicators $Y_{nr}(t)$ of being at risk of type r at time $t-$ of epoch n are predictable (even though not necessarily publicly observable) binary processes which are independent of $\boldsymbol{\theta}$.

Moreover, we assume for convenience that for any type of notional risk r in epoch n the corresponding hazard rate $h_{nr}(\cdot)$ satisfies the Cox's proportional hazard model

$$h_{nr}(t|\mathbf{Z}_n;\boldsymbol{\theta}) = h_{0r}(t;\gamma)\exp(\mathbf{Z}_n'\boldsymbol{\beta}_r). \quad r = 1,...,R,\ n = 1,...,N, \tag{3.4}$$

where $\mathbf{Z}_n = (Z_{n1},....Z_{np})'$ is the vector of covariates observed at the beginning of the nth epoch.[1] The parameter $\boldsymbol{\theta}$ has a semiparametric structure,

$$\boldsymbol{\theta} = (\boldsymbol{\beta}_1',....,\boldsymbol{\beta}_R';\gamma)',$$

where each $\boldsymbol{\beta}_r$, $r = 1,....,R$, has dimension p, whereas γ describing the shape of baseline hazard $h_{0r}(\cdot;\gamma)$ can generally belong to an infinitely dimensional functional space. The only assumption about the baseline hazards $h_{0r}(\cdot)$ that we need at this point is that $h_{0r}(\cdot)$ are non-negative with

$$\int_0^t h_{0r}(u)du < \infty. \quad r = 1,...,R. \tag{3.5}$$

for all $t \in [0;\tau)$.

Components of the joint process $N_n(t) = (N_{nr}(t))_{r=1}^R$ are assumed to be independent with respect to the filtration $(\mathcal{F}_n(t))_{t=0}^\tau$. This means that conditional on $(N_n(u))_{u=0}^t$ and the covariates $\mathbf{Z}_n$, we postulate the existence of the intensity

$$\lambda_{n\cdot}(t;\boldsymbol{\theta}) = \lim_{\delta\to 0+}\frac{1}{\delta}\Pr\{N_n(t+\delta) - N_n(t) > 0|\mathbf{Z}_n, (Y_{nr}(u))_{r=1}^R, 0 \le u \le t\}$$

of the joint counting process $N_n(\cdot)$ and the relationship

$$\lambda_{n\cdot}(t;\boldsymbol{\theta}) = \sum_{r=1}^R \lambda_{nr}(t;\boldsymbol{\theta})$$

between the intensities of the joint process and its components, which holds for all $n = 1,...,N$ and $0 \le t < \tau$.

[1] In this dissertation only the case of fixed covariates completely specified at $t = 0$ needs to be dealt with. Correspondingly, all theoretical results are formulated bearing in mind the specific context of this work. However, most of the analysis is extended without major changes to the more general case of predictable and locally bounded time-varying covariates.

To formulate the asymptotic properties of the hazard function estimates of individual events, we make some technical assumptions about the covariate process $(\mathbf{Z}_n)_{n=1}^N$. In particular, assume that it is stationary, ergodic, and supported by a compact subset in $\mathbf{R}^p$. Denote by $\Phi_r(\mathbf{z}, t)$, $r = 1, ..., R$, the conditional distribution functions of the covariate vector $\mathbf{Z}_n$, provided that t seconds after the beginning of a randomly chosen epoch $(T_{n-1}; T_n]$ the system is still at risk of type r,[2] that is,

$$\Phi_r(\mathbf{z}, t) = \Pr\{\mathbf{Z}_n \leq \mathbf{z} | Y_{nr}(t-) = 1\}.$$

If there exists an absolutely continuous non-increasing function

$$S_r(t) = \Pr\{Y_{nr}(t-) = 1\}.$$

which can be interpreted as the probability of survival by time t while being under the risk of type r at time $t-$, then the distribution function $\Phi_r(\mathbf{z})$ of the covariate $\mathbf{Z}_n$, conditional that the system is under the risk of type r, can be defined as follows

$$\begin{aligned} \Phi_r(\mathbf{z}) &= \Pr\{\mathbf{Z}_n \leq \mathbf{z} | Y_{nr}(t-) = 1 \text{ for some } t \in [0; \tau)\} \\ &= -\int_0^\tau \Phi_r(\mathbf{z}, t) dS_r(t). \end{aligned}$$

Finally, consider

$$\mathcal{F}_n(t) = \sigma\{(N_n(u))_{u=0}^t, (Y_n(u))_{u=0}^t, \mathbf{Z}_n\}$$

and define the filtration

$$\mathcal{F}(t) = \bigvee_{n=1}^N \mathcal{F}_n(t).$$

With the above definitions, $\lambda_n(t) = (\lambda_{nr}(t))_{r=1}^R$ is a predictable process having independent components with respect to $\mathcal{F}_n(t)$, and hence with respect to $\mathcal{F}(t)$, while the processes

$$M_n(t) = (M_{nr}(t))_{r=1}^R = (N_{nr}(t) - \Lambda_{nr}(t))_{r=1}^R$$

[2]We assume that all at-risk indicators Y_{nr}, $r = 1, ..., R$. are set to zero immediately at the end of epoch n.

with compensators

$$\Lambda_{nr}(t) = \int_0^t \lambda_{nr}(u)du$$

are square-integrable local martingales with respect to $\mathcal{F}(t)$, which also have independent components on the interval of durations $[0;\tau]$. More precisely,

$$\Lambda_n(t) = (\Lambda_{nr}(t))_{r=1}^R$$

is the compensator of vector process

$$N_n(t) = (N_{nr}(t))_{r=1}^R$$

with respect to the filtration $\mathcal{F}_n(t) \vee \mathcal{H}_\tau(t)$, where $\mathcal{H}_\tau(t)$ is the σ-field generated by the entire future of

$$Y_n(t) = (Y_{nr}(t))_{r=1}^R$$

within the given epoch n.

Before describing the methods that may be used for estimation of this rather abstract general model of event history, it is worthwhile to recast the key concepts in the context of two simple examples.

3.2.3 Example 1: Two Competing Risks

Suppose we observe a sample $(\mathbf{Z}_n, T_n, \delta_n)_{n=1}^N$, where $\mathbf{Z}_n$ is a vector of covariates, $\delta_n = \mathbf{1}\{T_{\text{Sell},n} < T_{\text{Buy},n}\}$ is the indicator of whether the nth epoch ends with a seller-initiated (vs. buyer-initiated) event, and $T_n = \min(T_{\text{Buy},n}, T_{\text{Sell},n})$. In the prototypical case of two competing risks without covariates, the only observables are the outcomes (T_n, δ_n) of the two latent duration variables $T_{\text{Buy},n}$ and $T_{\text{Sell},n}$ (Figure 3.2).

3.2.4 Example 2: Multiple Risks in a Dynamic Limit Order Market Environment

The central application of the competing risks theory developed in this chapter pertains to the analysis of trading history in an electronic limit order book, which can be completely characterized by the sample $(\mathbf{Z}_n, Y_n(t), T_n, r_n)_{n=1}^N$, where $\mathbf{Z}_n$ is a set of covariates that includes publicly observable characteristics of the limit order book at the beginning of epoch n (which is a subset of the information that appears on trading screens at the outset of epoch n), as well as a publicly available recent history of the book. As usual in a competing risks environment, denote by

$$T_n = \min_{r=1,\ldots,R} T_{nr}$$

the duration of epoch n, and by

$$r_n = \arg\min_{r=1,\ldots,R} T_{nr}$$

the type of event causing the end of epoch n,[3] and $Y_n(t) = (Y_{nr}(t))_{r=1}^R$ is the vector of at-risk indicator processes, which are usually functions of covariates $\mathbf{Z}_n$ and may also depend on the state of the limit order book, which may be hidden and undetectable on the trading screens. The simplified illustration for three types of risk, of which only two are assumed to be publicly observable, is given in Figure 3.3.

We conclude this section with three alternative forms of at-risk indicators $Y_{nr}(t)$. The first form characterized by the property $Y_{nr}(0+) \stackrel{\text{a.s.}}{=} 1$ corresponds to the typical situation where the market conditions at the beginning of epoch n do not preclude the occurrence of an event of type r at any moment within the given epoch.

The second form corresponds to the case $Y_{nr}(t) \equiv 0$ for all $t \in [0; \tau)$. For instance, if r is the risk that the next event will be submission of a limit order within the bid-ask spread, and the current value of the bid-ask spread equals the minimum possible value of one tick, then risk r can be effectively eliminated from the set of the risks competing in the current

[3]For the time being we rule out the tied durations, so that type r of the event identifying the end of epoch is determined uniquely.

epoch n since it can never occur.

The third form may be exemplified by the risk r that cancellation of a subsidiary limit order occurs one tick above the current best ask price, whereas the limit order book does not currently contain any sell orders at such a price. Therefore, risk r is not among those competing for time priority in the beginning of the current epoch. However, if submission of a subsidiary sell limit order one tick above the current best ask price occurs prior to any observable event, risk r can re-emerge and therefore should be included into the set of competing risks. For example, this can happen because the at-risk indicator $Y_{nr}(t)$, which equals zero at the beginning of the epoch, may switch its value once or several times before the epoch ends. Of course, any change of the indicator marks the start of a new sub-epoch, but not necessarily of a new epoch. (See the definitions of epoch and sub-epoch in the beginning of this section.)

3.3 Semiparametric Estimation of the Competing Risks Model

This section lays out statistical foundations of the likelihood-based approach to the semiparametric estimation of competing risks. The most popular methods of hazard rate estimation are either fully parametric or nonparametric. While the theory behind parametric methods is rather straightforward, the estimation results can be biased if the parametric model is misspecified. On the other hand, purely nonparametric techniques often suffer from the "curse of dimensionality" leading to unacceptably wide confidence bands. Therefore one may expect to take advantage of both parametric and nonparametric techniques by hitting the middle ground with the semiparametric model.

In this section we focus on a semiparametric likelihood-based estimation procedure which is robust and computationally efficient. Robustness can be a serious concern in applications since economics and financial theory usually leave at least part of the model

unspecified. Moreover, frequently there is a logical gap between the formulation of a highly stylized theoretical models of market microstructure delivering crisp predictions under the purist assumptions and the loosely specified data-driven econometric models used to test implications of the theory. Taking into consideration multiple sources of noise and errors in the empirical data, successful implementation of empirical methods often requires large data sets that might involve tens and even hundred thousands observations. Additionally, one should approach seriously the problem of computational efficiency, since the methods successfully applied to the samples of moderate size often become excessively cumbersome and impractical when adapted to large databases, which is usually the case in high-frequency financial econometrics.

The following review begins with the description of the profile likelihood method for hazard rate estimation. To illustrate this method we use the Cox proportional hazard specification which leads to an extremely simple form of the partial likelihood function and has attractive robustness properties. Moreover, the likelihood function can be efficiently maximized using the standard statistical software. The partial likelihood estimation naturally leads to the classical Nelson–Aalen estimator for the cumulative hazard functions. Next we review the idea of smoothing the increments of this estimator which was originally applied by Ramlau-Hansen [115] to estimation of the baseline hazard functions. After heuristic discussion of the deficiencies of the proposed "internal" estimator we present an alternative "external" estimator, which has the same first-order asymptotics as the Ramlau-Hansen estimator but is expected to perform better in the presence of heavily censored observations, which is a typical situation in the models with multiple competing risks. Finally we provide an extension of the locally constant "external" estimator to the locally linear "external" estimator. Results of this section lay the groundwork for the analysis of asymptotic properties of the proposed estimators in the next section.

3.3.1 The Nelson–Aalen estimators for the Cumulative Hazards and the Incidence Rates

Assume that the available dataset consists of the triplets of censored data $(\mathbf{Z}_n, T_n, r_n)$, $n = 1, ..., N$, where $\mathbf{Z}_n$ is a vector of covariates, and the pair (T_n, r_n) represents the time and type of the observed event,

$$T_n = \min_{r=1,...,R} T_{nr}, \quad r_n = \arg\min_{r=1,...,R} T_{nr}.$$

The stopping times T_{nr} of the underlying counting processes (3.1) are defined as

$$t_{nr} \equiv \inf\{t \in [0;\tau) : \Delta N_{nr}(t) > 0\}, \quad r = 1, ..., R; \; n = 1, ..., N.$$

The partial likelihood function of the data is proportional to the product integral

$$\prod_{n=1}^{N} \prod_{t\in[0;\tau)} \left\{ \prod_{r=1}^{R} (dH_{0r}(t) \exp(\mathbf{Z}_n'\boldsymbol{\beta}_r))^{\Delta N_{nr}(t)} \left(1 - \sum_{r=1}^{R} dH_{0r}(t) S_r^{(0)}(\boldsymbol{\beta}_r, t)\right)^{1-\Delta N_{n\cdot}(t)} \right\} \tag{3.6}$$

defined on the complete dataset (3.1) of vector processes $(N_{nr}, r = 1, ..., R; n = 1, ..., N)$, which are only partially observable. In formula (3.6) we use the notation

$$S_r^{(0)}(\boldsymbol{\beta}_r, t) = \sum_{n=1}^{N} Y_{nr}(t) \exp(\mathbf{Z}_n'\boldsymbol{\beta}_r), \quad r = 1, ..., R.$$

for the cumulative risk index of type r at time $t-$, and

$$H_{0r}(t) = \int_0^t h_{0r}(u)du, \quad r = 1, ..., R,$$

for the cumulative baseline hazard functions of type r.

After rewriting the partial likelihood in the form

$$\prod_{t\in[0;\tau)} \left\{ \prod_{n=1}^{N} \prod_{r=1}^{R} (dH_{0r}(t) \exp(\mathbf{Z}_n'\boldsymbol{\beta}_r))^{\Delta N_{nr}(t)} \right\} \exp\left[-\sum_{r=1}^{R} dH_{0r}(t) S_r^{(0)}(\boldsymbol{\beta}_r, t) \right] \tag{3.7}$$

and maximizing it with respect to $\Delta H_{0r}(t)$ for a fixed value of $\boldsymbol{\beta} = (\boldsymbol{\beta}_1', ..., \boldsymbol{\beta}_R')'$, we obtain

$$\Delta\widehat{H}_r(t, \boldsymbol{\beta}_r) = \frac{\Delta N_{\cdot r}(t)}{S_r^{(0)}(\boldsymbol{\beta}_r, t)} \tag{3.8}$$

and, therefore, for a fixed value of $\boldsymbol{\beta}$, $H_{0r}(t)$ is estimated by the **Nelson–Aalen estimator**

$$\widehat{H}_{0r}(t, \boldsymbol{\beta}_r) = \int_0^t \frac{J_r(u) dN_{\cdot r}(u)}{S_r^{(0)}(\boldsymbol{\beta}_r, u)}, \tag{3.9}$$

where $J_r(u) \equiv \mathbf{1}(Y_r(u) > 0)$, $Y_r = Y_{1r} + ... + Y_{Nr}$, and $N_{\cdot r} = N_{1r} + ... + N_{Nr}$.

Inserting the Nelson–Aalen estimator (3.9) into the partial likelihood formula (3.7) yields the profile likelihood which only depends on $\boldsymbol{\beta}$:

$$\begin{aligned} \mathcal{L}_{\text{part}}(\boldsymbol{\beta}) &= \mathcal{L}_{\text{Cox}}(\boldsymbol{\beta}) \prod_{t \in [0;\tau)} \prod_{r=1}^{R} \Delta N_{\cdot r}(t)^{\Delta N_{\cdot r}(t)} (1 - \Delta N_{\cdot \cdot}(t))^{1 - \Delta N_{\cdot \cdot}(t)} \\ &= \mathcal{L}_{\text{Cox}}(\boldsymbol{\beta}) \prod_{t \in [0;\tau)} \left\{ \prod_{r=1}^{R} \Delta N_{\cdot r}(t)^{\Delta N_{\cdot r}(t)} \right\} \times \exp(-N_{\cdot \cdot}(\tau)), \end{aligned}$$

where $N_{\cdot \cdot} = N_{\cdot 1} + ... + N_{\cdot R}$, and

$$\mathcal{L}_{\text{Cox}}(\boldsymbol{\beta}) = \prod_{t \in [0;\tau)} \prod_{n=1}^{N} \prod_{r=1}^{R} \left(\frac{\exp(\mathbf{Z}_n' \boldsymbol{\beta}_r)}{S_r^{(0)}(\boldsymbol{\beta}_r, t)} \right)^{\Delta N_{nr}(t)}$$

is the Cox partial likelihood. The value of $\boldsymbol{\beta}$ maximizing the log Cox partial likelihood

$$\begin{aligned} \ell_{\text{Cox}}(\boldsymbol{\beta}) &= \log \mathcal{L}_{\text{Cox}}(\boldsymbol{\beta}) \\ &= \sum_{r=1}^{R} \left[\sum_{n=1}^{N} \int_0^\tau \mathbf{Z}_n' \boldsymbol{\beta}_r dN_{nr}(t) - \int_0^\tau \log S_r^{(0)}(\boldsymbol{\beta}_r, t) dN_{\cdot r}(t) \right] \end{aligned}$$

will be denoted $\widehat{\boldsymbol{\beta}}$. Then the cumulative baseline hazard function $H_{0r}(t)$, $r = 1, ..., R$, is estimated by $\widehat{H}_{0r}(s, \widehat{\boldsymbol{\beta}}_r)$, which is the **Breslow estimator** of the baseline hazard of competing risk r.

Finally, the conditional probability that event of type r will be next to occur less than $t \leq \tau$ seconds after the previous event is

$$P_r(0, t; \mathbf{Z}_0) = \int_0^t P_0(0, u; \mathbf{Z}_0) \exp(\mathbf{Z}_0' \boldsymbol{\beta}_r) dH_{0r}(u), \quad r = 1, ..., R, \tag{3.10}$$

and can be estimated by

$$\widehat{P}_r(0, t; \mathbf{Z}_0) = \int_0^t \widehat{P}_0(0, u; \mathbf{Z}_0) \exp(\mathbf{Z}_0' \widehat{\boldsymbol{\beta}}_r) d\widehat{H}_{0r}(u, \widehat{\boldsymbol{\beta}}_r), \tag{3.11}$$

where $P_0(0,t;\mathbf{Z}_0)$ and $\widehat{P}_0(0,t;\mathbf{Z}_0)$ are respectively the conditional probability of survival at time t since last event,

$$\begin{aligned} P_0(0,t;\mathbf{Z}_0) &= \prod_{u\in[0;t)} \left(1-\sum_{r=1}^{R}\exp(\mathbf{Z}_0'\boldsymbol{\beta}_r)dH_{0r}(u)\right) \\ &= \exp\left[-\sum_{r=1}^{R}\exp(\mathbf{Z}_0'\boldsymbol{\beta}_r)H_{0r}(t)\right], \end{aligned}$$

and its estimator

$$\widehat{P}_0(0,t;\mathbf{Z}_0) = \exp\left[-\sum_{r=1}^{R}\exp(\mathbf{Z}_0'\widehat{\boldsymbol{\beta}}_r)\widehat{H}_{0r}(t,\widehat{\boldsymbol{\beta}}_r)\right].$$

3.3.2 Invariance Property of the Maximum Partial Likelihood Estimator

The partial likelihood function $\mathcal{L}_{\text{Cox}}(\boldsymbol{\beta})$ has a remarkable robustness property which makes the maximum likelihood estimator $\widehat{\boldsymbol{\beta}}$ an attractive choice in the following dynamic context. Specifically, it is invariant with respect to monotonic deformations of the time scale within the duration period $[0;\tau)$. This invariance property with respect to monotonic transformations $g(t)$ of parameter $t \in [0;\tau]$ follows from the fact that the numerical value of Cox partial likelihood function

$$\begin{aligned} \ell_{\text{Cox}}(\boldsymbol{\beta}) &= \log \mathcal{L}_{\text{Cox}}(\boldsymbol{\beta}) \\ &= \sum_{r=1}^{R}\left[\sum_{n=1}^{N}\int_0^{\tau}\mathbf{Z}_n'\boldsymbol{\beta}_r dN_{nr}(t) - \int_0^{\tau}\log S_r^{(0)}(\boldsymbol{\beta}_r,t)dN_{\cdot r}(t)\right] \end{aligned} \tag{3.12}$$

increases by a deterministic constant C_g after the parameter $t \in [0;\tau]$ is transformed monotonically into $\widetilde{t} = g(t) \in [0;g(\tau)]$. Indeed, the ordering of occurences of the failure events and their censorings is unaffected by the deterministic time scale deformation $g(\cdot)$, which implies the invariance of the second integral in (3.12). Invariance of the first integral in (3.12) with respect to the time deformation is obvious, since the integrands $\mathbf{Z}_n'\boldsymbol{\beta}_r$ are time-invariant.

3.3.3 Alternative Kernel Estimators of the Baseline Hazard Rates

The first kernel estimator for $h_{0r}(t)$ was proposed in Ramlau-Hansen [115]. It is defined as

$$\widehat{h}_{0r}(t|\mathcal{K}_r, b_r) = \frac{1}{b_r}\int_0^\tau \mathcal{K}_r\left(\frac{t-u}{b_r}\right) d\widehat{H}_{0r}(u, \widehat{\beta}_r), \tag{3.13}$$

where $\widehat{H}_{0r}(u, \widehat{\beta}_r)$ is the Nelson–Aalen martingale estimator (3.9) evaluated at point u, and the kernel function $\mathcal{K}_r(\cdot)$ satisfies the following standard conditions.

Condition K. The kernel functions $\mathcal{K}_r$, $r = 1, ..., R$, satisfy

$$\int_{-1}^{1} \mathcal{K}_r(t)dt = 1, \quad \int_{-1}^{1} t\mathcal{K}_r(t)dt = 0, \quad \int_{-1}^{1} t^2\mathcal{K}_r(t)dt = k_{2r} > 0.$$

It is easy to see that Ramlau-Hansen estimator is obtained by application of the kernel smoother to the increments of the Nelson–Aalen estimator (3.8). Since smoothing of the cumulative baseline hazard increments in (3.13) is performed prior to integration, the Ramlau-Hansen estimator (3.13) is frequently called "internal" estimator. This "internal" property implies that the estimator (3.13) depends only on the realized durations of risk r, and does not depend on the realized durations of the competing risks $r' \neq r$.

The heuristic idea behind the second method of baseline hazard estimation is that the value of $h_{0r}(\cdot)$ at the point t is likely to affect the process $N_{nr}(\cdot)$ only in a neighborhood of t. Thus in order to estimate $h_{0r}(t)$ it would be sufficient to consider a portion of the likelihood function that emphasizes behavior of the process in the locality of t. This is accomplished

by computing the kernel-weighted partial log-likelihood

$$\begin{aligned}
&(\ell_{\text{part},N}(\widehat{\boldsymbol{\beta}}) * \mathcal{K}_{\mathbf{b}})(t) \\
\equiv &\sum_{r=1}^{R} \int_{-\infty}^{+\infty} \left[\int_0^{s\wedge\tau} \sum_{n=1}^{N} Y_{nr}(u) dN_{nr}(u) (\log(h_{0r}(u)) + \mathbf{Z}_n' \widehat{\boldsymbol{\beta}}_r) \right. \\
&\left. - \int_0^{s\wedge\tau} S_r^{(0)}(\widehat{\boldsymbol{\beta}}_r, u) dH_{0r}(u) \right] \mathcal{K}_{r,b_r}(t-s) ds \\
= &\sum_{r=1}^{R} (\ell_{\text{part},r,N}(\widehat{\boldsymbol{\beta}}_r) * \mathcal{K}_{r,b_r})(t)
\end{aligned}$$

instead of the ordinary likelihood, where the kernel function $\mathcal{K}_{r,b_r}$ is defined as $\mathcal{K}_{r,b_r}(t) = \frac{1}{b_r}\mathcal{K}_r\left(\frac{t}{b_r}\right)$, and $\mathcal{K}_r(\cdot)$ satisfies the usual conditions **K**.

Let $\widehat{h}_{0r}(t)$ maximize the kernel-weighted partial likelihood

$$\begin{aligned}
(\ell_{\text{part},r,N}(\widehat{\boldsymbol{\beta}}_r) * \mathcal{K}_{r,b_r})(t) = \int_0^{\tau} \left[\sum_{n=1}^{N} Y_{nr}(u) dN_{nr}(u) \log(h_{0r}(u)) \mathcal{K}_{r,b_r}(t-u) \right. \\
\left. + \sum_{n=1}^{N} \mathbf{Z}_n' \widehat{\boldsymbol{\beta}}_r Y_{nr}(u) dN_{nr}(u) \mathcal{K}_{r,b_r}(t-u) - S_r^{(0)}(\widehat{\boldsymbol{\beta}}_r, u) h_{0r}(u) \mathcal{K}_{r,b_r}(t-u) du \right],
\end{aligned}$$

where $\mathcal{K}_{\mathbf{b}} \equiv (\mathcal{K}_{1,b_1}, \mathcal{K}_{2,b_2}, ..., \mathcal{K}_{R,b_R})'$ is an R-vector of non-negative smooth symmetric kernels $\mathcal{K}_{r,b_r}$ with support $[-1;1]$ and the vector of bandwidth parameters $\mathbf{b} = (b_1, b_2, ..., b_R)'$. Note that the kernel-weighted partial log-likelihood can be re-written so that the Cox partial log-likelihood (3.4) is factored out, yielding the "external" estimator

$$\widehat{h}_{0r}(t|\mathcal{K}_r, b_r) = \frac{\frac{1}{b_r}\int_0^{\tau} \mathcal{K}_r\left(\frac{t-u}{b_r}\right) J_r(u) dN_{\cdot r}(u)}{\frac{1}{b_r}\int_0^{\tau} \mathcal{K}_r\left(\frac{t-u}{b_r}\right) S_r^{(0)}(\widehat{\boldsymbol{\beta}}_r, u) du}. \tag{3.14}$$

This estimator was first proposed in the covariate-free case by Hjort [73] and studied by Nielsen and Linton [108], among others. The key feature of estimator $\widehat{h}_{0r}$ is its external property. Indeed, $\widehat{h}_{0r}$ incorporates the information about durations to the events of type r both directly (via realized occurences of risk events) and indirectly (via the risk index $S_r^{(0)}(\widehat{\boldsymbol{\beta}}_r, u)$ eliciting the information about censoring of risk r by its competitors $r' \neq r$).

Both Ramlau-Hansen and Hjort estimators are consistent and asymptotically normal under the mild regularity conditions (Andersen *et al.* [8]). Moreover, the second-order

properties of the asymptotic distributions of both proposed estimators are identical to each other.[4] The "internal" property of this estimator comes from the fact that its properties depend only on the realized durations of risk r, and are independent of the realized occurences of the competing risks $r' \neq r$. Therefore, despite its simplicity, the Ramlau-Hansen estimator is thought to be inferior to the Hjort estimator, especially in the situations when the realized durations of risk r are subject to heavy censoring by a large number of "competitors", which will be the case in our applications.

3.3.4 Local Linear Smoothing and Bias Reduction

The advantage of local linear smoothing approach to hazard rate estimation (Jones [80], Nielsen and Linton [108]) is that it fits locally the first-order Taylor approximation of the underlying pattern, whereas the local constant kernel estimator fits locally to the underlying pattern only a constant. Since adding in the first-order approximation usually leads to sufficient reduction of biases originating from the regular kernel smoothing, the presentation will be restricted to the local linear case.

Now the estimate of the baseline hazard rate $h_{0r}(t)$ can be obtained as follows.[5] Consider the solution to the minimization problem

$$\tilde{\boldsymbol{\theta}}(t) \equiv \arg\min_{\boldsymbol{\theta}} \sum_{n=1}^{N} \int_0^{\tau} \frac{1}{b_r} \left[\Delta N_{nr}(u) - \mathbf{x}'_{t-u}\boldsymbol{\theta}\right]^2 \mathcal{K}_r\left(\frac{t-u}{b_r}\right) Y_{nr}(u) \exp(\mathbf{Z}'_n \hat{\boldsymbol{\beta}}_r) du, \tag{3.15}$$

where $\mathbf{x}'_s\boldsymbol{\theta} = \theta_0 + \theta_1 s$ denotes the local linear trend with parameter $\boldsymbol{\theta} = (\theta_0, \theta_1)'$ and the regressors $\mathbf{x}_s = (1, s)'$. The optimal parameters $\tilde{\boldsymbol{\theta}}(t) = (\tilde{\theta}_0(t), \tilde{\theta}_1(t))'$ solve the system of

[4] This problem was discussed by Nielsen and Linton [108]. They considered a similar problem of estimating a semiparametric hazard rate with a parametric duration part and nonparametric dependence of hazard rates on a marker (covariate) process.

[5] This subsection describes only the local linear modification of the "internal" estimator of the baseline hazard rates. The local linear extension of the "external" estimator (3.14) can be obtained from an analogous procedure. The form of this estimator is similar to the nearest neighbor estimator (3.23) below.

equations

$$\sum_{n=1}^{N} \frac{1}{b_r} \int_0^\tau \mathbf{x}_{t-u} \left(\Delta N_{nr}(u) - \mathbf{x}'_{t-u} \boldsymbol{\theta} \right) \mathcal{K}_r \left(\frac{t-u}{b_r} \right) Y_{nr}(u) \exp(\mathbf{Z}'_n \widehat{\boldsymbol{\beta}}_r) du = 0, \tag{3.16}$$

which can be written as

$$\begin{aligned} & \sum_{n=1}^{N} \frac{1}{b_r} \int_0^\tau \mathbf{x}_{t-u} \mathbf{x}'_{t-u} \mathcal{K}_r \left(\frac{t-u}{b_r} \right) Y_{nr}(u) \exp(\mathbf{Z}'_n \widehat{\boldsymbol{\beta}}_r) du \cdot \boldsymbol{\theta} \\ = & \frac{1}{b_r} \int_0^\tau \mathbf{x}_{t-u} \mathcal{K}_r \left(\frac{t-u}{b_r} \right) J_r(u) dN_{\cdot r}(u) \end{aligned}$$

or, equivalently,

$$\begin{aligned} \widetilde{c}_0 \theta_0 + \widetilde{c}_1 \theta_1 &= \frac{1}{b_r} \int_0^\tau \mathcal{K}_r \left(\frac{t-u}{b_r} \right) J_r(u) dN_{\cdot r}(u), \\ \widetilde{c}_1 \theta_0 + \widetilde{c}_2 \theta_1 &= \frac{1}{b_r} \int_0^\tau (t-u) \mathcal{K}_r \left(\frac{t-u}{b_r} \right) J_r(u) dN_{\cdot r}(u), \end{aligned}$$

with

$$\widetilde{c}_m = \frac{1}{b_r} \int_0^\tau (t-u)^m \mathcal{K}_r \left(\frac{t-u}{b_r} \right) S_r^{(0)}(\widehat{\boldsymbol{\beta}}_r, u) du, \quad m = 0, 1, 2.$$

Then form the estimator

$$\begin{aligned} \widehat{h}_{0r}^{\text{lin}}(t) &= \frac{1}{b_r} \int_0^\tau \mathbf{X}_{t-u} \widetilde{\boldsymbol{\theta}}(t) \mathcal{K}_r \left(\frac{t-u}{b_r} \right) du \\ &= \frac{1}{b_r} \int_0^\tau \left[\widetilde{\theta}_0(t) + \widetilde{\theta}_1(t) \cdot (t-u) \right] \mathcal{K}_r \left(\frac{t-u}{b_r} \right) du, \end{aligned} \tag{3.17}$$

which also can be represented in the standard kernel form

$$\widehat{h}_{0r}^{\text{lin}}(t) = \frac{1}{b_r} \int_0^\tau \widetilde{\mathcal{K}}_r \left(\frac{t-u}{b_r} \right) J_r(u) dN_{\cdot r}(u), \tag{3.18}$$

with the local linear kernel

$$\widetilde{\mathcal{K}}_r(s) = \frac{\widetilde{c}_2 - \widetilde{c}_1 s}{\widetilde{c}_0 \widetilde{c}_2 - \widetilde{c}_1^2} \mathcal{K}_r(s). \tag{3.19}$$

Note the stochastic character of the corrected kernel (3.19) that makes automatic adjustment near the boundaries. The local constant estimators need special treatment near the endpoints of the sample, because of their bias for low durations and the excessive variability for high durations. The local linear estimators are much better in this respect, since it uses not only the end values but also internal changes to predict the value of hazard at the end points.

3.3.5 Adaptive Nearest Neighbor Estimation of the Baseline Hazard Functions

In empirical applications the durations used for hazard function estimation are rarely exponentially distributed, even after a large number of covariates have been extracted. When the sample of observed durations is heavily skewed, the usual kernel nonparametric estimates of baseline hazards with global bandwidth provide poor approximations to the true baseline functions in the low density domain of the observed durations. Estimation of the behavior of a system in low density regions usually represents a challenge for applied econometricians since the rare ("extreme") events are poorly modeled by conventional econometric techniques. In particular, precise estimation of functional forms of the hazard rates in these regions may be important both in simulation experiments and for global performance evaluation of such econometric models.

Two practically relevant problems associated with hazard rate estimation under the random censoring were studied extensively in the review paper by Müller and Wang [103]. The first problem concerns with the boundary effects near the endpoints of the duration support, in particular, around the origin. The second problem concerns with a substantial, often explosively growing variance of estimators in the range of extremely high durations, where the number of observations is relatively small. As demonstrated by Nielsen and Linton [108], the use of the adaptive kernel estimator defined as the ratio of the smoothed numerator to the smoothed denominator in the classical Nelson–Aalen formula (3.8) for

the hazard rate estimator substantially reduces the bias in the range of extremely short durations. The external smoother has also some advantages in the range of extremely high durations, since its denominator better accommodates the dynamics of the risk index at the censored durations which dominate the right tail of the duration range for any notional risk of interest.

The nearest-neighbor estimator, which engenders attractive properties of the fixed-bandwidth kernel estimators and is more flexible with regard to the choice of the smoothing parameter, can be proposed as a solution to the second problem. The combination of flexible choice of the bandwidth parameter and the enhanced efficiency of the external estimator can lead to substantial variance reduction in the regions of sparse sample design. However, one always should be aware that the variance reduction may be still insufficient to warrant precise estimation of baseline hazard functions in the presence of a large number of competing risks.

As has been emphasized in the beginning of this section, successful estimation based on high-frequency datasets requires fast and efficient computational algorithms. The adaptive version of a local kernel regression, which is a version of the kernel smoother procedure **ksm.ado** in *Stata*, represents an example of such algorithm. The estimator is defined as a weighted average of the k nearest neighbors to the point of interest t. The method selects $[k/2]$ neighbors to the left of the target duration t and $[k/2]$ neighbors to the right of t, assigning the weights according to the distance of neighbors to the target point. In contrast to the regular nearest neighbor estimator, the weights are assigned on the basis of distance (rather than the ranks of distance) between the neighbor and the target point t, which implies the validity (up to the higher-order terms) of the standard asymptotic formulas for kernel smoothers. If the number j^* of available durations to the left (right) of the target is less than $[k/2]$, the estimation is performed using all j^* observations to the left (right) and the usual number of $[k/2]$ durations to the right (left) of the estimation point. The weights for the internal durations (located more than $[k/2]$ sample points away from the boundaries)

are determined by one of the canonical kernels

$$\mathcal{K}_{\text{quartic}}(u) = (1-u^2)^2 \cdot \mathbf{1}_{[-1;1]}(u),$$

$$\mathcal{K}_{\text{triweight}}(u) = (1-u^2)^3 \cdot \mathbf{1}_{[-1;1]}(u), \tag{3.20}$$

$$\mathcal{K}_{\text{tricube}}(u) = (1-|u|^3)^3 \cdot \mathbf{1}_{[-1;1]}(u). \tag{3.21}$$

In this dissertation we will report estimation results only for the tricube kernel (3.21) (Cleveland [23]). It has the advantage of being twice continuously differentiable with the mass concentrated more heavily around the center of the bandwidth spectrum, which reduces the influence of distant observations when enough observations in the sample are close to the target point.[6]

This method adapts automatically to the random unbalanced sample design, where the heaviest proportion of the sample mass is concentrated at the short durations, and admits application of most results from the standard theory of kernel estimation, including the formulas for the asymptotic biases and variances of kernel estimators.[7] It can be programmed as a fast iterative procedure with a simple one-in-one-out updating of weights which correspond to $[k/2]$ left and right nearest-neighbor durations in a manner similar to the ordinary k-NN estimator described in Härdle [63], Chapter 3, for the uniform kernel, with the weighting

[6]The nearest-neighbor kernel can be modified to reduce the endpoint bias as explained above in subsection 3.3.4.

[7]Due to the random nature of sample design the proposed tricube kernel is randomly deformed so that the effective kernel is asymmetric even in the case of the local constant smoother. In fact, the kernel becomes randomly left- or right-skewed, depending on the relationship between the distances of the $[k/2]$ nearest left and right neighbors to the target points. The kernel tends to be right skewed for the sample densities with a negative first derivative and left skewed for the sample densities with a positive first derivative near the estimation point. The overall effect leads to the changes in the standard formulas for the asymptotic bias and variance in the random design case. However, in practice the size of distortions is usually small. We argue that the distortion of variance is of smaller order of magnitude and does not affect the asymptotic results of the next section. The distortion of bias (relative to the standard kernel) is of order j^*/k and is likely to change the formula for the asymptotic bias. In any event, these complications appear a relatively small price to pay for the enhanced computational efficiency of the algorithm.

window slowly moving across the data.

The formulas for the kernel smoothers used in the application in section 3.5 are given by

$$\widehat{h}_{0r}^{(\mathrm{NN})}(t_i|\mathcal{K}_r,k) = \frac{\sum_{j'=j-\lfloor k/2\rfloor}^{j+\lfloor k/2\rfloor} \mathcal{K}_r\left(\frac{t_{(j)}-t_{(j')}}{\max(t_{(j)}-t_{(j-\lfloor k/2\rfloor)},t_{(j+\lfloor k/2\rfloor)}-t_{(j)})}\right) J_r(t_{(j')})\Delta N_{\cdot r}(t_{(j')})}{\sum_{j'=j-\lfloor k/2\rfloor}^{j+\lfloor k/2\rfloor} \mathcal{K}_r\left(\frac{t_{(j)}-t_{(j')}}{\max(t_{(j)}-t_{(j-\lfloor k/2\rfloor)},t_{(j+\lfloor k/2\rfloor)}-t_{(j)})}\right) \sum_{n\in\mathcal{R}(t_{(j')})} \exp(\mathbf{z}_n'\widehat{\boldsymbol{\beta}}_r)\Delta t_{(j')}} \tag{3.22}$$

for the local constant asymmetric k-NN smoother, and by

$$\widehat{h}_{0r}^{\mathrm{lin(NN)}}(t_i|\mathcal{K}_r,k) = \frac{\sum_{j'=j-\lfloor k/2\rfloor}^{j+\lfloor k/2\rfloor} \widetilde{\mathcal{K}}_r\left(\frac{t_{(j)}-t_{(j')}}{\max(t_{(j)}-t_{(j-\lfloor k/2\rfloor)},t_{(j+\lfloor k/2\rfloor)}-t_{(j)})}\right) J_r(t_{(j')})\Delta N_{\cdot r}(t_{(j')})}{\sum_{j'=j-\lfloor k/2\rfloor}^{j+\lfloor k/2\rfloor} \widetilde{\mathcal{K}}_r\left(\frac{t_{(j)}-t_{(j')}}{\max(t_{(j)}-t_{(j-\lfloor k/2\rfloor)},t_{(j+\lfloor k/2\rfloor)}-t_{(j)})}\right) \sum_{n\in\mathcal{R}(t_{(j')})} \exp(\mathbf{z}_n'\widehat{\boldsymbol{\beta}}_r)\Delta t_{(j')}} \tag{3.23}$$

for the local linear asymmetric k-NN smoother, where the modified kernel $\widetilde{\mathcal{K}}_r(\cdot)$ is defined by formula (3.19) with

$$\widetilde{c}_m = \sum_{j'=j-\lfloor k/2\rfloor}^{j+\lfloor k/2\rfloor} (t_{(j)}-t_{(j')})^m \mathcal{K}_r\left(\frac{t_{(j)}-t_{(j')}}{\max(t_{(j)}-t_{(j-\lfloor k/2\rfloor)},t_{(j+\lfloor k/2\rfloor)}-t_{(j)})}\right) \sum_{n\in\mathcal{R}(t_{(j')})} \exp(\mathbf{z}_n'\widehat{\boldsymbol{\beta}}_r)\Delta t_{(j')}.$$

3.3.6 Spline-Based Technique for Hazard Rate Estimation

The spline-based method for hazard rate estimation, which can be considered an alternative to the semiparametric kernel techniques considered above, is a flexible parametric procedure, which accommodates a broad variety of shapes of the baseline hazard and mitigates the curse of dimensionality associated with nonparametric estimation of the baseline hazard, provided that the knots of spline intervals are chosen correctly. According to the approach described by Royston and Parmar [117], the logarithm of the baseline cumulative hazard function can often be modeled by the cubic spline function of log time. The general function $\log(H_{0r}(t))$

is approximated by the so-called *natural* cubic spline, which is by definition constrained to a linear form beyond the boundary knots $k_{\min} = \log(t_{\min})$ and $k_{\max} = \log(t_{\max})$. In addition to the boundary knots, m internal knots k_1, k_2, ..., k_m are specified, with the parameter m governing the model complexity.

The natural cubic spline is written as

$$\log(H_{0r}(\tau)) = \gamma_{0r} + \gamma_{1r}\tau + \gamma_{2r}\omega_1(\tau) + ... + \gamma_{m+1,r}\omega_m(\tau),$$

where the jth basis function is defined as

$$\omega_j(\tau) = (\tau - k_j)_+^3 - \lambda_j(\tau - k_{\min})_+^3 - (1 - \lambda_j)(\tau - k_{\max})_+^3$$

with

$$\lambda_j = \frac{k_{\max} - k_j}{k_{\max} - k_{\min}} \quad \text{and} \quad (\tau - a)_+^3 = \max\{0, (\tau - a)_+^3\}.$$

Royston and Parmar [117] propose selecting the parameter m by minimizing the Akaike information criterion, which is defined as -2 times the log likelihood plus twice the number of model parameters. In principle, the location of boundary and internal knots can be estimated since they can be treated as additional parameters. However, adding the new parameters even for a limited number of knots was found to slow down significantly the convergence of the optimization procedure. As a practical solution to this problem, the boundary knots are often placed at the most extreme observed durations, whereas the internal knots are chosen to divide the observed durations in approximately equal groups.

3.4 Asymptotic Theory of Competing Risks Estimators

In this section, we assume the Cox proportional hazard model. Extensions of most results to the general multiplicative hazard specification are not pursued in this chapter, even though the results are mostly analogous to those reviewed here. First, we lay out the conditions

for the asymptotic normality of the partial likelihood score statistics $\mathbf{U}_r(\boldsymbol{\beta}_r, \tau)$. Then the established results are used to derive the asymptotic distribution of the maximum likelihood estimator $\widehat{\boldsymbol{\beta}}_r$ of parameters $\boldsymbol{\beta}_{0r}$ and for the estimators $\widehat{H}_{0r}(t, \widehat{\boldsymbol{\beta}}_r)$ of the cumulative baseline hazard functions $H_{0r}(t)$. We also establish conditions for the asymptotic efficiency of partial likelihood estimates of the Cox regression coefficients. The approach follows closely the martingale-based techniques of Andersen and Gill [7] using the counting process machinery reviewed in Andersen *et al.* [8].

3.4.1 The Partial Likelihood Score Statistic

We need the following regularity conditions.

Condition P. $h_{0r}(\cdot)$ are non-negative with

$$\int_0^t h_{0r}(u)du < \infty, \quad r = 1, \ldots, R, \tag{3.24}$$

for all $t \in [0; \tau)$, where τ is a fixed positive number or infinity.

Condition A. There exists a neighborhood $\mathcal{B}_r$ of $\boldsymbol{\beta}_{0r}$ and the scalar, p-vector and $p \times p$ matrix functions $s_r^{(0)}$, $\mathbf{s}_r^{(1)}$, and $\mathbf{s}_r^{(2)}$, respectively, defined on $\mathcal{B}_r \times [0; \tau]$ such that for $m = 0, 1, 2$, $r = 1, \ldots, R$:

(A1) $\sup_{\boldsymbol{\beta}_r \in \mathcal{B}_r,\ t \in (0;\tau]} \left\| \frac{1}{N}\mathbf{S}_r^{(m)}(\boldsymbol{\beta}_r, t) - \mathbf{s}_r^{(m)}(\boldsymbol{\beta}_r, t) \right\| \xrightarrow{\mathsf{P}} 0$ as $N \to \infty$;

(A2) $\mathbf{s}_r^{(m)}(\cdot)$ are bounded on $\mathcal{B}_r \times [0; \tau)$ and continuous as functions of $\boldsymbol{\beta}_r \in \mathcal{B}_r$ uniformly in $t \in [0; \tau]$;

(A3) $\mathbf{s}_r^{(0)}(\boldsymbol{\beta}_{0r}, \cdot)$ are bounded away from zero on $[0; \tau]$;

(A4) $\mathbf{s}_r^{(1)}(\boldsymbol{\beta}_r, t) = \frac{\partial}{\partial \boldsymbol{\beta}_r'} s_r^{(0)}(\boldsymbol{\beta}_r, t)$, $\mathbf{s}_r^{(2)}(\boldsymbol{\beta}_r, t) = \frac{\partial^2}{\partial \boldsymbol{\beta}_r \partial \boldsymbol{\beta}_r'} s_r^{(0)}(\boldsymbol{\beta}_r, t)$ for $\boldsymbol{\beta}_r \in \mathcal{B}_r$, $t \in [0; \tau)$;

(A5) $\Sigma_r(\tau) = \int_0^\tau \mathbf{v}_r(\boldsymbol{\beta}_{0r}, t) s_r^{(0)}(\boldsymbol{\beta}_{0r}, t) h_{0r}(t)dt$ are positive definite, where $\mathbf{v}_r = \mathbf{s}_r^{(2)}/s_r^{(0)} - \mathbf{e}_r^{\otimes 2}$ and $\mathbf{e}_r = \mathbf{s}_r^{(1)}/s_r^{(0)}$.

Condition B. There exists a $\delta > 0$ such that

$$\frac{1}{\sqrt{N}} \sup_{r,n,t} |\mathbf{Z}_n| \cdot Y_{rn}(t) \cdot \mathbf{1}(\mathbf{Z}_n' \boldsymbol{\beta}_{0r} > -\delta |\mathbf{Z}_n|) \xrightarrow{\mathsf{P}} 0 \quad \text{as } N \to \infty. \tag{3.25}$$

Brief explanation of the above conditions is in order. The most significant constraints are conditions **P** and (A2). Weak convergence results obtained under these regularity conditions will be restricted to sub-intervals of the support of the distribution of the data, excluding the right-hand tail. Condition **B** is facilitates establishing the Lindeberg condition, and trivially holds for bounded covariates. Condition (A1) guarantees asymptotic stability of the functions $\mathbf{S}_r^{(m)}$, $m = 0, 1, 2$, while the regularity conditions (A3)–(A5) are standard in the asymptotic likelihood theory. As the result of these conditions, the limits with respect to N and differentiation with respect to $\boldsymbol{\beta}_r$ can be interchanged.

Now consider the vector of score statistics

$$\mathbf{U}_r(\boldsymbol{\beta}_r, \tau) = \frac{\partial}{\partial \boldsymbol{\beta}_r'} \ell_{\text{Cox}}(\boldsymbol{\beta}), \quad r = 1, \ldots, R,$$

which can be written as

$$\mathbf{U}_r(\boldsymbol{\beta}_r, \tau) = \sum_{n=1}^{N} \int_0^\tau \mathbf{Z}_n dN_{nr}(t) - \int_0^\tau \mathbf{E}_r(\boldsymbol{\beta}_r, t) dN_{\cdot r}(t),$$

where

$$\mathbf{E}_r(\boldsymbol{\beta}_r, t) = \frac{\mathbf{S}_r^{(1)}(\boldsymbol{\beta}_r, t)}{S_r^{(0)}(\boldsymbol{\beta}_r, t)}$$

and

$$\mathbf{S}_r^{(1)}(\boldsymbol{\beta}_r, t) = \sum_{n=1}^{N} Y_{nr}(t) \mathbf{Z}_n \exp(\mathbf{Z}_n' \boldsymbol{\beta}_r).$$

Note that

$$\mathbf{U}_r(\boldsymbol{\beta}_{0r}, t) = \sum_{n=1}^{N} \int_0^t (\mathbf{Z}_n - \mathbf{E}_r(\boldsymbol{\beta}_{0r}, t)) dM_{nr}(t),$$

the score statistics evaluated at $\boldsymbol{\beta}_{0r}$ and considered as processes in t, are linear combinations of stochastic integrals with respect to the local square integrable martingales

$$M_{nr}(t) = N_{nr}(t) - \int_0^t h_{0r}(u) Y_{nr}(u) \exp(\mathbf{Z}_n' \boldsymbol{\beta}_{0r}) du,$$

which are orthogonal with quadratic variation process

$$\langle M_{nr} \rangle (t) = \int_0^t h_{0r}(u) Y_{nr}(u) \exp(\mathbf{Z}_n' \boldsymbol{\beta}_{0r}) du.$$

Denote by $-\mathcal{J}_r(\boldsymbol{\beta}_r, \tau)$ the $p \times p$ matrices of second-order derivatives of $\ell_{\text{Cox}}(\boldsymbol{\beta})$ with respect to $\boldsymbol{\beta}_r$, $r = 1, ..., R$. Then $\mathcal{J}_r(\boldsymbol{\beta}_r, \tau)$ can be written as

$$\mathcal{J}_r(\boldsymbol{\beta}_r, \tau) = \int_0^\tau \mathbf{V}_r(\boldsymbol{\beta}_r, t) dN_{\cdot r}(t).$$

where

$$\mathbf{V}_r(\boldsymbol{\beta}_r, t) = \frac{\mathbf{S}_r^{(2)}(\boldsymbol{\beta}_r, t)}{S_r^{(0)}(\boldsymbol{\beta}_r, t)} - \mathbf{E}_r(\boldsymbol{\beta}_r, t)^{\otimes 2}$$

and

$$\mathbf{S}_r^{(2)}(\boldsymbol{\beta}_r, t) = \sum_{n=1}^N Y_{nr}(t) \mathbf{Z}_n^{\otimes 2} \exp(\mathbf{Z}_n' \boldsymbol{\beta}_r).$$

3.4.2 Covariate Effects

Theorem 1 *(Andersen et al. [8], Theorems VII.2.1-2.2) Assume Condition A and (3.24) hold for $t = \tau$. Then the probability that the system of equations $\mathbf{U}_r(\boldsymbol{\beta}_r, t) = \mathbf{0}$, $r = 1, ..., R$, has a unique solution $\widehat{\boldsymbol{\beta}} = \left(\widehat{\boldsymbol{\beta}}_1', ..., \widehat{\boldsymbol{\beta}}_R'\right)'$ tends to 1 and $\widehat{\boldsymbol{\beta}} \xrightarrow{\mathsf{P}} \boldsymbol{\beta}_0$ as $N \to \infty$. If, in addition, the Lindeberg-type Condition B is satisfied, then, as $N \to \infty$,*

$$\sqrt{N}(\widehat{\boldsymbol{\beta}}_r - \boldsymbol{\beta}_{0r}) \xrightarrow{\mathcal{D}} \mathcal{N}(\mathbf{0}, \Sigma_r^{-1}(\tau))$$

and

$$\sup_{t \in (0;\tau]} \left\| \frac{1}{N} \mathcal{J}_r(\widehat{\boldsymbol{\beta}}_r, t) - \Sigma_r(t) \right\| \xrightarrow{\mathsf{P}} 0$$

hold for each $r = 1, ..., R$.

3.4.3 Cumulative Hazard Functions

Theorem 2 *(Andersen et al. [8], Theorems VII.2.3-2.6) Assume Conditions A and B and that (3.24) hold for $t = \tau$. Then $\sqrt{N}(\widehat{\boldsymbol{\beta}}_r - \boldsymbol{\beta}_{0r})$ and the processes*

$$W_r(t) = \sqrt{N}\left(\widehat{H}_{0r}(t, \widehat{\boldsymbol{\beta}}_r) - H_{0r}(t)\right) + \sqrt{N}\left(\widehat{\boldsymbol{\beta}}_r - \boldsymbol{\beta}_{0r}\right)' \int_0^t \mathbf{e}_r(\boldsymbol{\beta}_{0r}, u) h_{0r}(u) du,$$

$r = 1, ..., R$, *are asymptotically independent. The limiting distribution of* $W_r(t)$ *is that of a zero-mean Gaussian martingale with variance function*

$$\omega_r^2(t) = \int_0^t \frac{h_{0r}(u)du}{s_r^{(0)}(\boldsymbol{\beta}_{0r}, u)}.$$

The process $\left(\sqrt{N}\left(\widehat{H}_{0r}(t, \widehat{\boldsymbol{\beta}}_r) - H_{0r}(t)\right)\right)_{r=1}^R$ *converges weakly to an* R*-variate Gaussian process with mean zero and covariance function* $\Gamma(s,t)$ *with the elements* $\Gamma_{rr'}(s,t)$ *given by*

$$\delta_{rr'}\omega_r^2(s \wedge t) + \int_0^s \mathbf{e}_r'(\boldsymbol{\beta}_{0r}, u)h_{0r}(u)du \cdot \Sigma_r^{-1}(\tau) \cdot \int_0^t \mathbf{e}_{r'}(\boldsymbol{\beta}_{0r'}, u)h_{0r'}(u)du, \quad r, r' = 1, ..., R,$$

which can be estimated uniformly consistently in $(s,t) \in [0;\tau) \times [0;\tau)$ *by*

$$N\left\{\delta_{rr'}\int_0^{s\wedge t} \frac{dN_r(u)}{S_r^{(0)}(\widehat{\boldsymbol{\beta}}_r, u)^2} + \int_0^s \frac{\mathbf{E}_r'(\widehat{\boldsymbol{\beta}}_r, u)dN_r(u)}{S_r^{(0)}(\widehat{\boldsymbol{\beta}}_r, u)} \cdot \mathcal{J}_r^{-1}(\widehat{\boldsymbol{\beta}}_r, \tau) \cdot \int_0^t \frac{\mathbf{E}_{r'}(\widehat{\boldsymbol{\beta}}_{r'}, u)dN_{r'}(u)}{S_{r'}^{(0)}(\widehat{\boldsymbol{\beta}}_{r'}, u)}\right\}.$$

The asymptotic covariance of $\sqrt{N}(\widehat{\boldsymbol{\beta}}_r - \boldsymbol{\beta}_{0r})$ *and* $\sqrt{N}\left(\widehat{H}_{0r}(t, \widehat{\boldsymbol{\beta}}_r) - H_{0r}(t)\right)$ *is*

$$-\Sigma_r^{-1}(\tau)\int_0^t \mathbf{e}_r(\boldsymbol{\beta}_{0r}, u)h_{0r}(u)du.$$

which can be estimated uniformly consistently in $t \in [0;\tau)$ *by*

$$-N\mathcal{J}_r^{-1}(\widehat{\boldsymbol{\beta}}_r, \tau) \cdot \int_0^t \frac{\mathbf{E}_r(\widehat{\boldsymbol{\beta}}_r, u)dN_r(u)}{S_r^{(0)}(\widehat{\boldsymbol{\beta}}_r, u)}.$$

Finally, the integrated hazard of type r *for the market conditions characterized by fixed covariates* $\mathbf{Z}_0$*, can be estimated by* $\widehat{H}_{0r}(t, \widehat{\boldsymbol{\beta}}_r)\exp(\mathbf{Z}_0'\widehat{\boldsymbol{\beta}}_r)$*. The asymptotic variance of*

$$\sqrt{N}\left(\widehat{H}_{0r}(t, \widehat{\boldsymbol{\beta}}_r)\exp(\mathbf{Z}_0'\widehat{\boldsymbol{\beta}}_r) - H_{0r}(t)\exp(\mathbf{Z}_0'\boldsymbol{\beta}_{0r})\right)$$

can be estimated uniformly consistently in $t \in [0;\tau)$ *by*

$$N(\exp(\mathbf{Z}_0'\widehat{\boldsymbol{\beta}}_r))^2\left\{\int_0^t \frac{dN_r(u)}{S_r^{(0)}(\widehat{\boldsymbol{\beta}}_r, u)^2}\right.$$
$$\left. + \int_0^t (\mathbf{E}_r(\widehat{\boldsymbol{\beta}}_r, u) - \mathbf{Z})'d\widehat{H}_{0r}(u, \widehat{\boldsymbol{\beta}}_r) \cdot \mathcal{J}_r^{-1}(\widehat{\boldsymbol{\beta}}_r, \tau) \cdot \int_0^t (\mathbf{E}_r(\widehat{\boldsymbol{\beta}}_r, u) - \mathbf{Z})d\widehat{H}_{0r}(u, \widehat{\boldsymbol{\beta}}_r)\right\}.$$

3.4.4 Local Constant External Estimator of Baseline Hazard

Consider the local constant external estimator $\widetilde{h}_{0r}(t)$ of the baseline hazard $h_{0r}(t)$ for each $r = 1, ..., R$. Define its compensator $h^*_{0r}(t)$ as

$$h^*_{0r}(t) = \frac{\frac{1}{b_r}\int_0^\tau \mathcal{K}_r\left(\frac{t-u}{b_r}\right) d\Lambda_{nr}(u;\boldsymbol{\theta})}{\frac{1}{b_r}\int_0^\tau \mathcal{K}_r\left(\frac{t-u}{b_r}\right) S_r^{(0)}(\widehat{\boldsymbol{\beta}}_r, u)du} = \frac{\frac{1}{b_r}\int_0^\tau \mathcal{K}_r\left(\frac{t-u}{b_r}\right) \sum_{n=1}^N Y_{nr}(u)h_{0r}(u)\exp(\mathbf{Z}_n\widehat{\boldsymbol{\beta}}_r)du}{\frac{1}{b_r}\int_0^\tau \mathcal{K}_r\left(\frac{t-u}{b_r}\right) S_r^{(0)}(\widehat{\boldsymbol{\beta}}_r, u)du}$$

Theorem 3 *(Andersen et al. [8], Theorem VII.2.7) Let t be an interior point in $[0;\tau]$ and assume that, for each $r = 1, ... R$, h_{0r} is twice continuously differentiable in a neighborhood of t. Assume that Conditions K, A, and B hold and that $b_r = b_r^{(N)} \to 0$ as $N \to \infty$ in such a way that $Nb_r^{(N)} \to \infty$ and $\limsup N^{1/5}b_r^{(N)} < \infty$. Then random variables*

$$\sqrt{Nb_r^{(N)}}\left(\widehat{h}_{0r}(t) - h_{0r}(t) - \frac{1}{2}b_r^2 h''_{0r}(t)_{2r}\right)$$

are asymptotically independent and

$$\sqrt{Nb_r^{(N)}}\left(\widehat{h}_{0r}(t) - h_{0r}(t) - \frac{1}{2}b_r^2 h''_{0r}(t)_{2r}\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \omega_r^2(t)) \quad as\ N \to \infty,$$

where the asymptotic variances

$$\omega_r^2(t) = \frac{h_{0r}(t)}{s_r^{(0)}(\boldsymbol{\beta}_{0r}, t)}\int_{-1}^{1} \mathcal{K}_r^2(u)du$$

can be estimated consistently by

$$\frac{N}{b_r}\int_0^\tau \mathcal{K}_r^2\left(\frac{t-u}{b_r}\right)\frac{dN_r(u)}{S_r^{(0)}(\widehat{\boldsymbol{\beta}}_r, u)^2}.$$

The second derivative of the hazard functions h''_{0r} in the bias term can be estimated consistently (although with a bias of its own) by

$$\widehat{h}''_{0r} = \frac{1}{b_r^{*3}}\int_0^\tau \mathcal{K}_r^{*\prime\prime}\left(\frac{t-u}{b_r^*}\right) d\widehat{H}_r(u, \widehat{\boldsymbol{\beta}}_r),$$

where $\mathcal{K}_r^*$, $r = 1, ... R$, are twice differentiable kernels satisfying condition K and b_r^* are corresponding bandwidths that may be different from $\mathcal{K}_r$ and b_r.

3.5 Estimation of Baseline Hazard Rates for Sell Events

To provide an example of application of nonparametric baseline hazard estimation techniques developed in this chapter, consider estimation of the baseline hazard functions for selected types out of the pull of events identified as seller-initiated competing risks in Chapter 4.[8] The risk types and the numbering of corresponding graphs are summarized in Table 3.1. The shapes of baseline hazard estimates for the symmetric buyer-initiated events are not reported since they look similar to those shown on the graphs. To conserve the space, we also omit the graphs for some baseline hazard estimates of seller-initiated events, when they are very similar to those shown on the graphs.

Figures 3.4–3.14 display the baseline hazard estimates $\widehat{h}_{0r}(t)$ (shown in white bold pattern) for the Cox proportional hazard specification (3.4) with the full set of covariates shown in Table 4.2. To illustrate the effect of the reduced covariate set on the shape of the baseline hazard rates, we also display (in darker bold pattern) the baseline hazard estimates when the full set of covariates from Table 4.2 is replaced by the smaller set of five covariates *Spread*, Q_{ask}, Q_{bid}, *Side*, and *Slippage*, which may be considered most popular characteristics of market liquidity in the microstructure literature and a good starting point in empirical model building.

The asymptotic standard errors for all estimators are based on the full and restricted sets of covariates and shown by shaded grey patterns and unshaded light curves, respectively, in the lower sections of the graphs. The standard errors of estimates are calculated by the k-nearest neighbor formula for asymptotic variance $\omega_r^2(t)$ of the external smoother, which are obtained from the fixed-bandwidth formulas of subsection 3.4.4 by rescaling the window size so that it covers exactly k sample points. All estimators are based on the locally constant tricube kernel (3.21) with $k/N = 20\%$ of the sample size, and reported for the range of durations up to 30 seconds since last publicly observable event (identified by the occurrence

[8] Description of seller-initiated events can be found in section 4.2. A brief summary of seller-initiated events is given by Table 4.1.

of markers of type A1 through A6, B1 through B6, AC6, or BC6, corresponding to the publicly observable events).[9]

The larger covariate set reduces the steepness of negative slope for all estimated baseline hazard functions, especially near the origin. One obvious explanation to this effect can be based on the observation that the larger set of covariates captures a larger portion of unaccounted heterogeneity due to time-varying market conditions and reduces the variance of mixing distribution over this unobserved heterogeneity. However, for almost all competing risks there remains substantial unaccounted heterogeneity, as the hazard rate patterns remain monotonically decreasing even after the large set of covariates has been accounted for in the Cox proportional hazard specification. Even though the tentative explanation may be misspecification of the parametric part of our model, or misclassification of events, the more serious potential source of unaccounted heterogeneity in our model is the lack of public and private information variables, such as messages from Reuters newswire, news from the futures and forward segments of the market, etc. Since it would be unrealistic to capture all unobserved heterogeneity in a single empirical model, it is important to be aware of this effect, constructing the models and using the estimation techniques that take this possibility into account.

3.6 Simulation of Competing Risks

The semiparametric bootstrap is a powerful resampling device used to approximate the distribution of statistics of interest. Even though standard asymptotic results are available for estimators of hazard rates, survivor functions, and the covariate effects in the Cox regression specification, there is no widely accepted method for inference about quantiles of the empirical duration distribution conditional on the specified vector of covariates. In the bandwidth

[9]See the formal explanation of observable events in section 3.2. The complete classification of buyer-initiated events B1–B14 and BC6–BC14 is analogous to the classification of seller-initiated events A1–A14 and AC6–AC14 given in Table 4.1.

selection problem the distribution of the mean integrated standard error (MISE), or any other criterion of choice, has to be approximated for a range of bandwidths. Bootstrapping might be the only feasible technique to obtain this distribution in practice.

3.6.1 Alternative Simulation Procedures

Since the bootstrap method of forming confidence intervals is not uniquely defined, there are usually several ways to draw the bootstrap samples when dealing with complex data. Consider, for example, the simple bootstrap procedure, which can be outlined as follows:

A. Paired bootstrap procedure

Randomly draw the triplets $(T^{(b)}, r^{(b)}, \mathbf{Z}^{(b)}) = (T_n^{(b)}, r_n^{(b)}, \mathbf{Z}_n^{(b)})_{n=1}^N$ with replacements from the original data set $(T_n, r_n, \mathbf{Z}_n)_{n=1}^N$.

Unfortunately, this simple procedure turns out numerically unstable in the context of Cox regression model with competing risks. Therefore we present two alternative bootstrap algorithms based on the general ideas of Marron [98] and Gonzalez-Manteiga *et al.* [51], who recommended to use "smoothed bootstraps" in order to approximate biases in hazard rate estimation with censored observations. There are several possible modifications of their smoothing procedures, as the bootstrapping tends to be more complicated in the presence of multiple sources of risk. Bearing in mind potential applications of these methods to large financial data sets, additional issues such as computational efficiency and numerical stability in the range of thin sample design must be taken into consideration.

The main difference between the bootstrap algorithms presented here is with regard to when the "smoothing" is done. Our algorithms parallel those described in Gonzalez-Manteiga *et al.* [51], except that we use the external smoother instead of the internal kernel estimator at the smoothing steps of the algorithms.

B. Constrained smoothed-censored (CSC) resampling plan

This method employs a variant of Efron's [38] original bootstrap idea. The method proceeds as follows:

1. Calculate the kernel smoothed versions $\widetilde{h}_{0r}(t|\mathcal{K}_r, b_r)$ of the baseline hazard rates of independent competing risks $r = 1, 2, ..., R$.

2. Draw the bootstrap samples $\mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}, ..., \mathbf{Z}^{(B)}$ of covariate vectors from the empirical distribution of covariates. Each $\mathbf{Z}^{(b)}$ represents a single bootstrap simulation of the market conditions and consists of N observations $\mathbf{z}_1^{(b)}, \mathbf{z}_2^{(b)}, ..., \mathbf{z}_N^{(b)}$.

3. Draw the bootstrap samples $\mathbf{T}_r^{(1)}, \mathbf{T}_r^{(2)}, ..., \mathbf{T}_r^{(B)}$ of the latent durations $r = 1, 2, ..., R$ generated by the CPH competing risks with the baseline hazard rates $\widetilde{h}_{0r}$ and covariates $\mathbf{z}_n^{(b)}$. Each $\mathbf{T}_r^{(b)}$ represents a single bootstrap simulation of type-r latent durations and consists of N observations $T_{1r}^{(b)}, T_{2r}^{(b)}, ..., T_{Nr}^{(b)}$.

4. For each realization $(\mathbf{T}_r^{(b)})_{r=1}^R$, $b = 1, 2, ..., B$, construct the bootstrap pair $(\mathbf{T}^{(b)}, r^{(b)}) = (T_n^{(b)}, r_n^{(b)})_{n=1}^N$, where $T_n^{(b)} = \min\limits_{r=1,...,R} T_{nr}^{(b)}$ and $r_n^{(b)} = \arg\min\limits_{r=1,...,R} T_{nr}^{(b)}$.

C. Constrained censored-smoothed (CCS) resampling plan

This method modifies Efron's original resampling scheme in such a way that the smoothing is performed *at the last step*. The sequence of steps is following:

1. Draw the bootstrap samples $\mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}, ..., \mathbf{Z}^{(B)}$ of covariate vectors from the empirical distribution of covariates. Each $\mathbf{Z}^{(b)}$ represents a single bootstrap simulation of the market conditions that consists of N observations $\mathbf{z}_1^{(b)}, \mathbf{z}_2^{(b)}, ..., \mathbf{z}_N^{(b)}$.

2. Using the raw (unsmoothed) Nelson–Aalen baseline hazards $\widehat{H}_{0r}(t, \widehat{\boldsymbol{\beta}}_r)$ and the bootstrapped covariates $\mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}, ..., \mathbf{Z}^{(B)}$, draw the bootstrap samples $\mathbf{T}_r^{(1)}, \mathbf{T}_r^{(2)}, ..., \mathbf{T}_r^{(B)}$

of the unsmoothed latent durations $r = 1, 2, ..., R$. Each $\mathbf{T}_r^{(b)}$ represents a single bootstrap simulation of type-r latent durations and consists of N observations $T_{1r}^{(b)}$, $T_{2r}^{(b)}$, ..., $T_{Nr}^{(b)}$.

3. For each bootstrap realization $(\mathbf{T}_r^{(b)})_{r=1}^R$, $b = 1, ..., B$, compute $r^{(b)} = (r_n^{(b)})_{n=1}^N$, where $r_n^{(b)} = \underset{r=1,....,R}{\arg\min} T_{nr}^{(b)}$.

4. Draw the bootstrap samples $\tilde{\mathbf{T}}^{(1)}, \tilde{\mathbf{T}}^{(2)}, ... \tilde{\mathbf{T}}^{(B)}$ of the realized durations. Each $\tilde{\mathbf{T}}^{(b)}$ consists of N observations $\tilde{T}_1^{(b)}, \tilde{T}_2^{(b)}, ..., \tilde{T}_N^{(b)}$ representing the draws of the latent durations $\tilde{T}_n^{(b)} = \tilde{T}_{nr}^{(b)}$ generated by a smoothed empirical distribution of the notional risk $r = r_n^{(b)}$ and the covariates $\mathbf{z}_n^{(b)}$.

3.6.2 Performance of Bootstrap Procedures

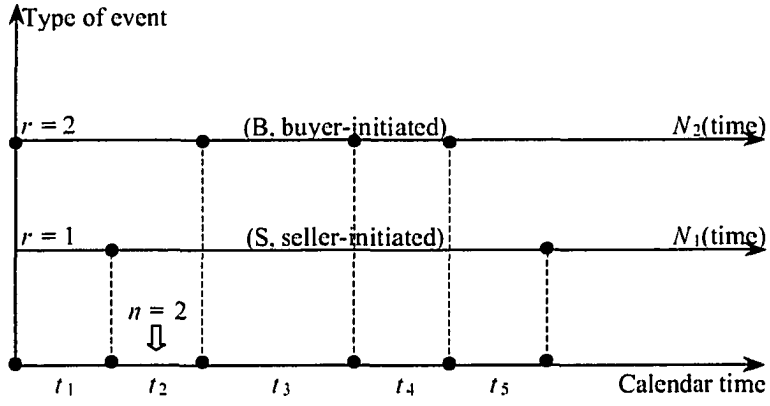
Burr [18] is the first systematic Monte Carlo study of the performance conducted for a number of bootstrap procedures and bootstrap confidence intervals in the Cox proportional hazard model. In the case of uninformative censoring Burr [18] finds that robust asymptotic confidence intervals generally perform better for the covariate effects, while the bootstrap methods provide better approximations to the confidence intervals for the survivor functions and the median survival times. The best procedure appears to depend on the design, but the model-based resampling generally outperforms the paired sampling. The bootstrap-t intervals were found to be consistently outperformed by one of the simpler bootstrap intervals, even though the bootstrap quantile intervals have better asymptotic convergence properties. Also, a non-trivial interaction effect between the method of drawing the sample and the method of forming confidence intervals was detected by Burr [18].

The performance of bootstrap procedures requires more thorough investigation in the competing risks environment, especially when the number of risk types R grows asymptotically with N at the slower rate. These issues will remain on the research agenda in the near future.

3.7 Conclusion

This chapter presents the review of asymptotic properties of the hazard function estimators in the framework of Cox proportional hazard model. Even though in practical applications to high-frequency data, such as the one presented in the next chapter, the asymptotic theory appears to work well, its relevance in applications to the samples of moderate size needs to be analyzed more carefully. Also, it is desirable to establish an automatic bandwidth selection procedure that works well for large samples as well as samples of moderate size. These issues are left for the future research and will not be pursued in this thesis.

3.8 Appendix: Figures and Table



Observe the sequence of events:

Types: $\{2, 1, 2, 2, 2, 1, \ldots\}$

$\{B, S, B, B, B, S, \ldots\}$

Durations: $\{t_1, t_2, t_3, t_4, t_5, \ldots\}$

Figure 3.1: Hypothetical event history in the case of two competing risks

The types B and S of two competing risks on Figure 3.1 are identified by the occurences of buyer- and seller-initiated events (quotes or transactions).

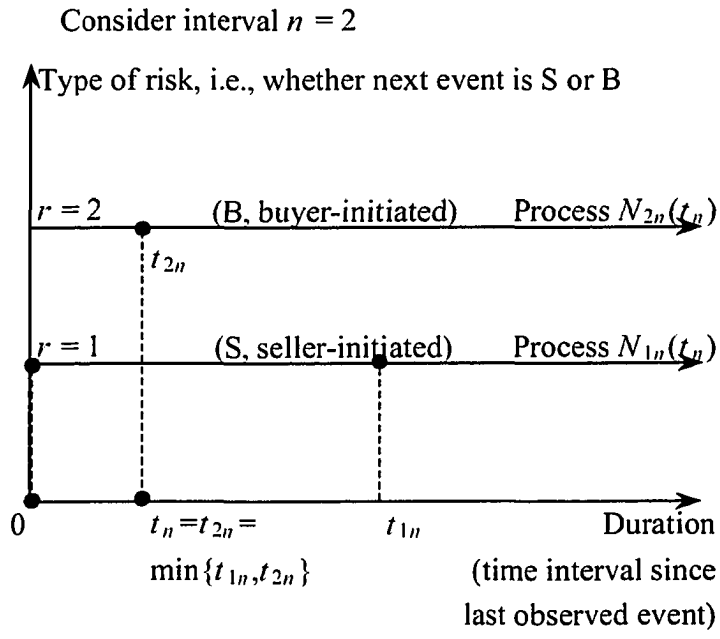


Figure 3.2: Stylized graphical presentation of the competition between two types of risk

Figure 3.2 displays in detail epoch $n = 2$, which is initiated by the event of type S and ends with an event of type B, as shown on Figure 3.1. The event of type S that could have occured t_{1n} seconds after the beginning of epoch $n = 2$ is censored by the early realization of the competing risk B at time $t_{2n} < t_{1n}$.

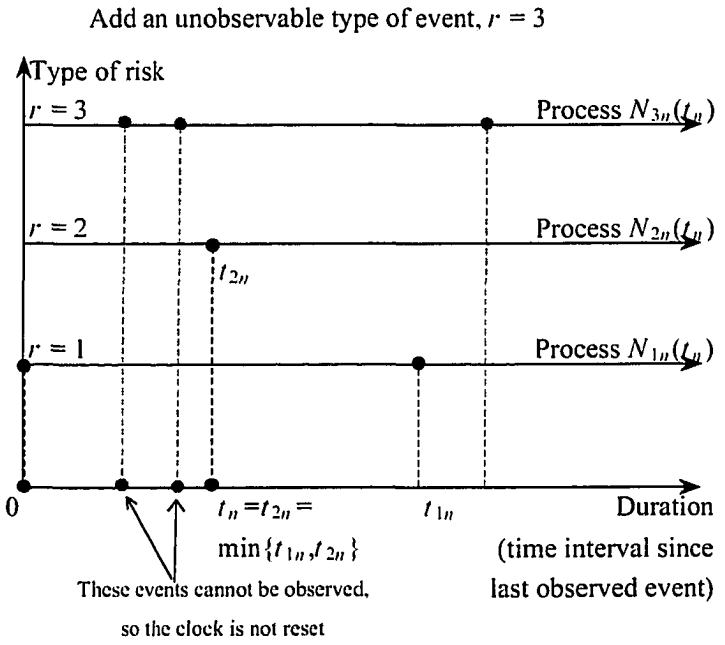


Figure 3.3: Stylized graphical presentation of the competition between two observable and one unobservable types of risk

Figure 3.3 displays epoch $n = 2$, which is initiated by the event of type S and ends with an event of type B, as shown on Figure 3.1. The event of type $r = 2$ marking the end of epoch is preceded by two occurrences of unobserved events of type $r = 3$, that remain undetected on the screens but might play important role for the future market dynamics.

Table 3.1: Selected seller-initiated events and graphs of their baseline hazard functions

Type of sell-side event	Graph and page number
A2: Sell market order	Fig.3.4, p.105
A3: Sell limit order at P_{bid}	Fig.3.5, p.106
A5: Sell limit order at $P_{\text{ask}} - 1$	Fig.3.6, p.106
A6: Sell limit order at P_{ask}	Fig.3.7, p.107
A7: Sell limit order at $P_{\text{ask}} + 1$	Fig.3.8, p.107
A9: Sell limit order at $P_{\text{ask}} + 3$	Fig.3.9, p.108
A12: Sell in $(P_{\text{ask}} + 5;\ P_{\text{ask}} + 10]$	Fig.3.10, p.108
AC6: Cancel at P_{ask}	Fig.3.11, p.109
AC7: Cancel at $P_{\text{ask}} + 1$	Fig.3.12, p.109
AC9: Cancel at $P_{\text{ask}} + 3$	Fig.3.13, p.110
AC12: Cancel in $(P_{\text{ask}} + 5;\ P_{\text{ask}} + 10]$	Fig.3.14, p.110

Figures 3.4–3.14 on the next six pages display the baseline hazard estimates $\widehat{h}_{0r}(t)$ (shown in a bold light pattern) for the Cox proportional hazard specification (3.4) with the full set of covariates from Table 4.2. and the baseline hazard estimates (shown in a bold grey pattern) for a smaller set of covariates selected as explained in section 3.5. The absolute values of asymptotic standard errors for the estimated hazard rates (shown by a shaded pattern for the full set of covariates and by a lighter curve above this pattern for the smaller set of covariates) are calculated from a k-NN version of formula for the asymptotic variance of external smoother in subsection 3.4.4.

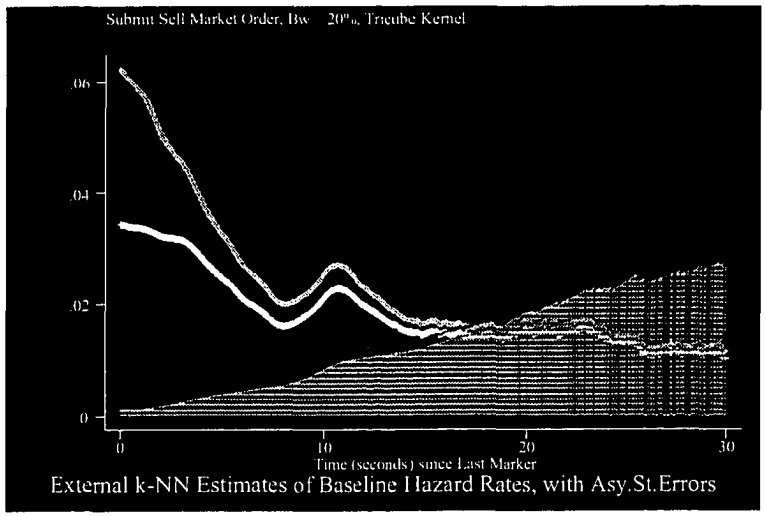


Figure 3.4: Baseline hazard rates for sell market order arrival events (A2)

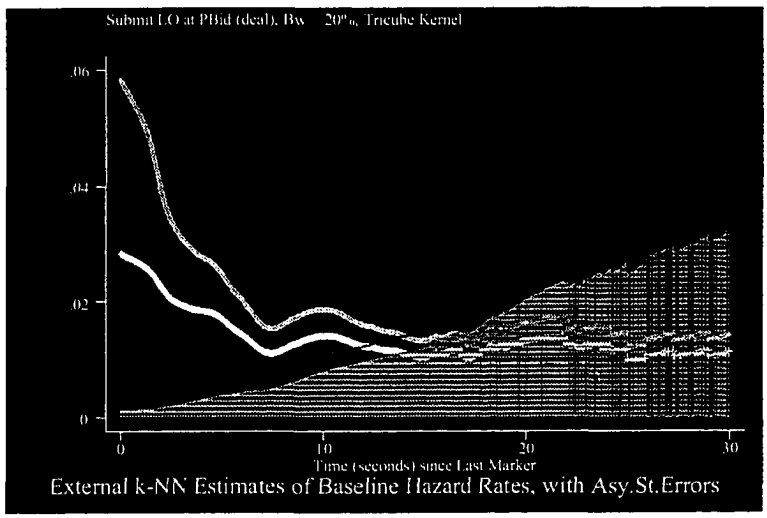


Figure 3.5: Baseline hazard rates for sell limit order arrival events at $P^* = P_{\text{bid}}$ (A3)

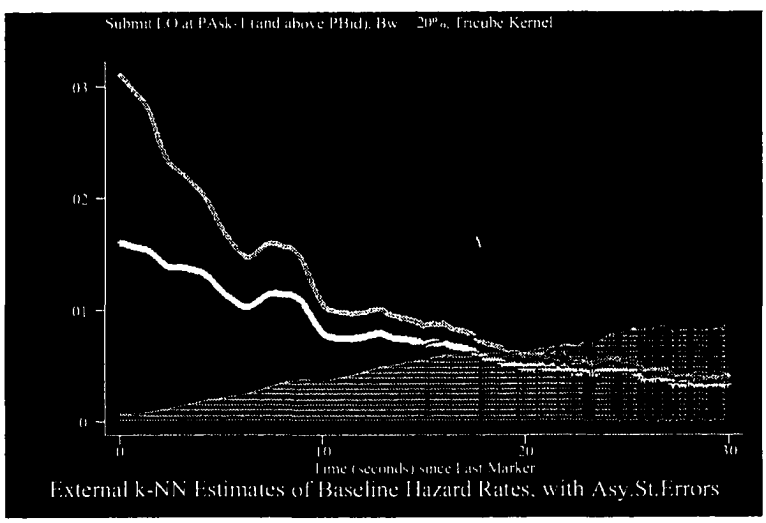


Figure 3.6: Baseline hazard rates for sell limit order arrival events at $P^* = P_{\text{ask}} - 1$ (A5)

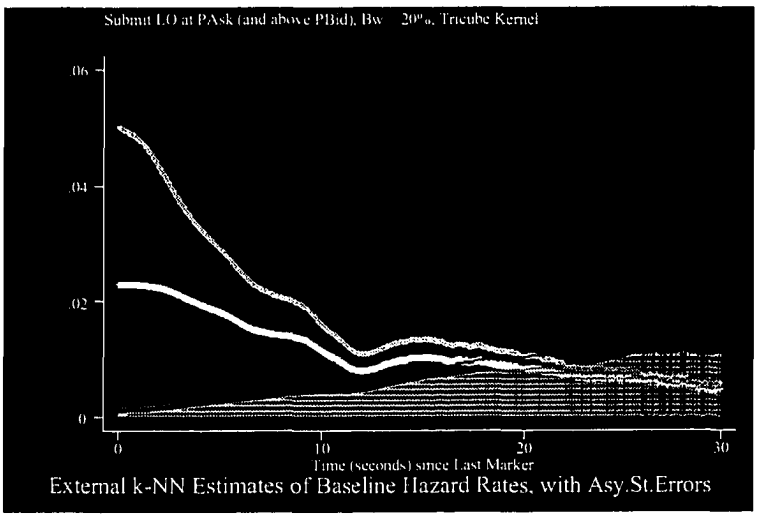


Figure 3.7: Baseline hazard rates for sell limit order arrival events at $P^* = P_{\text{ask}}$ (A6)

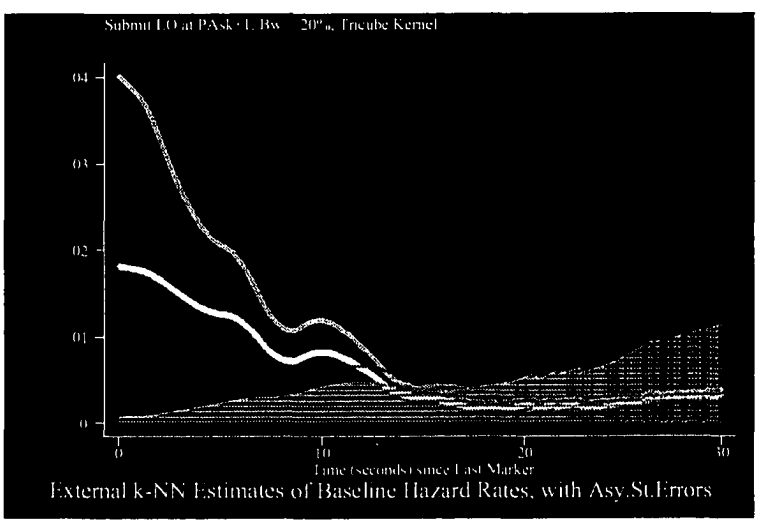


Figure 3.8: Baseline hazard rates for sell limit order arrival events at $P^* = P_{\text{ask}} + 1$ (A7)

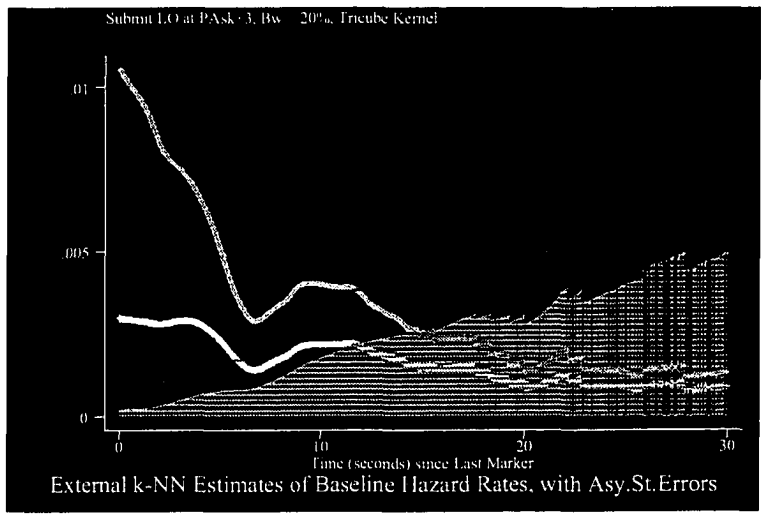


Figure 3.9: Baseline hazard rates for sell limit order arrival events at $P^* = P_{\text{ask}} + 3$ (A9)

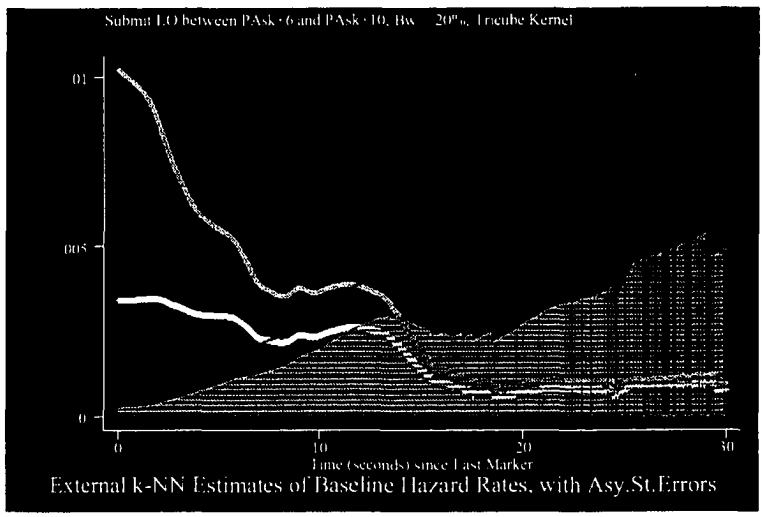


Figure 3.10: Baseline hazard rates for sell limit order arrivals between $P_{\text{ask}} + 6$ and $P_{\text{ask}} + 10$ (A12)

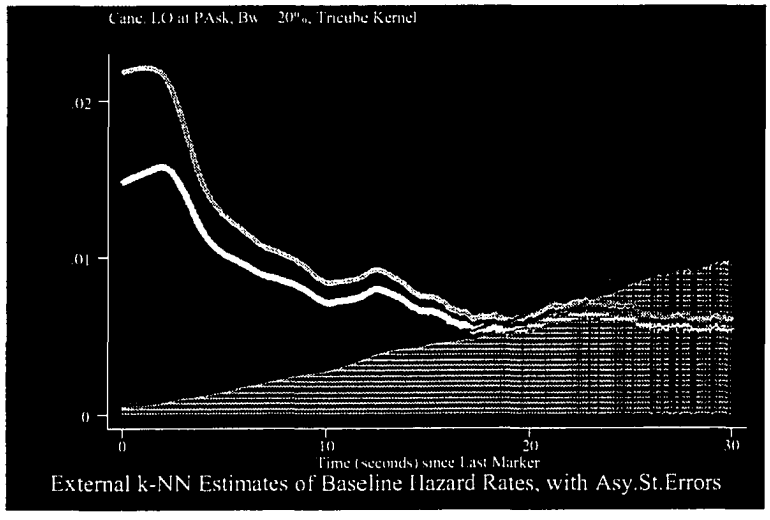


Figure 3.11: Baseline hazard rates for limit order cancellation events at $P^* = P_{\text{ask}}$ (AC6)

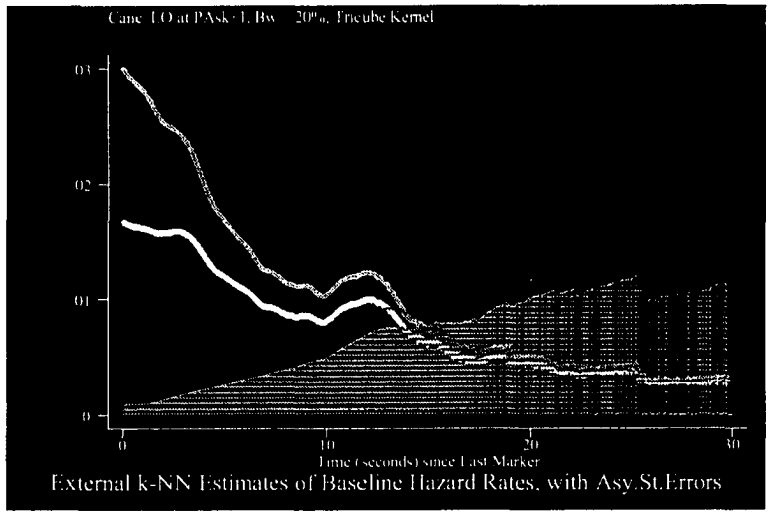


Figure 3.12: Baseline hazard rates for limit order cancellation events at $P^* = P_{\text{ask}} + 1$ (AC7)

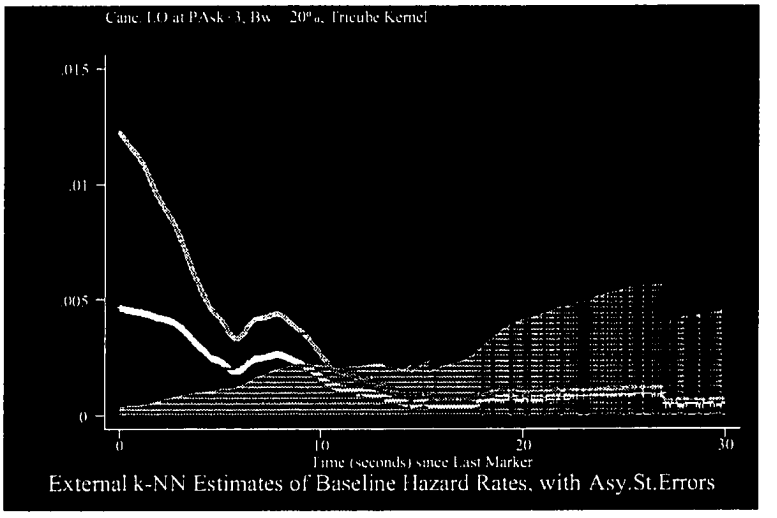


Figure 3.13: Baseline hazard rates for limit order cancellation events at $P^* = P_{\text{ask}} + 3$ (AC9)

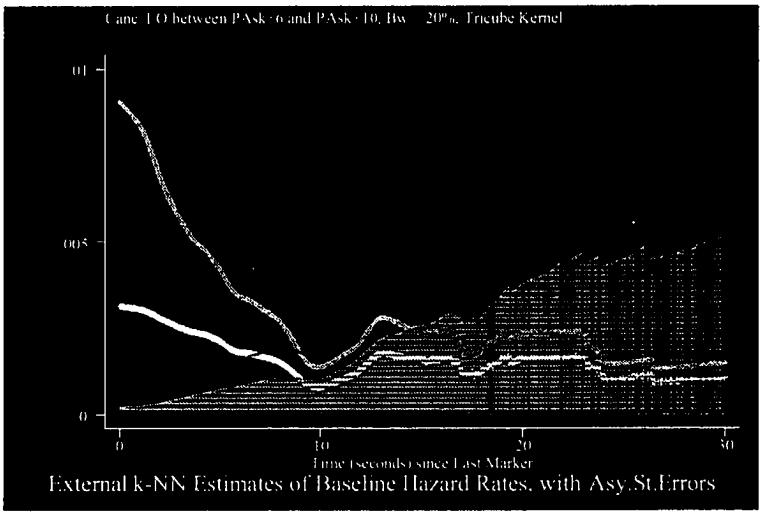


Figure 3.14: Baseline hazard rates for limit order cancellations between $P_{\text{ask}} + 6$ and $P_{\text{ask}} + 10$ (AC12)

Chapter 4

Endogenous Determination of Liquidity in Continuous Time: An Index-Based Model of Activity in the Electronic Limit Order Book

The effect of information on price determination and liquidity has a long history of research in market microstructure literature. In the foreign exchange market, the main source of differential information is the customer order flow. The dealer accepting a large customer order takes a long or short inventory position which is associated with the risk of depreciation or appreciation. If the information contained in the customer order is short-lived, the dealer would usually attempt to close his position fairly quickly (sometimes in a matter of minutes) by making a trade on the interdealer market, through traditional brokers, or via an electronic brokerage system. However, to benefit from his private information, the dealer may split the order and trade its portions gradually over time using market orders. Alternatively, the dealer may attempt to receive a better price by placing one or several limit orders and being prepared to deal with the risk of uncertain execution. Finally, dumping a large quantity

on the market by placing the large limit order entails the risk of getting adverse price in exchange for the chance of speedy execution.

Three measures are commonly used in the market microstructure literature to characterize liquidity: (1) the *bid-ask spread*, defined as the difference between the lowest ask and the highest bid prices currently available on the market; (2) the *market depth* on the ask and bid sides, usually defined by the quantities available for immediate purchase or sale at the best ask and bid quotes; and (3) the *time* elapsing before the trade can be performed at the "favorable" price. If arrival of a large order signals high probability of private information on the market, then a large buy limit order would lead to a temporary loss of liquidity on the ask side resulting in higher cancellation probabilities for the existing sell orders, and the larger spreads that would not be restored until the ask price moves up to a higher level. However, since dealers can rarely be certain that large customer orders come from informed customers and contain private information not incorporated in the market price, it can also be argued that large orders originate from uninformed liquidity traders. The possible argument can be based on the observation that uninformed traders have more incentives to place large orders thereby increasing the competition and tightening the bid-ask spread. The dataset on quotes and transactions used in this chapter can shed some light on this issue and discriminate empirically between these two hypotheses.

The time dimension brings in additional flavor to this dichotomy. In early microstructure literature the evolution of market was studied in activity time, and the duration between consecutive quotes or transactions was usually dismissed as uninformative, or treated as an additional exogenous explanatory variable. More recently it has been argued that the lack of activity and absence of transactions can represent the events that could be used to predict market dynamics, at least in the near future. Among the first attempts to investigate analytically the trade-off between executing a trade using market versus limit orders analytically stands out the paper by Parlour [112] who models the behavior of an agent submitting a limit or market order depending on the state of the limit order book. The agent placing an order in Parlour's model evaluates the trade-off between the market impact of her trading

strategy and the probability of her own order execution.

Empirical research on the limit order markets has long been hampered by the lack of detailed order-level data. Until recently, most order-level data came from the stock exchanges organized as electronic limit order books. Examples are Biais, Hillion, and Spatt [12], Hamao and Hasbrouck [59], Harris and Hasbrouck [65], Hollifield, Miller, and Sandas [74], who investigate empirical properties of limit order markets in Paris, Tokyo, New York, and Stockholm. Even though there are now several electronic systems trading currencies, corporate and government bonds, and other financial instruments, the information disclosure restrictions on the providers of such systems in the foreign exchange markets usually make the detailed order-level data unavailable for academic researchers.

The goal of the present chapter is to enhance our understanding of the short-run liquidity in the order driven segment of the foreign exchange market. It presents an empirical study of interaction between the trade flow, market bid-ask spread, and depth on the bid and ask sides of Reuters D2000-2, which was one of the major liquidity providers on the foreign exchange market in the nineties, and investigates the roles played by these variables in the process of price formation. The competing risks model is extended and applied to obtain a decomposition the order flow into the components identified by direction and aggressiveness of trading activity. While the decomposition of the order flow is determined endogenously from the past trading activity, part of the information hidden in the baseline hazard rates of competing risks can be extracted by application of the econometric techniques reviewed in Chapter 3. The signed order flow defined in this dissertation as the number of buyer-initiated transactions less the number of seller-initiated transactions serves as the major vehicle of information discovery (but not necessarily the ultimate source of information).

Since the data set used in this chapter lacks information on traders' identities, the empirical model does not tackle explicitly the agents' trading strategies.[1] Instead, the model can be viewed as a comprehensive description of partially disaggregated trading activity,

[1]The data set used in the empirical analysis does not allow explicit replication of traders' behavior since it does not contain any information on the identities of market participants.

where the level of disaggregation is determined by the data. It represents an agnostic behavioral framework for learning about empirical regularities in the markets populated by multiple interacting agents who may possess heterogeneous information and employ complex and diverse trading strategies. It will be shown how the model can be used to generate probability forecasts of short-term market dynamics that might facilitate the development of dynamic trading strategies by dealers and their sponsors.

This chapter proceeds by the following plan. Section 4.1 presents a stylized model of trading activity in the idealized limit order market reminiscent of the market studied in the empirical application. Then it provides motivation and poses several empirical questions that will be investigated in the subsequent sections. Section 4.2 expands the covariate structure of the basic competing risks model from Chapter 2 that incorporates a broad range of observable limit order book characteristics. The estimated covariate indices capture a large portion of cross-sectional and serial correlation between the various types of the limit order book trading activity. Section 4.3 discusses the estimates of the covariate effects drawing special attention to the publicly observable limit order book information and the lagged activity measures such as the past order flow and the cumulative counts of recent transactions. Application of the principal component analysis (PCA) to the covariate indices in section 4.4 identifies five pervasive factors capturing 85% of the observed activity and leading to substantial data compression. Each of the first five pervasive PCA factors has distinct characteristics that facilitate their interpretation as alternative measures of aggregate market activity. Section 4.5 outlines the results of simulation experiments and cast some light on the strengths and deficiencies of the modeling methodology. Finally, section 4.6 reviews the probability forecasting formulas and applies some of the graphical tools to evaluation of goodness-of-fit and out-of-sample performance of the index-based model of limit order book activity. It is found that the model has good predictive power, at least relative to popular simple moving average-type forecasting rules. In summary, the competing risks methodology is found to be a valuable tool for short-term forecasting of market activity and its outcomes, and for understanding the behavior of heterogeneous agents in a competitive

market environment.

4.1 The Anatomy of a Pure Limit Order Market

To motivate the modeling strategy and provide the background for the empirical results of this chapter, we first give a stylized description of the automated limit order book. While technical details and the peculiarities of actual limit order markets must always be accounted for in specific applications, our presentation here is deliberately simplified. Even though it has been developed bearing in mind an application to the specific limit order market in section 4.2, in this informal review we draw reader's attention to the salient features shared by all markets organized as pure electronic limit order books.

4.1.1 Stylized Description of Electronic Limit Order Book

In the forthcoming development, the *limit order* is defined as an instruction to sell (or buy) a certain number of units of financial asset at a certain price, which is called the *limit order price*. We define *market order* as an instruction to sell (or buy) a certain number of units of financial asset immediately at the best available market price. While market orders always face full and immediate execution, the limit orders may face only partial execution, or may not be filled at all, in which case the untraded portion of the limit order is placed into one of the two queues (separate for sell and for buy limit orders) and kept in the queue until the order is explicitly cancelled by the trader or hit by the arriving market or limit order.

A bank dealer can enter a buy or sell limit order into the system at any moment of time, indicating the limit order price and the quantity of foreign exchange (usually an integer number, in millions of US dollars) that he wants to trade. After an attempt to match the incoming order with outstanding orders submitted by other traders on the opposite side of the limit order book, the new order is entered into the system. Additionally, traders have an option to submit a buy or sell market order indicating the price and the quantity. In this

case, after an attempt to match the incoming market order with outstanding limit orders at the price *equal* to the incoming market order, an unexecuted portion of the market order is cancelled automatically and a confirmation message to the trader is sent. The system allows traders to monitor their limit orders that can be removed by a hit of a "Cancel" key at any moment of time.

The market structure which consists of the two queues for buy and sell limit orders, along with the specified trading protocol describing the priority of limit order execution is called the *limit order book*. The queues of sell and buy limit orders can be interpreted as approximate representations of excess supply and demand curves for the traded financial asset (Figure 4.1). The priority of limit orders in the book depends on the details and the trading protocol of the particular system. All orders are submitted anonymously as the identity of a foreign exchange trader is considered a strictly confidential information and never disclosed.

Only a handful of real financial markets are organized as pure electronic limit order books, such as the one described above. In practice virtually any real limit order market represents a hybrid system, which would be hard to analyze using the simple model presented in this chapter. For example, the trader submitting a market order in Reuters D2000-2 must provide not only the quantity but also the price, which does not have to be (but usually is) the best market price available on the buy or sell side of the market at the time of submission. After the market order is submitted it is matched only with the limit orders submitted at the prices *equal* to the arriving market order price. The unfilled portion of the market order is cancelled automatically. Unlike the market orders, aggressive limit orders submitted at the prices that are different from the current best price can obtain price improvement, as the outstanding limit orders in the book submitted at the better prices receive priority in execution. The traders demanding early execution may be willing to submit a market order and sacrifice the difference between the best bid and ask quotes (the *bid-ask spread*), or submit a limit order which gives them a chance of getting a better price at the expense of execution uncertainty. The seller who is willing to receive a better execution price may

submit a limit order with a relatively high price and be prepared to wait longer and deal with the risk of execution uncertainty in the event that the market price moves in the opposite direction, and with the risk of being "picked off" in the event of a sudden increase in the market price beyond the level considered reasonable by this limit order trader.

The dynamics of a limit order market are illustrated on a series of graphs (Figures 4.2–4.8). Figure 4.2 provides graphical representation of the situation when a limit order trader submits a bid to buy one million dollars at the price 1.7510 DEM per US dollar, which exceeds by 0.01 Pfennig the previous best market price 1.7509 DEM, but still 0.02 Pfennig short of the price that guarantees immediate execution. Alternatively, the trader can submit a less aggressive limit buy order at the current best market price (Figure 4.3). In such event the arriving limit order receives a lower priority in comparison to the limit order that has been previously submitted to the limit order book at the same price. In the first case the traders watching monitors of the trading screens notice immediately the increase in the best bid market price, while in the second case they will only observe the increase in the quantity available at the unchanged best bid market price. Whether and how this public information arrival affects the behavior of other traders remains one of the central issues in the empirical and theoretical market microstructure literature.

The limit order buyer has yet another option. He can avoid revealing any information to the market about his intentions and willingness to trade if he submits a *subsidiary limit order*, which can be an order to buy at the price just below the current bid market price (Figure 4.4) or two ticks (minimal increments) below the current bid market price (Figure 4.5). In both cases the best market prices and quantities available to sellers and buyers at these prices remain unchanged. Therefore, the information on the screens is not affected by arrivals of subsidiary limit orders.

The last three diagrams provide illustrations of the effects of a subsidiary limit order cancellation (Figure 4.6), a market order-initiated transaction followed by immediate cancellation of the portion of order which cannot be matched at the best sell market price (Figure 4.7), and a similar transaction initiated by an aggressive limit order buyer (Figure

4.8). Note the difference between the effect of market versus aggressive limit order arrival on the market liquidity. While the market order buyer only consumes the liquidity available at the best market sell price of DEM 1.7512 per dollar, the limit order buyer also provides liquidity at this price, once it becomes clear that the quantity available for sale at this price is insufficient to satisfy his demand for liquidity. Since the information content and the role played by aggressive limit orders in liquidity provision can be different from the role played by market orders, it is important to differentiate between these two types of events as they might be used by traders possessing distinct information or having different beliefs and risk attitude.

In the liquid financial markets such as the segment of foreign exchange market studied in this work, interactions between the limit order arrivals, cancellations, and transactions, similar to those described above, occur virtually every second. Figure 4.9 shows a representative subsample of the continuous time, which was the result of such interaction. The time period on Figure 4.9 covers ten minutes of moderately active morning trading on Monday, October 6, 1997. The light solid curves show the time series plots of best bid and ask prices available on the market at any moment of time. The white crosses and circles illustrate, respectively, the times and the prices of buyer- and seller-initiated transactions. The chart on Figure 4.9 provides an illustration of the empirical fact that much of the activity in the foreign exchange market is not accompanied by any trades, but rather represents the reaction of traders submitting and cancelling limit orders to prior market events. The analysis of dynamic interaction of limit order traders using the screen information on the limit order book will be the main object of investigation in the remainder of this dissertation.

4.1.2 Motivation and Limits of the Econometric Approach

The microstructure approach to foreign exchange determination[2] postulates that the order flow (broadly defined as the sequence of buyer- and seller-initiated trades) is the only variable

[2]This approach was popularized by Lyons [96]: see also a brief review of the key ideas in Chapter 1 of this dissertation, and the policy implications in Lyons [97].

explaining the long-range dynamics of the foreign exchange. The microstructure approach, which can be justified theoretically by the ability of the order flow to capture the prevailing motives of the buyers and sellers who form their beliefs and exploit information on the future price dynamics, has found a solid support in the recent empirical literature uncovering the ability of the order flow to explain a large share of movements in the major floating exchange rates[3]. It is commonly assumed that the beliefs of market participants formed on the basis of differential information find their outlet in the form of order flow. Even though the order flow ultimately impounds the traders' interpretation of private and public information about the fundamentals, we take a stand that in the short run the order flow represents the dominant mechanism transforming the beliefs of traders into the dynamic pattern of buyer- and seller-initiated transaction prices. However, for a number of reasons, no attempt to model this mechanism explicitly will be made in this work.

Since this chapter does not attempt to capture dynamic effects of the lags longer than 15 minutes, the key order flow-related variables driving the dynamics of notional hazard rates will be the counts of buyer- and seller initiated trades in the five-second periods immediately preceding the sub-epoch, and the similar counts from the earlier history. We deliberately bypass macroeconomic fundamentals as there is remarkably little evidence that macroeconomic variables have any consistent effect at the ultra-short horizons which are the focus of this study.[4] Similarly, we restrict the role of the past prices emphasized in the literature on technical analysis of foreign exchange markets, to the price differences between the last and the second-to-last trades that occur prior to the beginning of the analyzed epoch.

It is important to stress that no *a priori* assumptions will be made about the valuation of foreign exchange. In our opinion, the ultimate value of foreign exchange is impossible to define without the reference to the "efficient price" and strong assumptions about its time series properties. In the financial markets literature it is often assumed that the "efficient

[3]See, for example, Evans and Lyons [43].

[4]The interest rate news announcements appear to be the only exception but even those take usually several hours to be absorbed in the market price.

price" is a continuous martingale driven by a Brownian motion with constant volatility.[5] This assumption is very intuitive and appears well justified when observation intervals are monthly, weekly, or even, with some reservations, daily, but appears to be at odds with the main objective of market microstructure literature which aims to study the process of price formation.[6] It is especially important to avoid making such a strong assumption at the high and ultra-high frequencies when even the very notion of "efficient price" becomes less transparent because of the price discreteness, illiquidity, bid-ask bounce, asynchronous trading,[7] and other effects which are at the center of market microstructure research agenda. The broadly recognized and accepted empirical fact that even the most sophisticated agents may have differential beliefs makes virtually any *a priori* assumption about the underlying fundamental value difficult to defend. However, the agents' information and beliefs, which are intrinsically unobservable, cannot be structured without making additional identifying assumptions about the process of information discovery.

In our agnostic approach, we focus attention on the hazard rates of alternative events and formulate the model using only a handful of observable characteristics that can be inferred from a subset of publicly available data. Thus, we postulate that the limit order book is the only medium for the price dynamics. No matter what foreign exchange rates – bid, ask, or the actual transaction prices – are considered, their dynamics are thought to be driven by interaction of supply and demand of multiple agents with differential information, horizons, beliefs, and trading strategies.

As the exact moments of information arrivals are unavailable from the data, and rarely can be identified in practice, difficulties arise in determining a good proxy for information. Trading volume has been by far the most frequently used proxy in the empirical studies of

[5] In the modern literature, more realistic assumptions about dynamic properties of the volatility and the drift term are usually made.

[6] Hasbrouck [67] contains a nice discussion of the role of the "efficient price" assumption in the market microstructure literature.

[7] Chapter 3 in Campbell, Lo, and MacKinlay [19] summarizes the early econometric studies that made attempts to resolve these problems.

stock returns. However, Jones, Kaul, and Lipson [79] in a comparative study of different information proxies demonstrate that trading volume has the same informational content as the number of trades. Similarly, Marsh and Rock [99] show that the net order flow (the number of seller-initiated minus the number of buyer-initiated transactions) explains as much of the price variation as does the signed volume of trade. Geman and Ané [48] show that the moments of increments of the time deformation process that makes returns on a stock market index normal closely match the moments of the number of trades for that index per minute. All this and other evidence indicates that the signed number of trades could be a better proxy for the information arrival than trading volume and may be an important factor behind the market volatility, and its persistence. Moreover, the occurrence and direction of trades are readily observable on Reuters D2000-2 trading screens, to the extent that a trader can distinguish between the flashes on the screen that accompany any new transaction that occur in the system. In view of this positive evidence and to the extent that other information proxies such as the number of quote changes, price changes, and so on, used in the empirical literature produced mixed results, we take the general premise of the order flow approach to exchange rate determination (Lyons [96]) and accept the signed number of trades as a primary vehicle behind the information arrival process.

4.1.3 Empirical Questions

In this chapter the following issues will be investigated:

1. dynamic links between the order flow and various types of market activity;
2. the price formation mechanism in electronic limit order markets and the role of order flow components, in particular, passive versus aggressive order flow;
3. the out-of-sample forecasting power of the order flow and other variables associated with limit order book trading.

To keep this work focused on a relatively limited set of issues and avoid additional methodological and practical complications, we do not model seasonal effects and clustering of orders at the round numbers (multiples of five ticks, in the case of DEM/USD exchange rate). The intraday seasonality is undoubtedly an important empirical feature of virtually any financial data. Even though, in principle, the mechanical introduction of time-of-the-day diurnal effects or a simple deterministic trend in the specification of hazard rates may lead to a slightly improved fit of the model, this will involve additional methodological and practical problems. A substantial part of the detected seasonality is likely to be spurious, since the trading history covered by our data set includes only five full trading days. In turn, this may lead to incorrect inferences about the effects of other covariates, which are the main focus of the present chapter.

We also refrain from modeling explicitly the empirically relevant features of real trading process such as random communication delays and failures, lack of mutual credit agreement among counterparties, occasional violations of order priority, potential implications of the complex architecture of communication networks, and so on. For most of this study we also do not distinguish between the market information and individual trader information which may be displayed in different sections of the Reuters D2000-2 trading screen, or come from alternative sources. Clearly, ignoring the peculiarities of actual trading process may play a crucial role for the success or failure of model's predictive performance and for the relevance of simulated trading histories to real market data. However, at this point we accept the lack of realistic representation of some aspects of actual limit order trading as a price to pay for the relative simplicity, analytical tractability, and methodological generality of the competing risks specification developed in this chapter.

4.2 Empirical Analysis of Liquidity and Order Flow in the Electronic Forex Limit Order Market

4.2.1 Statistical Methodology

Since the key methodological ideas behind the competing risks specification have been described at length in Chapters 2 and 3, in this section we pay special attention to the details that are not discussed in the previous chapters. As usual in the competing risks framework, the choice of the Markov space depends on the range of prices, quantities, and other market characteristics, whose effect on the point processes will be investigated. Even though the specification of the Markov space is ultimately determined by the objectives of study and often is severely restricted because of the data limitations, it is desirable in the loosely structured problems like this to start with as broad set of variables as could be reasonably possible. Among the dimensions of the Markov space should be the variables that might be linked to the market factors identified from theoretical considerations. Since the variables in the Markov space may be discrete or continuous, it may have sometimes a fairly complicated topological structure.[8]

The full covariate vector $\mathbf{z}$ selected for the analysis in this chapter is described in the next subsection (Table 4.2). Along with the "usual suspects" such as the bid-ask spread and the depth of limit order book on the bid and ask sides, components of the covariate vector $\mathbf{z}$ represent deeper dynamic characteristics of market liquidity, including several types of price and quantity changes, the side of the book where the recent transactions occur, and so on. No doubt, this set of covariate could be modified if the alternative data on traders' identity or relevant information from other markets had been available.

[8]For instance, if the purpose of research is to study the effect of order clustering at the multiples of five ticks, the most natural specification of the study window would distinguish the range of prices such as multiples of 5 ticks, one plus multiples of 5 ticks, two plus multiples of 5 ticks, and so on, like it was done in Osler [111].

The classification and identifying properties of event types will also be reviewed in the next subsection (Table 4.1). Each of the R identified competing risks can be associated with occurrence of alternative limit order book events. The hazard rates of these risks are modelled using the Cox proportional hazard (CPH) specification

$$h_r(t|\mathbf{z}) = h_{0r}(t)\exp(\mathbf{z}'\boldsymbol{\beta}_r), \quad r = 1, ..., R, \tag{4.1}$$

where $h_{0r}(t) = h_r(t|\mathbf{0})$ is the baseline hazard function, t is the time at risk, and $\mathbf{z}$ is the covariate vector. Covariates in proportional hazard models always act multiplicatively on the hazard rate of the specified type of event. The model (4.1) can be rewritten in the integrated form as follows

$$H_r(t|\mathbf{z}) = \int_0^t h_r(u|\mathbf{z})du = \int_0^t h_{0r}(u)du\exp(\mathbf{z}'\boldsymbol{\beta}_r) = H_{0r}(t)\exp(\mathbf{z}'\boldsymbol{\beta}_r),$$

where $H_r(t|\mathbf{z})$ is the cumulative hazard function of risk r. The third equivalent representation of the model (4.1) is based on the expression for the survival function

$$S_r(t|\mathbf{z}) = \exp(-H_r(t|\mathbf{z})) = \exp(-H_{0r}(t)\exp(\mathbf{z}'\boldsymbol{\beta}_r)) = (S_{0r}(t))^{\exp(\mathbf{z}'\boldsymbol{\beta}_r)},$$

where $S_{0r}(t) = S_r(t|\mathbf{0}) = \exp(-H_{0r}(t))$ is the baseline survivor function of risk r.

4.2.2 Description of Event Types and Covariates

The choice of the event types in the competing risks approach that has been briefly discussed in section 4.1 is somewhat arbitrary and depends on the questions to be answered, data limitations, and the prior theoretical considerations about the data generating process. The richness and limitations of our data allow specification of $R = 46$ types of buyer- and seller-initiated events and selection of $S = 14$ "observable" types of events that can be identified on the trading screens by all traders. Table 4.1 provides the summary of the sell side events introduced in this chapter. The rows of Table 4.1 describe 23 types of events triggered by sellers' actions. The type of event is defined as a combination of order type (market or limit order), character of activity (submission or cancellation of limit order), and the distance

between the price P^* of the limit order and the prevailing best bid and ask quotes P_{bid} and P_{ask} prior to the event (column 2). The consequences for the best bid and ask prices and for the liquidity of the limit order book (market depth) at these prices are indicated in columns 3 and 4 of Table 4.1. The events associated with arrivals of market and limit *sell* orders are denoted by letter "A" (which means the activity occurs on the *ask side* of the limit order book) followed by a numerical index corresponding to the sell order aggressiveness. Similarly, the cancellations of sell limit orders are denoted by "AC" followed by a numerical index that depends on the distance between the limit order price and the best market ask quote. The events marked in the first column of the table by single and double stars can be observed by all market participants. The events marked by double stars, which are also observable by all market participants, typically trigger immediate trade executions.[9] The unmarked types of events are associated with limit order arrivals or cancellations at suboptimal prices that cannot be observed on the D2000-2 trading screens and therefore constitute private information of traders.

The buy side events (which occur on the *bid side* of the book) are denoted by letters "B" and "BC" followed by numerical indices. The definitions of the buy side events are similar to the sell side events.

The events associated with changes of subsidiary quotes and the quantities available at these quotes are not included into the public information domain, even though some of these events can be potentially observed by market participants. Therefore, we follow the general logic of the approach developed in Chapter 3 according to which the subsidiary events do not restart the "internal clock" of the "race" between competing risks and assume that all types of events except A1 through A6, AC6, B1 through B6, and BC6 are unobservable.

The number and types of events are chosen to allow a fine partition across the types of events to capture heterogeneity of traders' behavior on the one hand, and to provide a sufficient number of markers for flexible estimation of a sufficiently rich covariate structure

[9]Note that occasionally the trades will not be executed automatically following order crossings because of the lack of mutual credit among the counterparties, communication delays, etc.

on the other hand. Given the data limitations, for each type r of the competing risks the components of vector $\mathbf{z}_n$ $(n = 1, ..., N_r)$ are chosen to be associated with the market conditions at the time of event and the recent trading history prior to the event as shown in Table 4.2.

The components of the vector $\mathbf{z}_n$ of covariates (explanatory variables) incorporate the pieces of information on the screen (or information that can be easily inferred from the quotes on the screen) that might be closely monitored by traders (c.f. Figure 2.1). In view of the analysis of the previous sections, the covariates include variables characterizing the recent quoted price dynamics, market liquidity, and the general level and momentum of market activity, including directional characteristics of recent transactions.[10] The components of the vector $\mathbf{z}_n$ of covariates are divided into the three categories.

A. Price covariates

- *Slippage* $\equiv$ difference between the current midquote $\frac{1}{2}(P_{\text{bid}} + P_{\text{ask}})$ and the last transaction price (measured in 0.0001 DEM);
- $Spread_{>0} \equiv$ quoted size of the bid-ask spread (measured in 0.0001 DEM) when it is positive, zero otherwise;
- $\Delta P_{\text{ask}} = P_{\text{ask}} - P_{\text{ask},\text{-}1} \equiv$ change of the best ask price (in 0.0001 DEM) between the last and second-to-last observable events;
- $\Delta P_{\text{ask},\text{-}1} = P_{\text{ask},\text{-}1} - P_{\text{ask},\text{-}2} \equiv$ change of the best ask price (in 0.0001 DEM) between the second-to-last and third-to-last observable events;
- $\Delta P_{\text{bid}} = P_{\text{bid}} - P_{\text{bid},\text{-}1} \equiv$ change of the best bid price (in 0.0001 DEM) between the last and second-to-last observable events;

[10] The determinants of limit order submission strategies are briefly discussed in O'Hara [109] and Goodhart and O'Hara [52]. The main theoretical contributions on the topic are Chakravarty and Holden [20], and Parlour [112]. Biais et al. [12], Handa and Schwartz [61], Harris and Hasbrouck [65], and Brown et al. [17] contain interesting empirical results.

- $\Delta P_{\text{bid},-1} = P_{\text{bid},-1} - P_{\text{bid},-2} \equiv$ change of the best bid price (in 0.0001 DEM) between the second-to-last and third-to-last observable events.

The price covariates characterize the short-term dynamics of the best bid and ask quotes and their interactions with the most recent transaction prices. These covariates accommodate short-term deviations of the quoted bid and ask quotes from the long-run equilibria as well as the potential errors that could be committed in the reconstruction of the trading history. For instance, the *Slippage* variable defined as the shift in the market price given by the midpoint of bid-ask spread relative to the last transaction price may be interpreted as the midquote positioning bias. One can think of it as a profit accrued to the trader participating in the last transaction if she liquidates her last trade position at the midpoint of the current bid-ask spread.[11] Since the absolute value of positioning bias is expected to be larger during the periods of changes in the bid and ask quotes without transaction activity, the *Slippage* variable can capture the tradeless price discovery mechanism that might prevail around the public news announcement. The size of the market bid-ask spread is often associated with the intuitive notion of illiquidity in the market microstructure literature, and is expected to have a strong impact on the types of submitted orders, as has been emphasized in the empirical microstructure literature.

B. Depth covariates

- $\log(Q_{\text{ask}}) \equiv$ natural logarithm of the market depth quoted on the ask side, i.e., the value of currency (expressed in \$ mln.) available at the best ask price;[12]

[11] This interpretation of positioning bias disregards the transaction cost, which is always incurred by aggressor (the counterparty initiating the trade) according to the trading protocol of the D2000-2 trading system.

[12] During the time period before 1998, the exact value of market depth on the ask and bid sides in the Reuters trading system was unobservable to market participants when it was in double digits (\$10 mln. or larger). Traders could see only the "R" indicator in the depth part of the screen. Therefore, we set $Q_{\text{ask}} = 10$ and $Q_{\text{bid}} = 10$ every time when the actual market depth is at least 10 million US dollars.

- $Q^{+}_{\text{ask}} \equiv$ indicator of large depth on the ask side, equals to unity if and only if the ask market depth is at least \$10 mln.;

- $\Delta \log(Q_{\text{ask}}) = \log(Q_{\text{ask}}) - \log(Q_{\text{ask},-1}) \equiv$ last change of logarithm of the market depth quoted on the ask side if the best ask price did not change between the last and second-to-last observable events, zero otherwise;

- $\Delta \log(Q_{\text{ask},-1}) = \log(Q_{\text{ask},-1}) - \log(Q_{\text{ask},-2}) \equiv$ second-to-last change of logarithm of the market depth quoted on the ask side if the best ask price did not change between the last and third-to-last observable events, zero otherwise;

- $\log(Q_{\text{bid}}) \equiv$ natural logarithm of the market depth quoted on the bid side, i.e., the value of currency (expressed in \$ mln.) available at the best bid price;

- $Q^{+}_{\text{bid}} \equiv$ indicator of large depth on the bid side, equals to unity if and only if the bid market depth is at least \$10 mln.;

- $\Delta \log(Q_{\text{bid}}) = \log(Q_{\text{bid}}) - \log(Q_{\text{bid},-1}) \equiv$ second-to-last change of logarithm of the market depth quoted on the bid side if the best bid price did not change between the last and third-to-last observable events, zero otherwise;

- $\Delta \log(Q_{\text{bid},-1}) = \log(Q_{\text{bid},-1}) - \log(Q_{\text{bid},-2}) \equiv$ second-to-last change of logarithm of the market depth quoted on the bid side if the best bid price did not change between the last and third-to-last observable events, zero otherwise.

The two market depth variables represent the second dimension of liquidity identified in the introduction to this chapter, specifically, how many units of asset can be bought (or sold) at the current ask (or bid) market prices. The depth covariates are also expected to be significant for the risks of cancellations since the likelihood of a cancellation event is expected to be positively related to the total number of active limit orders, and the latter number is correlated with the quoted depth at the best market price. Similarly to the quoted prices, the changes of quoted quantities capture the more subtle traders' reaction to changes

in the publicly available information on the limit order book and private information from the customer orders. The indicators of large depth Q^{+}_{ask} and Q^{+}_{bid} accommodate potential nonlinearities in the dependence of market activity on the depth variables (which might be partially justified by unobservable exact levels of depth when the market depth exceeds \$10 mln.)[13]

C. Recent trade activity and order flow covariates

- *Side* $\equiv$ directional indicator of the last transaction (+1 for seller-initiated trades, −1 for buyer-initiated trades);
- $Side_{-1} \equiv$ directional indicator of the second-to-last transaction (+1 for seller-initiated trades, −1 for buyer-initiated trades);
- $F_{0-5''} \equiv$ the signed order flow (measured as the difference between the number of seller- and buyer-initiated transactions) in the five-second period prior to the last observable event;
- $F_{5-10''}$, $F_{10-15''}$, $F_{15-30''}$, $F_{30-60''}$, $F_{1-2'}$, $F_{2-5'}$ and $F_{5-15'}$ are similarly defined as the signed order flow over the time periods five to ten seconds, ten to 15 seconds, and so on, prior to the last observable event;
- $T_{0-5''} \equiv$ the trade (measured as the total number of transactions) in the five-second period prior to the last observable event;
- $T_{5-10''}$, $T_{10-15''}$, $T_{15-30''}$, $T_{30-60''}$, $T_{1-2'}$, $T_{2-5'}$ and $T_{5-15'}$ are similarly defined as the number of transactions in the electronic system in the periods five to ten seconds, ten to 15 seconds, and so on, prior to the last observable event.

Side, which is the indicator of aggressor in the most recent transaction, characterizes the asymmetry in the impact of *completed transactions* on the hazard rates as opposed to the

[13]See the previous footnote.

asymmetry in the impact of aggressive quotes captured by other variables, since the quotes that occur without transactions only indicate the *intention to trade*, not the actual trades. There is a strong evidence that the buy-sell indicator has a high predictive power for the direction of future transactions on the foreign exchange market (Goodhart *et al.* [54]) and on the stock markets (Hausman *et al.* [70], Lo *et al.* [89], Huang and Stoll [75]).

The additional activity and order flow covariates attempt to capture some of the lower-frequency serial dependence in the market dynamics. Such variables should incorporate the influence of common factors contributing to the unobserved heterogeneity that cannot be captured by the current state of the limit order book. Since the failure to account for the unobserved common factors may invalidate the conditional independence assumption which is one of the foundations of the competing risks framework, the variables representing the trade and activity history are chosen to capture a substantial part of the lower-frequency serial dependence, at the same time striking a balance between the correct specification and empirical tractability of the model.[14]

4.2.3 Estimation Results

Tables 4.3, 4.4, and 4.5 report the estimated coefficients of the Cox proportional hazard covariates for the competing risks of arrivals and cancellations of sell-side limit orders and for the sell market orders. Only events recorded during the first three days of the week, October 6–8, 1997, between 6 a.m. and 5 p.m. GMT, which are the most liquid trading hours in the Reuters D2000-2 trading system, have been used for the analysis.[15] Significance of the covariate coefficients is determined from the robust t-statistics (Lin and Wei [87]) at the 99% level, and the statistically insignificant coefficients are shown in small script. The

[14] An alternative lagged activity measure given by the amplitude of transaction price fluctuations in a given time interval leads to qualitatively similar estimation results. Extension of the model to incorporate the dynamic error correction terms in the spirit of the ACD model (Engle and Russell [41]) is currently under investigation. Results of that study will be reported in a separate paper.

[15] The last two trading days of the sample are reserved for out-of-sample evaluation of the forecasts.

qualitative effects of covariates for buy-side events are determined analogously. They closely mirror the covariate effects for sell-side events (Tables 4.6, 4.7, and 4.8).

The empirical regularities uncovered by the Cox proportional hazard regressions can be assigned to one of the three groups depending on whether they are associated with the shifts in price levels, changes in market depth, or changes in the signed order flow. Accordingly, they are reported in sections A, B, and C below. The location of coefficients in Tables 4.3–4.8 providing evidence in support of these empirical facts is indicated in parentheses.

A. Empirical regularities associated with changes in the best bid and ask quotes and transaction prices

(A1) (*The top three entries of column 2 in Tables 4.3 and 4.6.*) The hazard rates of aggressive limit order arrivals tend to be more sensitive to midquote shifts than the hazard rates of market order arrivals, even though the reaction of hazard rates is qualitatively similar for the market and aggressive limit orders. This is consistent with the theoretical prediction of Foucault [45] and the empirical finding of Daníelsson and Payne [31]. In other words, the stronger upward price adjustment without trades implies smaller proportion of market orders in the seller-initiated flow of transactions and larger proportion of market orders in the buyer-initiated flow of transactions.

(A2) (*Column 2 in Tables 4.3 and 4.6.*) Negative signs of *Slippage* coefficients for the submission and cancellation rates of subsidiary sell limit orders and positive signs of *Slippage* coefficients for the submissions and cancellation rates of subsidiary buy limit orders support the view that the general level of subsidiary limit order submission and cancellation activity declines when the midpoint of bid-ask spread moves in the direction of the limit order price and away from the price of the previous transaction. The previously submitted orders are less likely to be cancelled, and the new orders are less likely to arrive when the market price moves closer to the given price level and the move is unaccompanied by transactions. However, the overall effect of the *Slippage*

variable is less obvious when the last transaction leads to asymmetric changes in the best quoted prices and the depth at these prices. In this case, the interpretation becomes complicated because of migration of submission and cancellation events between alternative classes of risk every time the reference quotes (P_{ask} or P_{bid}) are modified.

(A3) (*The top section of column 3 in Tables 4.3 and 4.6.*) There is a strong negative association between the transaction intensity and bid-ask spread, a strong positive association between the rate of large price improvements and bid-ask spread, and a fairly strong negative association between the subsidiary order arrival activity and the size of the bid-ask spread.

(A4) (*Column 3 in Tables 4.3 and 4.6.*) Even though the fresh supply of liquidity (in the form of subsidiary limit order arrivals) and fresh demand for liquidity (in the form of aggressive limit and market order arrivals) are negatively related to the size of bid-ask spread, there is also negative association between the cancellation rates of previously submitted subsidiary limit orders and the size of bid-ask spread. The effect of spread on the supply of subsidiary liquidity is close to neutral. The only source of liquidity positively associated with the size of bid-ask spread are large price improvements that occur more often when the spread exceeds two ticks. Other than that, the liquidity supply appears to be fairly steady and driven primarily by short-term price fluctuations and trends and changes in the depth of the limit order book.

(A5) (*The top three entries of column 5 in Tables 4.3 and 4.6.*) Sell market orders are discouraged by bid price improvement; buy market orders are discouraged by ask price improvement. This "contrarian" property of market orders contrasts the properties of aggressive limit orders that appear to be more frequent following the price improvements on the opposite side of the book. However one should be aware of the possibility of mechanical misclassification of limit orders as being "aggressive" since a larger proportion of sell limit orders submitted at the same prices overlap with higher bid quotes, and a larger proportion of buy limit orders overlap with lower ask quotes, even though

the quote changes might be transitory price swings without any information content.

(A6) (*The top five entries of column 4 in Tables 4.3 and 4.6.*) Arrivals of market sell orders and aggressive limit sell orders, as well as ask price improvements (but not the ask depth improvements) are more likely to occur after recent ask quote deteriorations (increases). A similar regularity is observed on the buy side of the market with regard to recent bid quote deteriorations (decreases).

(A7) (*The mid-section of columns 4 and 5 in Tables 4.3 and 4.6.*) The arrival and cancellation rates of subsidiary limit orders positioned several ticks above the current ask and several ticks below current bid market prices are negatively affected by recent deteriorations of best quoted prices on the same side and recent improvements of best quoted prices on the opposite side of the limit order book. A similar negative reaction to recent quoted price improvements on the opposite side of the book is observed for the arrival rates of price and quantity improving orders. In summary, price improvements on one side of the market are more likely to be followed by reduced liquidity provision on the other side of the market.

B. Empirical regularities associated with depth of the limit order book at the best bid and ask quotes

(B1) (*The top six entries of column 6 in Tables 4.3 and 4.6.*) The arrival rates for sell market orders and aggressive sell limit orders, as well as the rates of ask price and quantity improvements are higher when depth on the ask side of the limit order book is high. A similar effect is observed on the buy side of the limit order book. This points to the competition among aggressive limit order traders for time priority in trade execution as one of the driving forces behind the price improvement and transaction activities.

(B2) (*The seventh and eighth entries of column 6 in Tables 4.3 and 4.6.*) The arrival rates for subsidiary limit orders at the price levels next to the best market bid and ask quotes

are negatively affected by large depth at the best market quotes. This points to the competition among less aggressive limit order traders as the abundant liquidity on the market discourages limit order submission at inferior prices.

(B3) (*The third and seventh entries of column 7 in Tables 4.3 and 4.6.*) Depth improvements at the best market quote become more likely after another depth improvement event at the same price market quote. This provides an evidence that liquidity tends to accumulate at the best market quotes once the market confirms these quotes as the new price levels. Depth improvements at the ask market price also encourage sell market order submission but do not affect the rates of ask price improvement. In fact, this mechanism is predominantly at work when the possibilities of price improvement have been already exhausted (and the market spread is one tick or smaller).[16]

(B4) (*The top three entries of column 8 in Tables 4.3 and 4.6.*) A tentative evidence that buy market orders arrive more actively during the periods of intermediate depth (less than \$10 mln.) and less actively during the periods of large depth (at least \$10 mln. or more) at the best ask price was not found for the arrivals of aggressive buy limit orders. Similar effects are observed for sell market orders and aggressive sell limit orders. On the other hand, the arrival intensity of very aggressive limit orders (submitted at the prices above the best ask and below the best bid quotes) appears to decline as the level of depth on the opposite side of the limit order book increases. This can be easily justified from the observation that very aggressive limit orders are more likely to be used during the periods of scarce liquidity (low depth) at the best market bid and ask quotes.

(B5) (*Column 9 in Tables 4.3 and 4.6.*) There is strong evidence that bid and ask depth improvement events without price improvements encourage cancellations of subsidiary limit orders on the opposite side of the book and discourage liquidity demand coming

[16] Additional Cox regressions for these events run separately for the periods of small and large spread confirm this hypothesis.

in the form of aggressive limit and market orders from the opposite side of the market. Moreover, the deteriorating level of depth at the best market quotes encourages traders on the opposite side of the market to chase more aggressively the remaining liquidity at these quotes and avoid cancellations of subsidiary limit orders on the opposite side of the market.

(B6) (*The bottom section of column 9 in Tables 4.3 and 4.6.*) The limit order cancellation rates at the best bid and ask prices tend to decrease after increases in the observed depth on the opposite side of the limit order book. The increasing depth encourages limit order traders on the opposite side of the book to keep his order at the best market price only if he knows that he will be the first trader to receive this price.[17]

C. Empirical regularities associated with the lagged signed order flow and lagged transaction activity

(C1) (*The top section of columns 2–4 in Tables 4.4 and 4.7.*) The arrival rates of sell market orders, aggressive sell limit orders, and market ask price improvements increase after seller-initiated transactions and decrease after buyer-initiated transactions. A similar effect is detected on the opposite side of the limit order book. This clustering of the buyer and seller pressure is consistent with the ample evidence of strong high-frequency directional momentum observed on foreign exchange and stock markets and reported in previous studies.[18]

(C2) (*The bottom section of columns 2–5 in Tables 4.4 and 4.7.*) There is evidence of higher cancellation rates for subsidiary sell limit orders up to ten seconds after seller-initiated

[17]This is confirmed by additional Cox regressions (not shown in Tables 4.3 and 4.6) run separately for cancellation events leading to price deterioration and leading to depth deterioration without change of the best quoted price.

[18]Lo **et al.** [89] report similar results for limit order data collected from stock markets. Hausman **et al.** [70] reports similar findings for transaction level data.

transactions and lower cancellation rates for subsidiary sell limit orders up to ten seconds after buyer-initiated transactions. A similar effect is detected for subsidiary buy limit orders. This can be generally interpreted as an evidence of limit order book updating by order anticipators receiving signal that the trading pressure is shifting market price in the opposite direction.

(C3) (*The mid-section of columns 5–9 in Tables 4.4 and 4.7.*) The prolonged periods of low submission rates for subsidiary ask limit orders are more likely to occur following the periods of massive seller-initiated transaction activity. However, only a weak evidence of similar effects is found for subsidiary bid limit orders. An explanation can be based on the common perception of the US dollar in the late 1990s as a currency with stronger fundamentals than the Deutsche Mark, which could be translated into the lower sensitivity of limit order bids to strong "Buy" signals. However, this apparent asymmetry may also be period-specific.

(C4) (*Tables 4.5 and 4.8.*) The model provides ample empirical support for the observation that order submission and order cancellation rates increase after the periods of high transaction activity. The effect of transaction activity is very persistent at all price levels in the limit order book. Interestingly, the persistence of activity appears to be stronger for the arrival rates of subsidiary limit orders relative to other types of events. This also points to subsidiary limit orders and, more generally, to stop-loss order execution activity (Osler [110]) as a possible transmission mechanism of shocks and the source of memory and fat tails in the foreign exchange returns.

4.3 Discussion and Extensions of the Model

4.3.1 Implications of Partial Observability of the Limit Order Book

The users of Reuters D2000-2 instantaneously observe the best market bid and ask quotes and the market depth at these quotes,[19] plus the information about direction and price of the most recent transaction. However, in reality the information that can be used by market participants to elicit the state of the limit order book and market conditions is obviously much larger, even if relevant information from other segments of the foreign exchange market and economic and political news announcements are ignored for the moment. What additional information on the state of the limit order book can be obtained from the Reuters D2000 trading screens?

First, the users can switch at any time between the modes of display showing the best prices and quantities in the entire system (in the "Market" mode), or only those available to the individual subscriber (in the "Trader" mode), or both (see Figure 2.1 and the explanation attached). For active traders, these two sets of quotes are closely related to each other most of the time, even though they do not have to be identical. Recall that subscribers to Reuters D2000 must provide the list of potential counterparties (banks) that they are willing to deal with. Therefore, if a quote arrives from a bank that is not on this list, the dealer would be unable to trade at the quote observed on the "Market" part of display. Additionally, after noticing the discrepancy between quotes available on the "Trader" and "Market" parts of display, the subscriber might get additional information on the current and future market liquidity, even if he does not intend to trade immediately on either quote. The multilevel and potentially complex structure of the information available to traders suggests that at

[19] In October 1997, when the data set was collected, a Reuters D2000-2 trader could only observe the value of depth at the best bid and ask quotes only when it did not exceed nine million dollars. Whenever the depth was 10 million US dollars or more, the trader could see litera "R" in the corresponding entry without precise information about the value of foreign exchange available at the best price. The design of Reuters D2000-2 was modified slightly since then.

least some of them might be better informed about the current elasticities of the market demand and supply curves around the best market bid and ask quotes simply because they voluntarily restrict the list of potential counterparties and as a consequence observe some subsidiary orders that remain hidden for the rest of the market.

Second, the dealers monitoring the market before a quote improvement event would normally know the quantity available at the previous best bid and ask quotes, at least for the first few seconds after change in the previous quoted price. The information inferred from the subsidiary quote quickly becomes obsolete if the best market quote does not bounce back, but sometimes it can be useful at least for the first few seconds after quote improvement. A similar explanation can be offered to support the view that traders may know the market depth better than what is revealed by the large quantity indicator. Indeed, the traders who see only this indicator on the screen not only know for sure that the total quantity available at the current market price is at least ten million dollars, but also might be able to evaluate this quantity more precisely from the previous state of the market and additional information available on the "Trader" part of their displays.

These observations suggest an alternative specification of the Cox proportional hazards for competing risks with two additional covariates given by the proxies of instantaneous price elasticities for the market supply and demand curves. The precise definitions of these covariates are following:

- $\varepsilon^{Q}_{\text{ask}} \equiv$ log of the ratio of quantity available at the second-best market ask quote to the difference between the second-best and first-best market ask prices;

- $\varepsilon^{Q}_{\text{bid}} \equiv$ log of the ratio of quantity available at the second-best market bid quote to the difference between the first-best and second-best market bid prices.

A bit surprisingly, no qualitative conclusions made in subsection 4.3.2 about the impact of covariates on the hazard rates of alternative events are affected after the change of model specification incorporating the bid- and ask-side liquidity supply elasticities. Even though, occasionally, the effects of these elasticities are marginally significant, their informational role

for the behavior of traders appears to be miniscule. For this reason, future specifications of the model will not include the elasticity covariates.

4.3.2 An Outline of the Actuarial Approach to Event History

Developing a unified approach to cancellation events remains a perennial challenge for econometricians trying to model these events as part of a truly dynamic limit order market environment. Cancellations of existing limit orders should be treated differently from the arrival events for several reasons.[20] First, the concept of "epoch" as a time period between two events admits several alternative definitions, which are closely related to alternative definitions of the clock time. In the *stochastic process approach* to event history, a new epoch is initiated every time the information on the trading screen is updated (i.e., after any event leading to change in the best bid or ask quote, or change in the depth at the best bid or ask quote). For cancellation risks, there is an alternative view according to which epochs are associated with the lifetimes of limit orders. Thus, at any moment of time, the number of overlapping epochs equals the number of outstanding bid and ask limit orders.[21] We call this treatment of event histories the *actuarial approach*, since there is a clear analogy between the risks of cancellation or execution for an outstanding limit order and the risks of contract termination or death for the beneficiary of life insurance policy. It is reasonable to assume that one of the determinants of cancellation risk for a given limit order might be closely related to its "age", i.e., the time elapsed since the limit order was submitted to the book. Additionally, the risks of events might depend on other factors that can be summarized as the "market conditions" in the limit order market case, the "environmental factors" in biometric and medical experiments, or the "operating environment" in engineering and relia-

[20] See also the discussion of the general philosophy of dynamic survival modeling in the expository paper by Singpurwalla [120].

[21] A version of the actuarial approach is implemented by Lo, MacKinlay, and Zhang [89], who use the tools of survival analysis to estimate the times to limit-order executions for a sample of 100 largest stocks in the S&P500.

bility studies. In particular, market conditions include any information on prices, quantities, and recent transaction activity that is updated in real time on the trading screens. Additionally, the risks of execution and cancellation events are likely to depend on unobserved idiosyncratic components characterizing the traits of agents submitting the limit orders (such as their identities, preferences, sponsors, and trading strategies) as well as the public news announcements in the analyzed period.

As mentioned before, the practical implementation of a full-fledged actuarial approach in a dynamic limit order environment can be computationally prohibitive, especially when the depth of the limit order book is large. Since any market event leading to modification of observable covariates may also affect the hazard rates of competing risks for *each* outstanding limit order in diverse fashion, the most straightforward approach would be to write the model that allows time-varying covariates, and estimate it for a large dynamic panel of buy and sell limit orders. As the computational burden associated with this approach for most practical situations can be extremely costly, several shortcuts to bypass the difficulties in estimation of the complete actuarial model are all based on the idea of dramatic reduction of the number of distinct time periods in the analysis.

One example of such a reduction is the proposal to replace time-varying covariates with their values at the moment of limit order submission, and hold them constant throughout the rest of the limit order history. Then the baseline hazard in the CPH model can be estimated as a function of time elapsed since the limit order submission. A version of this approach was implemented by Lo, MacKinlay, and Zhang [89] in their study of probabilities of limit order execution for a number of stocks in TORQ database. Application of the method of Lo, MacKinlay, and Zhang [89] to our data set yields qualitatively similar results, providing some evidence of similarity in the performance of trading systems based on manual and automated limit order execution. However, since the thrust of our work is on the dynamic links between the continuously updated screen information and traders' activity, rather than the static impact of fixed covariates, as if they are chosen in a controlled laboratory environment, we do not emphasize these results. For the time being, the best approach appears to model the

limit order book history in the old time scale, restarting the calendar clock every time the information on the screen is modified, and restrict the large number of limit orders subject to the cancellation risk to a relatively few types identified by the difference between the limit order and current market price. This is exactly what has been done in section 4.2 of the present chapter.

4.3.3 Discrete Time VAR Approach to the Joint Dynamics of Price, Volume, and Market Depth

The vector autoregression is a popular econometric approach to high-frequency data based on the information aggregated at discrete time intervals. However, it is well known that temporal aggregation can lead to inconsistent estimates of parameters, which is usually the case when a continuous time model is estimated on a sample of data collected in discrete time (Lo [88], Aït-Sahalia and Mykland [5]). In this subsection we compare implications of our model with the estimates obtained by Daníelsson and Payne [31] in the discrete-time VAR framework.

The raw data set in this chapter is identical to the one used by Daníelsson and Payne [31] in their study of the Reuters D2000-2 limit order book, even though the details of algorithms matching the limit and market orders might be slightly different. After pre-processing the data set by a proprietary algorithm which is similar to the one described in section 2.1, Daníelsson and Payne discretize the data in physical time at the 20-second sampling frequency and estimate their model on this discretized dataset. In fact, Daníelsson and Payne apply VAR to the new data set, which is comprised by snapshots of the limit order book and its summary liquidity characteristics recorded at fixed frequency. The discretization approach has many advantages, such as substantial data compression and the possibility to apply a wide range of standard econometric techniques to the regularly spaced observations in the discretized time series. However, the discrete time approach may be of limited value when the main object of study is instantaneous reaction of multiple market participants to

the continuously updated screen information. The main reason that makes interpretation of discrete time model parameters problematic is the aliasing problem that was recognized in the engineering literature long time ago, and relatively recently drew the attention of econometricians who focused mostly on the estimation of continuous time diffusion processes from discrete sample observations (Phillips [114], Lo [88], Hansen and Scheinkman [62], and [5]). Conceptual difficulties arise from the fact that the high-frequency dynamics of trading activity, which is the main research object in the empirical literature on market microstructure, are unidentifiable from the coefficients of discrete time VAR or any other time series model estimated in discrete time. As the reaction time of a typical trader to new information on the trading screens usually does not exceed one or two seconds, the frequency of 20 seconds used in Danielsson and Payne [31] appears inadequate for capturing almost instantaneous reaction of multiple traders to new information. Increasing the frequency of observations would be an obvious way to avoid this problem. However, reformulating of the VAR in discrete time at higher frequency does not provide an easy solution as the higher-frequency sample becomes dominated by inactivity periods, ultimately leading to biases in the estimated coefficients. Explicit formulation and estimation of the empirical model of activity in continuous time appears to be the only reliable way to avoid the aliasing problem.

4.3.4 Summary of Results

The nature of competition among aggressive and non-aggressive traders can be summarized by the following digest of the main empirical implications of the continuous time model. Special attention is drawn to the role of price and depth improvements with and without trade for the price discovery process.

1. The continuous time model of this chapter implies that the buy market order activity is spurred by depth improvements due to higher competition of limit order traders for time priority on the bid side. However, the buy market order activity is deterred if depth improvements on the ask side are not accompanied by simultaneous price

improvements (i.e., spread reductions).

2. The buy market order activity is encouraged by bid quote improvements originating from the competing limit order buyers, and by ask quote improvements originating from the providers of liquidity on the sell side of the market. However, the effect of a bid quote improvement on the buy market order activity is captured almost entirely by the concurrent reduction of the bid-ask spread, and appears to be much smaller in magnitude.

3. In the time period covered by our data, price improvements by competitors were generally considered good news, while depth improvements of competing limit order traders were generally treated as bad news that made market order traders to defer their transactions and consider taking different actions.

4. Overall, the prevalence of buyer-initiated trades among the recent transactions encouraged the buy market order activity and appeared to be good news for traders using such orders. Similar effects are observed on the opposite side of the market for seller-initiated trades.

4.4 Principal Components of the Competing Risk Indices

The main purpose of principal component analysis (PCA) in the present context is reduction of a large set of competing risk indices to a much smaller set of indices (factors) that still generate most of short-term market dynamics. Besides the data compression, PCA represents an important step toward automatic and efficient generation of short-term forecasts for the market activity. In retrospect, careful inspection of the market activity around the times of news arrivals reveals a typical market reaction pattern, which starts with a sudden and numerous withdrawals of orders on the "weaker" side of the market as it comes under

pressure. The limit order withdrawals are accompanied by larger than usual values of bid-ask spread and strong directional trading volume driven primarily by market orders. This activity creates a considerable imbalance between the number of buy and sell orders in the limit order book, which is gradually restored after the price adjustment that occurs in a staggering wave-like fashion and is accompanied by higher than usual volatility of the price. No matter whether the public news announcement confirms dealers' expectations about Bundesbank cutting German interest rates, the severity of government crises in Italy, or the vulnerability of East Asian financial markets, as long as the pattern of traders' reaction to such information releases has some common characteristics, it may be captured by a small number of factors that could potentially be much smaller than the number of market participants or the number of initially identified market components.

The version of PCA conducted in this section is based on the analysis of the sample variance-covariance matrix $\widehat{\Sigma}$ of competing risk indices $Y_r = \mathbf{Z}'\boldsymbol{\beta}_r$, $r = 1, ..., R$. Matrix $\widehat{\Sigma}$ is formed by R sample variances of risk indices

$$\widehat{\sigma}_r^2 = \frac{1}{N}\sum_{n=1}^{N}(y_{nr} - \overline{y}_r)^2, \quad r = 1, 2, ..., R,$$

as well as $R(R-1)/2$ sample covariances among the pairs of risk indices

$$\widehat{\sigma}_{rr'} = \frac{1}{N}\sum_{n=1}^{N}(y_{nr} - \overline{y}_r)(y_{nr'} - \overline{y}_{r'}), \quad r, r' = 1, 2, ..., R, \quad r \neq r',$$

and contains a large amount of information about the contemporaneous variations in the components of covariate vector $\mathbf{Z}$ characterizing the state of the limit order book. The data compression can be achieved since PCA approximates the parametric covariate-dependent competing risk indices $Y_r = \mathbf{Z}'\boldsymbol{\beta}_r$ by alternative risk indices $\widetilde{Y}_r = \mathbf{U}'\boldsymbol{\gamma}_r$ that depend on a limited set of common factors $\mathbf{U} = (u_1, ..., u_Q)'$. These factors are formed as simple linear combinations of the covariates with the coefficients chosen in such a way that the first factor u_1 explains the largest portion of the sample variance-covariance matrix[22] of the competing

[22]Since the ultimate goal of analysis in this chapter is construction of short-term forecasts of trading activity, our main concern to the end of this chapter will be approximating the short-run variances and

risk indices, the second factor u_2 explains the next largest portion, and so on. The factors constructed in such a manner and normalized to have unit variance are called the *principal components*, or *PCA risk factors*. Then the factor loadings of these observable factors (which are, by construction, the linear combinations of covariates) can be obtained by application of standard regression techniques.

4.4.1 Relative Contribution of PCA Factors

Since, mathematically, the principal components are normalized eigenvectors of the variance-covariance matrix $\hat{\Sigma}$, their number is equal to the number of competing risk indices R. As the last few eigenvectors of $\hat{\Sigma}$ point to the directions where the risk indices Y_1, Y_2, ..., Y_R jointly exhibit little or no variation, most of the information content of the data is likely to be represented by some smaller number of PCA risk factors $u_1, ..., u_Q$, $Q < R$. In fact, the situation when $Q \ll R$, and the principal components $u_1, ..., u_Q$ result in a much smaller data set, also cannot be ruled out.

The importance of PCA risk factors u_1, u_2, ..., u_Q across the set of competing risk indices is measured by the eigenvalues λ_1, λ_2, ..., λ_Q of the extracted factors. In particular, the proportion of the total variance described by the first Q principal components is given by the ratio

$$\frac{\sum_{q=1}^{Q} \mathrm{Var}(u_q)}{\sum_{r=1}^{R} \mathrm{Var}(u_r)} = \frac{\mathrm{tr}(\mathrm{Var}(\mathbf{UU}'))}{\mathrm{tr}(\Sigma)} = \frac{\sum_{q=1}^{Q} \lambda_q}{\sum_{r=1}^{R} \lambda_r}.$$

The natural question arises as to how many principal components should be retained to capture systematic variation in the original data set and avoid capturing what is likely to be a random noise. Since there is no single universally accepted statistical approach in the statistical literature to the number of PCA factors to be retained, it would be reasonable to take an eclectic approach applying a spectrum of alternative criteria to this problem.

covariances of the competing risk indices. Therefore, the analysis throughout the rest of this chapter will be based primarily on the properties of short-run variances and covariances. It remains to be seen whether the similar data compression performed at the lower frequencies leads to useful forecasts.

Figure 4.10 represents the so-called *scree plot* showing the eigenvalues corresponding to the first 13 principal components against the number of those components. Inspection of the graph suggests that the first five PCA factors capture the major portion of variation in the risk indices, even though the sixth and seventh factors might also be marginally important. Table 4.13 shows that the first seven PCA factors capture almost 90% of variation in the indices, whereas the first five PCA factors capture almost 86% of the variation. The "subjective" choice of $Q = 5$ appears to be reasonable in the present context.

There is a large number of "objective" decision rules frequently applied to decide on the number of "significant" principal components to be retained. Jolliffe [78], Chapter 6 gives a good survey of formal and informal approaches, while Jackson [76] investigates performance of alternative decision rules applied to some artificial and real data. Even though a large amount of research has been done on the rules for choosing the number of retained components, there is no universally accepted rule that is applied in the literature in all circumstances.

One of the popular statistical decision rules which is strongly favored in Jackson [76] can be derived from the so-called broken stick model (Frontier [47]). According to this model, if the total variance, represented by the sum of the eigenvalues of the variance-covariance matrix of indices, can be divided randomly among Q components, then the distribution of components follows a "broken-stick" distribution, with the expected kth largest eigenvalue calculated as

$$\lambda_k^* = \frac{1}{Q} \sum_{q=k}^{Q} \frac{1}{q}, \tag{4.2}$$

when the number of components Q is large enough. One way of deciding whether the proportion of variance accounted by the kth PCA factor is sufficiently large for this component to be retained is to compare this proportion with λ_k^* given by (4.2). The test based on λ_k^* leads to the conclusion that the first four components should be retained. However, since there are no systematic results on the size and power properties of the "broken stick" rule based on the expected eigenvalues, and in view of frequently cited evidence that the rule

retains too few components, we check for the number of components using a bootstrap procedure that recovers the entire distribution of the kth largest eigenvalue under the null of equal eigenvalues for the last $Q - k + 1$ PCA factors. This alternative procedure leads to the conclusion that $Q = 5$ components should be retained, with the fifth component being only marginally significant at the 95% confidence level.

Yet another, rather subjective procedure for selection of the number of retained components uses the method of log-eigenvalue (LEV) diagrams (Wilks [126], Chapter 9). The method is motivated by the idea that, if the last $R-Q$ principal components pick up random noise, then the magnitudes of their eigenvalues decrease exponentially with the component number. The Q retained PCA factors correspond to the log-eigenvalues deviating from the straight-line portion of the plot on the LEV diagram. Figure 4.11 shows that the LEV plot deviates from the linear pattern implied by exponential decay of the plot under the null hypothesis of random noise for the number of factors as low as five. Even though it cannot be unambiguously seen on the plot that $Q = 5$, it appears to be a reasonable choice again. Therefore, the further analysis will be conducted for five PCA factors. However, to ensure robustness of our results, we repeated the analyses for seven retained PCA factors and obtained similar results.

4.4.2 Interpretation of PCA Factors

Table 4.9 gives the representations of covariates in the original Cox proportional hazard regressions for the competing risks in terms of the extracted PCA factors. Tables 4.10, 4.11, and 4.12 report the estimates $\widehat{\gamma}_r$ of factor loadings on the first five PCA factors in the semiparametric Cox proportional hazard model for the competing risk indices $\widetilde{Y}_r = \mathbf{U}'\boldsymbol{\gamma}_r$ of sell order arrivals, buy order arrivals, and cancellation activity in the limit order book, respectively. The estimation period is between 6 a.m. and 5 p.m. GMT on the week of October 6-8, 1997. The t-statistics for the estimates of Tables 4.10, 4.11, and 4.12 are shown in parentheses, and statistically significant factor loading estimates coefficients (at the 95%

level) are marked by stars. To facilitate interpretation of the extracted PCA factors, we also show the diagonal and off-diagonal terms of cross-correlogram (Figures 4.13 and 4.12) for the first five PCA factors.

The first principal component, which is obviously nonstationary, dominates the dynamics of risk indices. The first PCA factor captures approximately 44% of the total variation in the covariate indices of the competing risks. Since all types of risk have large positive factor loadings on the first factor, the first component can be readily interpreted as the general level of *limit order book activity*.

The second principal component contributes less than 16% to the total variation of risk indices dynamics. Inspection of its autocorrelogram reveals its stationarily, and its cross-correlogram with the first principal component shows that they interact only marginally at all leads and lags. The factor loadings on the second factor are uniformly positive for log-hazard rates of buyer-initiated events (associated with submission and cancellation of bid limit orders and buy market orders) and almost uniformly negative for the log-hazards of seller-initiated events. Therefore, we can interpret the second factor as the *short-term activity momentum*, which identifies the more active side of the limit order book (buy or sell) without differentiating across the types of activity (whether it is submission or cancellation of limit or market orders). Since the active order-driven market intrinsically represents the dynamic interaction of buyers and sellers, it should be no surprise that buyers' or sellers' actions cannot dominate the market for long periods of time. Indeed, the autocorrelation of this factor becomes indistinguishable from zero for lags as low as 50 time periods (epochs), where each epoch is assumed to be terminated by an observable limit order book event and lasts about one second on average during the hours of liquid trading.

The third principal component capturing 12.4% of variation in the hazard rate dynamics has slowly decaying autocorrelogram and therefore has long memory. The rate of decay of its autocorrelogram on Figure 4.13, characterizing persistence of the third PCA factor, is close to the rate of decay of the autocorrelogram for the first factor. Similarly to the second principal component, the third PCA factor is uncorrelated to the first factor at all

leads and lags but interacts rather non-trivially with the second principal component. Since the factor loadings of the third component are negative for submissions of non-aggressive bids and aggressive seller-initiated orders, and positive for submissions of non-aggressive ask orders and aggressive buyer-initiated orders, we can think of the third factor as the "*buying pressure*" on the market. The cancellation activity pattern that can be uncovered from the factor loadings of cancellation risks generally countervails those for the second PCA factor (short-term activity momentum) and conforms to the intuitive notion of "buying pressure." Aggressive buyers who tend to cancel buy limit orders several ticks below the market bid price, keep limit orders just below the market bid price in the hope of price reversal, and do just the opposite on the ask side of the market. Interestingly, since the cancellation activity patterns at the best market bid and ask quotes are similar to the submission patterns at the more competitive prices (leading to price improvement and narrowing the spread), this also conforms to the activity of aggressive sellers who might be testing the market before resubmitting their orders if their earlier offers have not been hit promptly.

The analysis of cross-correlations between factors 2 and 3 is conducted using the augmented graph of autocorrelogram highlighting the interaction between factors 2 and 3 (Figure 4.14). The graph shows that the unusually low activity on the ask-side of the limit order book relative to the bid side (small values of factor 2) precede the aggressive buyer pressure (high values of factor 3), leading to moderately higher ask-side activity in the short run that eventually reverts to persistently lower ask-side activity in the long run. This observation reveals the non-trivial interaction between the two factors, which implies a richer story than most theoretical microstructure models can tell.

Since the interpretation of the consecutive principal components becomes increasingly difficult, we make an attempt to interpret only the fourth and fifth components that contribute, respectively, to 9% and 4.7% of variation in the parametric parts of the log-hazard rates of competing risks. The loadings on factor 4 appear symmetric for buyer- and seller-initiated events, and are positive for limit order submissions within the spread that do not cause immediate transactions, as well as for limit order at least five ticks away from the

prevailing market bid-ask spread. The loadings on factor 4 are negative for the arrivals of market and limit orders causing transactions, as well as for submissions and cancellations of subsidiary orders in the vicinity of the market bid-ask spread. Therefore, high values of factor 4 can be associated with the tendency to quote more competitive prices within the spread, while high values of factor 4 are associated with the tendency to take limit orders on the opposite side of the book without much bargaining. Therefore, low values of factor 4 can be associated with "choppy" markets when trades tend to occur without much bargaining, which often happens at the high levels of the market spread and might be associated with the relatively high adverse selection component (Harris [64]). We may attach the term "*adverse selection*" to the fourth PCA factor, and stick to the terminology in the future.

Note that the "adverse selection" factor does not appear to interact much with any of the other major principal components except the first one, which provides an illustration of the frequently reported phenomenon that the general level of trading activity is higher when some information is present in the market (i.e., the "adverse selection" component is high). Indeed, active markets lead to higher competition among traders who tend to submit more quotes before making a deal. The inverse causality also appears to be at work. Trading at the relatively large levels of bid-ask spread might be a signal of "choppy market," at least for some, presumably uninformed, traders, causing adjustments of their quotes until the market returns to the "smoother" state.

Finally, we make an attempt to interpret the fifth PCA factor. The loadings on this factor are positive for all aggressive buy limit order arrivals and cancellations, for submissions of buy limit orders well below the bid market price, for submissions of sell limit orders just above the ask market price, and for all cancellations of subsidiary limit orders on the ask side. In all other cases, the loadings on factor 5 are negative and their signs and magnitudes are roughly symmetric to the signs and magnitudes of this factor loadings for similar risks on the opposite side of the book. Even though, in many respects, factor 5 behaves similarly to factor 3, large values of factor 5 imply that the aggressive buyer activity tends to be accompanied by cancellations of subsidiary sell limit orders rather than arrivals of subsidiary sell limit

orders, in contrast to factor 3. Also, factor 5 does not appear to be as persistent as factor 3 (Figures 4.13 and 4.14), and generally is more sensitive to the changes of depth (Table 4.9). All these properties might be interpreted as a tentative evidence of directional information on the market that signals a permanent shift of the market price in the near future. For this reason we will call the fifth PCA factor the "*bull market momentum*" throughout the rest of this chapter. In a small forecasting exercise of section 4.6 we pay closer attention to the second, third, and fifth principal components, that seem to be major determinants of directional activity in the limit order book.

4.5 Simulation Experiments

To validate the estimation results of section 4.2 and get a better idea about the strengths and limitations of the proposed competing risks model, we generate simulate the sequences of events originating from interactions of limit and market orders as described by the Cox proportional hazard specification of competing risks. The estimation period covers the active trading hours (between 6 a.m. and 5 p.m. GMT) of the first three days in the sample and includes almost 114000 observable and unobservable events.[23] As a by-product of the simulation experiment, we obtain long histories of bid-ask spreads, transaction prices, and other major characteristics of the trading process. Then the time-series properties of simulated prices, quantities, order flow, transaction price volatility, and their predictive ability with respect to each other can be compared to their analogues for actual trajectories.

In the second part of this section, we simulate the cumulative distribution functions for the time elapsed to specified event of interest, when the covariate vector and the initial state of the limit order book are specified in advance, and given the future transaction price at the moment when the event occurs. The non-trivial interaction patterns between alternative types of limit orders appear to be responsible for a variety the shapes of conditional

[23]The last two trading days of the sample are reserved for out-of-sample evaluation of the forecasts in section 4.6.

distributions given the future transaction prices. Even though the results of the simulated exercise in subsection 4.5.2 may be state specific, they point to the significant role of the limit order book for the price formation and future transaction activity, and indicate the importance of its continuous monitoring whenever the timing of trading decisions is a major concern.

4.5.1 Simulation of Quote and Transaction Histories

The simulation procedure generates the trajectories of transaction prices, signed directional flow of transactions, and the limit order book imbalance such as the ones shown in Figure 4.15. The simulation experiment was conducted with the parameters of competing risks equal to the estimates reported in Tables 4.3–4.8. It was assumed that the initial bid and ask prices are $P_{\text{bid}} = 1.7498$ and $P_{\text{ask}} = 1.7502$ DEM per dollar, the quantities available at these prices are, respectively, $Q_{\text{bid}} = \$2$ mln. and $Q_{\text{ask}} = \$2$ mln., and the last two recent transactions (both initiated by sellers of US dollars using market orders) were for \$1 mln. each and occured within the last two seconds at the price $P^* = 1.7499$. It was also assumed that no other limit order book activity has been reported after the two most recent transactions, no transactions occured prior to the two most recent trades over the period of 15 minutes, and the bid and ask sides of the limit order book consisted of eleven limit orders of size \$2 mln. each evenly distributed in the interval $[P_{\text{bid}} - 0.001; P_{\text{bid}}]$ on the bid side and in the interval $[P_{\text{ask}}; P_{\text{ask}} + 0.001]$ on the ask side.[24] All simulations were performed under the assumption that only incremental events[25] are observable and change the state of the

[24] In all simulations we assumed that the size distribution of all arriving and cancelled limit orders at any given price level is independent of the market conditions; in particular, the ratios of hazard rates for the limit orders of sizes Q and Q' arriving at P_{bid} are constant and independent on other market conditions for all Q and Q'. In view of the evidence of information content for very large limit orders in stock and currency markets, it remains to be seen whether and to what degree the assumption we make is justified by the empirical data.

[25] Incremental events are defined as the events occuring within the market bid-ask spread or at the bid and ask touch, and therefore associated with immediate changes of information on all Reuters screens.

limit order book.

The visual similarities between the plots of Figure 4.15 and those shown on Figures 1.1 and 1.7 in Chapter 1 are confirmed by the comparison of descriptive statistics for the trajectories of actual and simulated processes. The comovement of the graphs reflects strong positive dependence between the signed order flow and the transaction returns. Moreover, the signed order flow has memory and a reasonably good predictive power in comparison to the transaction returns, which appear virtually uncorrelated with the past order flow activity at time horizons larger than five seconds. In all simulation experiments, the artificial trading histories show persistence in the order flow time series and the ability of order flow to predict the aggressor and the very short-run dynamics of future transaction prices.

On the other hand, the order flow has virtually no explanatory power for the midpoint of future quoted bid-ask spread. Despite the strong evidence of directional momentum in the transaction activity driven by past order flow, there are significant negative low-order autocorrelations in the time series of returns. In fact, the evidence of large first-order negative autocorrelations in transaction returns is consistent with the results of earlier studies (see Hasbrouck and Ho [68], Goodhart *et al.* [54], among others).which reflect the widespread phenomenon in almost all tick-by-tick quote and transactions data that have been traditionally explained by the bid-ask bounce and asynchronous trading (see Chapter 3 in Campbell, Lo, and MacKinlay [19]).

Several aspects of the simulated trajectories do not appear to be perfectly realistic and point to possible sources of model misspecification. First, the simulated time series of best bid and ask prices do not fluctuate as much as their empirical counterparts. This price inertia may be caused by the accumulation of limit orders at and near the best bid and ask prices where the cancellation rates are on average lower than the rates of new limit order arrivals. Such accumulation of limit orders gradually leads to the formation of barriers at and near the best bid and ask quotes that turn out very difficult to break through. While formation of such barriers at certain levels is reminiscent of the clustering

of stop-loss orders in the real markets studied by Osler [111],[26] the model in the present form does not contain a well-articulated self-enforcing mechanism for the price to break through these barriers. As a result, the simulated prices tend to bounce between two clusters of limit orders much longer than the real prices in the market do. Even though various adjustments of the model are possible, and its deficiencies may be fixed on *ad hoc* basis, most of the attempts to do this broke the subtle ecological balance between the order arrivals, cancellations, and transactions. We make a preliminary conclusion that the model in its present form is inapplicable as a realistic simulator of very long trading histories, if the goal is to simulate the price sequences that would reflect all important aspects of the real market prices. However, this does not preclude future applications of the model amendments as testbeds for newly launched electronic trading systems and for validation of new trading strategies, provided that there is a continuous feedback to the simulator from the real market data.

Another source of model misspecification may provide an alternative explanation of the excessive price inertia observed in the artificial trajectories. Specifically, the present model specification ignores a large amount of information which is likely to be observable by most traders but is conspicuously absent from the set of covariates. Examples of critical pieces of such information include the messages from the Reuters electronic newsline, information from the forward market, information from other active currency markets, and, last but not least, the knowledge of depth of the limit order book at second-best prices. The design of the Reuters D2000-2 trading system allows its subscribers to use the "Trader" section of the screen in addition to the "Market" section (see Figure 2.1) and thereby peek slightly deeper than the very top of the book. Finally, traders may remember or at least have a limited knowledge of the limit order book depth at subsidiary price levels, while such information is not instantaneously available on the trading screens, since quotes at the subsidiary price levels get concealed as a result of the recent price improvements.

[26] Some evidence of clustering of the limit order quotes and transaction prices can also be found in our data set.

The fact that the volatility of simulated trajectories tends to be smaller than the volatility of actual transaction prices over the time intervals of similar length poses another challenge to the present form of the model. By itself, this failure to fit the volatility pattern may be not so surprising as it can be a manifestation of the fact that the major portion of intraday price variability should be attributed to events outside of the Reuters D2000-2 limit order book. In other words, since trading in the analyzed segment of the foreign exchange market represents only a fraction of foreign exchange trading activity for the same period, the quote and transaction activity in the Reuters limit order book can hardly be responsible for a major portion of variability in foreign exchange rates that we typically observe in the real world. Therefore, the question about the ultimate cause of high intraday volatility in the foreign exchange market cannot be affirmatively answered without major modifications of the proposed framework. One obvious suggestion (that would be nevertheless difficult to implement in practice) includes collecting high-frequency data from all essential segments of the foreign exchange market over the same sample period, and expanding the model to accommodate the interaction of trading activity in the multiple segments of foreign exchange market.

4.5.2 Simulation of Time to Event

Generating multiple scenarios of alternative events from the statistical distributions of competing risks represents the first step in the development of dynamic strategies to exploit profit opportunities, or reduce the liquidation cost for open positions. As usual in the proposed competing risks environment, we assume that all market participants react to the conditions captured by the vector of covariates, and the intensity of their reaction can be represented by a system of hazard functions of these covariates. In the order-driven financial market viewed as an arena of competition among multiple agents for time priority, the non-homogeneous reaction of multiple traders to sequences of buyer- or seller-initiated transactions (market momentum) might be a potential source of profit opportunities or reductions in liquidation

costs when the existing open positions are closed.

Figure 4.16 shows the simulated plots of cumulative distribution functions for the *time to the next seller-initiated transaction*, conditional on the price level of this first transaction. Apparently, all simulations are performed under the initial conditions, given by the state of the limit order book and the recent history. In this experiment we assume that the initial bid and ask prices are $P_{\text{bid}} = 1.7498$ and $P_{\text{ask}} = 1.7502$ DEM per US dollar, the quantities available at these prices are, respectively, $Q_{\text{bid}} = \$2$ mln. and $Q_{\text{ask}} = \$2$ mln., and the last two recent transactions (both initiated by a seller of US dollars using market orders) were for \$1 mln. each and occured within the last two seconds at the price $P^* = 1.7499$ DEM per US dollar. It is also assumed that no other limit order book activity has been observed between the time of the last transaction and the present (i.e., the current epoch was triggered one second ago by a seller-initiated transaction), that no trade occured at least for 15 minutes prior to the previous two transactions, and that the bid and ask sides of the limit order book consist of eleven limit orders of size \$2 mln. each that are evenly spread within one tick of each other in the intervals $[P_{\text{bid}} - 0.001; P_{\text{bid}}]$ and $[P_{\text{ask}}; P_{\text{ask}} + 0.001]$, respectively. The experiment is performed 500 times using the competing risks calibrated by the parameter estimates obtained in section 4.2.

Provided that the initial state described above corresponds to the actual market conditions, the results may be interesting for market order traders and their sponsors contemplating the effect of their trades on the future evolution of the market. For instance, consider a trader who needs to close his short US dollar position that has been originated by two \$1 mln. transactions at the price $P^* = 1.7499$ DEM per US dollar. He might consider placing a bid limit order and would like to evaluate the probability that the market could be driven down temporarily far enough to make a purchase of US dollars at bid price a bargain. However, even though the short-term directional momentum set off by the two recent seller-initiated trades moves down the market bid price (at least temporarily), it is highly unlikely that it would be strong enough to carry the ask price below the price of the most recent transaction. In the experiment of Figure 4.16, the first seller-initiated transaction occured at the price

below $P^* = 1.7499$ DEM in 315 out of 500 cases (63%), showing that the simple strategy of selling aggressively and trying to "catch the wave" with a buy limit order submitted at the price below $P^* = 1.7499$ DEM might pay off, at least under the specified market conditions. However, this strategy can be risky since (a) the market conditions (in particular, the state of the limit order book) are usually unknown, (b) even when the market conditions are known, the bid limit order may never get hit, and the market price actually can go up and not return to $P^* = 1.7499$ DEM in the foreseeable future. A simple way to reduce the non-execution risk associated with open limit order positions would be to buy using a market order once the ask market price goes down below $P^* = 1.7499$ DEM. However, this event does not happen frequently enough in our simulation experiments to justify this strategy.[27] Taking into account the $25 transaction fee (paid in this market by aggressors), the attempt of creating profitable momentum is likely to be self-defeating, at least in a relatively quiet market environment such as the one described above, as it leaves the trader with an open position that he would be forced to carry over or close with a loss.

4.6 Order Flow Forecasts Based on the Limit Order Book Information

The market efficiency has been an issue of primary concern for both academics and practitioners since the seminal work of Bachelier [9]. In its most popular form, the efficient market hypothesis is a statement that prices in financial markets fully and correctly reflects all public information. As a consequence, the price changes in an informationally efficient market can only occur in response to the new information, and therefore they are unpredictable. The early academic studies summarized by Fama [44] found some evidence in support of the view that major financial market are very efficient. At the same time, since Cowles [25]

[27] The market ask price went down to $P_{\text{ask}} = 1.7498$ or below prior to the first seller-initiated transaction only in 24 cases out of 500. It is unlikely that it would happen more frequently if the trader waits longer.

and Niederhoffer and Osborne [105], there was a substantial body of literature questioning the view of market efficiency on both empirical and theoretical grounds. It was also demonstrated that the predictability of financial returns and other quantities is related but not synonymous to the market inefficiency. Specifically, the predictability does not necessarily imply market inefficiency in the sense that there exist viable dynamic strategies making positive riskless returns.[28]

The growing consensus in the economics profession is that it is possible, though extremely difficult, to predict the changes in prices of financial assets. Given the ample evidence about dependence between the order flow and the dynamics of returns in the foreign exchange positions, the natural question is whether the magnitude of this effect is high enough to generate reliable short-term forecasts. Even though structural and time series models of foreign exchange traditionally faced considerable difficulties generating reliable forecasts of exchange rates (Meese and Rogoff [101]), more recent studies suggest a greater degree of success in generating intraday forecasts at intradaily frequencies. For example, Zhou [128] pre-processes the tick-by-tick DEM/USD data using a "de-volatilization" algorithm to construct the trading signals based on the direction of "abnormal" (larger than two standard deviations) "de-volatilized" returns. Evans and Lyons [43] present an econometric model of joint price and order flow determination and show that application of this model to daily foreign-exchange returns produces good out-of-sample forecasts at short horizons. Apart from the substantial anecdotal evidence of profitability of technical analysis, the recent systematic surveys of practitioners in the interbank foreign exchange market (Cheung and Chinn [21], Cheung and Wong [22]) provide a support for the view that the foreign exchange market may be forecastable in the very short run. As a consequence, the signed order flow is a viable indicator of traders' willingness to back their beliefs with real money, and should be taken seriously as a signal of future direction of exchange rate movements.

[28] On the other hand, the market timing and design of efficient trading strategies is an issue of practical importance in both efficient and inefficient markets, since traders often have the commitments to liquidate their positions and would like to do it at a minimum loss.

The issue of predictability of financial markets has always been especially relevant for foreign exchange dealerships and portfolio managers who spend their time in constant pursuit of elusive profit and earn substantial fees executing orders of their clients. In view of the recent interest of practitioners to developing dynamic trading rules and active asset management strategies, evaluation of forecasting accuracy becomes one of the critical issues of modern research. For example, Levich and Thomas [86] demonstrated the statistical significance of some technical trading rules against conventional parametric models of exchange rate determination. More recently, Lo, Mamaysky, and Wang [90] applied kernel nonparametric regression techniques to technical pattern recognition and found that several technical indicators provide incremental information of practical importance for stock market valuation. In application to high-frequency foreign exchange data, Dacorogna *et al.* [29] examines real-time trading models under heterogeneous trading strategies and makes the conclusion that the identification of the heterogeneous market microstructure in their models leads to excess risk-adjusted return.[29]

In this section, we give the forecasting formulas for the probabilities of next limit order book events. Then we briefly discuss alternative measures of forecasting performance (including the monetary measures based on expenditure analysis) drawn predominantly from meteorology and quality control literature. The selected graphical tools for forecast evaluation will be applied to measure the goodness-of-fit of the five-factor principal component competing risks model (section 4.4) calibrated to the data collected over the liquid trading period (6 a.m. to 5 p.m. GMT) on October 6–8, 1997, and evaluate the out-of-sample predictive performance of this model in the liquid trading period on October 9–10, 1997. This exercise is especially interesting since our out-of-sample period is dominated by highly volatile trading following the Bundesbank announcement about an increase in the repo rate around 11:30 a.m. GMT on Thursday, October 9, 1997. The volatility and spread over a large portion of the out-of-sample period remained much higher than they were during the

[29] Chapters 9 and 11 of Dacorogna **et al.** [30] give a summary of the framework for evaluation of real time forecasting performance of trading models and their applications to foreign exchange markets.

period October 6–8, 1997. The ability of the competing risks model calibrated on historical data to generate credible out-of-sample probability forecasts will be a clear indicator of its promise in improvement of short-term probability forecasts of market events and in their applications to the analysis of alternative scenarios in real time.

4.6.1 Forecasting the Probability of Next Event

Our probability forecasts will be based on the expressions (3.10) and (3.11) with the version of kernel estimator of the baseline hazard function used in section 3.5. Throughout this subsection we also maintain the regularity assumption

$$\lim_{\tau \to +\infty} \int_0^{\tau} h_{0r}(u)du = \infty, \quad \text{for some } r = 1.2, ..., R, \tag{4.3}$$

which guarantees that the probability of no event in interval $[0; \tau]$ converges to zero as the period without arrivals of new observations expands. Then t_0 seconds after the arrival of a new observable event the updated risk index $Y_{nr} = \mathbf{z}_n' \boldsymbol{\beta}_r$ for the risk of type r feeds into the expression of the hazard rate to yield

$$H_r(t; Y_{nr}) = H_{0r}(t) \exp(Y_{nr})$$

and into expression (3.10) for the incidence rate to obtain the conditional probability that event of type r will be next to occur less than t seconds after the previous observable event provided that the observation period has started t_0 seconds after the previous observable

event[30]

$$P_r(t_0, t; Y_{n\cdot}) = \int_{t_0}^{t} P_0(t_0, u; Y_{n\cdot}) \exp(Y_{nr}) dH_{0r}(u), \quad r = 1, ..., R, \tag{4.4}$$

where

$$\begin{aligned} P_0(t_0, t; Y_{n\cdot}) &= \prod_{u \in [t_0; t)} \left(1 - \sum_{r=1}^{R} \exp(Y_{nr}) dH_{0r}(u)\right) \\ &= \exp\left[-\sum_{r=1}^{R} (H_{0r}(t) - H_{0r}(t_0)) \exp(Y_{nr})\right], \end{aligned}$$

is the conditional probability that no event occurs over the period $[t_0; t)$ after the previous observable event given that no event has occured over the period $[0; t_0)$ following the previous observable event. The conditional probabilities (4.4) are estimated by the incidence rates (3.10), using the formula

$$\widehat{P}_r(t_0, t; \widehat{Y}_{n\cdot}) = \int_{t_0}^{t} \widehat{P}_0(t_0, u; \widehat{Y}_{n\cdot}) \exp(\widehat{Y}_{nr}) d\widehat{H}_{0r}(u, \widehat{Y}_{nr}), \quad r = 1, ..., R, \tag{4.5}$$

where

$$\widehat{P}_0(t_0, t; \widehat{Y}_{n\cdot}) = \exp\left[-\sum_{r=1}^{R} (\widehat{H}_{0r}(t, \widehat{Y}_{nr}) - \widehat{H}_{0r}(t_0, \widehat{Y}_{nr})) \exp(\widehat{Y}_{nr})\right],$$

$\widehat{H}_{0r}(\cdot, \widehat{Y}_{nr})$ are obtained by integration of one of the kernel hazard rate estimators (3.13), (3.14), (3.18), (3.22), or (3.23), discussed in section 3.3, and $\widehat{Y}_{n\cdot} = (\widehat{Y}_{nr})_{r=1}^{R}$ is the estimated vector of risk indices. In particular, if only the subset of observable types of events $r = 1, ..., S$ $(S < R)$ and the epochs terminated less than τ seconds $(0 < \tau \leq \infty)$ after the last observable

[30] A more parsimonious alternative characterization of the risk indices could be obtained in terms of the PCA factors $\mathsf{u}_n = (u_1, u_2, ..., u_Q)'$ (Section 4.4), since

$$Y_{nr} = \mathsf{u}_n' \gamma_r + \tilde{v}_{nr} = \tilde{Y}_{nr} + \tilde{v}_{nr}, \quad r = 1, ..., R.$$

In this case the set of baseline cumulative hazard functions $H_{0r}(u)$ need to be recalculated to incorporate the portion $\tilde{v}_{nr}$ of the covariate index $Y_{nr} = \tilde{Y}_{nr} + \tilde{v}_{nr}$ assigned to the noise component. The conditional probability formulas look similar, except that the indices $\tilde{Y}_{nr}$ and baseline hazard functions $\tilde{H}_{0r}(u)$ are captured by the more parsimonious factor structure.

event are considered in the analysis (see section 3.2), then the probability evaluated $t_0 < \tau$ seconds after the last event that the given epoch will be terminated with an event of an observable type $r = 1, ... S$ is given by the ratio

$$\pi_r(t_0, \tau; Y_{n\cdot}) = \frac{P_r(t_0, \tau; Y_{n\cdot})}{\sum_{s=1}^{S} P_s(t_0, \tau; Y_{n\cdot})}. \tag{4.6}$$

Finally, after substitution of the expression for the incidence rates $P_s(t_0, u; Y_{n\cdot})$ into the integral (4.4) and some manipulations, we obtain

$$\begin{aligned} P_r(t_0, \tau; Y_{n\cdot}) &= \int_{t_0}^{\tau} \exp\left[Y_{nr} - \sum_{s=1}^{S} \exp(Y_{ns}) H_{0s}(u)\right] dH_{0r}(u) \\ &= \int_{t_0}^{\tau} \frac{\exp(Y_{nr}) h_{0r}(u)}{\sum_{s=1}^{S} \exp(Y_{ns}) h_{0s}(u)} \exp(-H(u, Y_{n\cdot})) dH(u, Y_{n\cdot}) \\ &= \int_{H(t_0, Y_{n\cdot})}^{H(\tau, Y_{n\cdot})} \frac{\exp(Y_{nr}) \tilde{h}_{0r}(\theta)}{\sum_{s=1}^{S} \exp(Y_{ns}) \tilde{h}_{0s}(\theta)} \exp(-\theta) d\theta \\ &= \int_{\exp(-H(t_0, Y_{n\cdot}))}^{\exp(-H(\tau, Y_{n\cdot}))} \frac{\exp(Y_{nr}) \tilde{h}_{0r}(-\log(v))}{\sum_{s=1}^{S} \exp(Y_{ns}) \tilde{h}_{0s}(-\log(v))} dv, \end{aligned} \tag{4.7}$$

where

$$H(u, Y_{n\cdot}) = \sum_{s=1}^{S} \exp(Y_{ns}) H_{0s}(u)$$

is the cumulative hazard function for the termination risk (triggered by an arrival of any observable event $s = 1, ..., S$), and $\tilde{h}_{0r}(\cdot) = h_{0r}(H^{-1}(\cdot, Y_{n\cdot}))$ is the baseline hazard rate of risk r for all $r = 1, ..., S$, expressed in the units of "intrinsic time" $H(\cdot, Y_{n\cdot})$. Provided that the last integral in (4.7) is well approximated by the logistic functional form

$$\frac{A_r(t_0, \tau) \exp(Y_{nr})}{\sum_{s=1}^{S} A_s(t_0, \tau) \exp(Y_{ns})},$$

then the odds ratio (4.6) takes the form of conventional logistic function (adjusted for the passage of time) and forecasts can be based on the conventional multinomial logit estimates. On the other hand, if the main interest is to obtain the forecasts of instantaneous relative risks exactly t_0 seconds after the last observable event, one should directly substitute the covariate and baseline hazard estimates into the ratio

$$\frac{\widehat{h}_{0r}(t_0) \exp(\widehat{Y}_{nr})}{\sum_{s=1}^{S} \widehat{h}_{0s}(t_0) \exp(\widehat{Y}_{ns})}.$$

More generally, the probability forecasts can be based on the sample analogues

$$\widehat{P}_r(t_0, \tau; \widehat{Y}_{n\cdot}) = \int_{t_0}^{\tau} \widehat{h}_{0r}(u) \exp(\widehat{Y}_{nr} - \widehat{H}(u, \widehat{Y}_{n\cdot}))du, \quad r = 1, ..., S, \tag{4.8}$$

$$\widehat{\pi}_r(t_0, \tau; \widehat{Y}_{n\cdot}) = \frac{\widehat{P}_r(t_0, \tau; \widehat{Y}_{n\cdot})}{\sum_{s=1}^{S} \widehat{P}_s(t_0, \tau; \widehat{Y}_{n\cdot})} \tag{4.9}$$

for the expression (4.7) and formula (4.6).

The probability estimates (4.8) and (4.9) form the backbone of our forecasting results. When the probabilities are evaluated immediately after the moment when the limit order book gets updated, we can set $t_0 = 0$ in the formulas (4.8) and (4.9). The forecasts constructed by these formulas are dynamic in the sense that they take into account the passage of time as the internal ("epoch") time t_0 increases from zero until the moment of epoch termination.

4.6.2 Evaluating the Performance of Probability Forecasting Models

The predictive performance of any forecasting model is judged on its ability to assign to the forecast event a meaningful probability as a function of the observed covariates $\mathbf{z}$. Consider event A, which occurs in population with probability $P(A|\mathbf{z})$, conditioned on the vector of observable market characteristics $\mathbf{z}$. Let $O_n = \mathbf{1}\{A\}$, which is one or zero if A actually occurs or does not occur in observation n. When the outcome O of event is dichotomous, and the occurence or non-occurence of event can always be verified, the most popular measure used to assess the probability assignment $\widehat{P}(A; \mathbf{z})$ provided by a forecasting model is the *mean probability score* (Brier [16]) based on the population least-squares differences between forecast probabilities and the subsequent binary outcomes,

$$\text{Score}(\widehat{P}(A)) = E[(O - \widehat{P}(A; \mathbf{z}))^2] = E[E[(O - \widehat{P}(A; \mathbf{z}))^2|\mathbf{z}]]. \tag{4.10}$$

The mean probability score can be estimated by its empirical analogue

$$\frac{1}{N}\sum_{n=1}^{N}(O_n - \hat{P}(A;\mathbf{z}_n))^2, \tag{4.11}$$

called the *Brier score*, and denoted $\mathrm{BS}(\hat{P}(A))$. The Brier score is close to zero if the probability estimates provide near-perfect forecasts, and close to unity if the forecasts almost always miss the correct outcomes. While attractive from a practical standpoint, any scalar measure such as (4.11) is necessarily incomplete and gives only limited characterization of the forecasting performance. Despite other obvious drawbacks of the Brier score, such its failure to attach an explicit monetary performance measure to actions that could be taken by the decision maker in conjunction with the forecast probabilities, the Brier score and its modifications are used in the quality control and forecasting literature nearly universally.

Murphy [104] derived an instructive algebraic decomposition of (4.11), which is emphasizes the contributions of different aspects of the Brier score to prediction quality (see also Yates [127]). The Murphy decomposition is written for the binned data, calculated after assigning to J bins the observations of N outcomes O_1, O_2, ... O_N sorted in the increasing order of their forecast probabilities $\hat{P}(A;\mathbf{z}_1)$, $\hat{P}(A;\mathbf{z}_2)$, ... $\hat{P}(A;\mathbf{z}_N)$. Let $\overline{P}_j$ be the average probability forecast for observations in the jth bin ($j = 1, 2, ..., J$), $\overline{O}_j$ be the average conditional outcome for the subsample of outcomes O_n assigned to the jth bin, and $j(n)$ be the index of the bin corresponding to observation n. Then the discretized (or Sanders modified) Brier score

$$\mathrm{BS}_J(\hat{P}(A)) = \frac{1}{N}\sum_{n=1}^{N}(O_n - \overline{P}_{j(n)})^2,$$

can be decomposed as follows

$$\begin{aligned}
\mathrm{BS}_J(\hat{P}(A)) &= \frac{1}{N}\sum_{n=1}^{N}(O_n - \overline{P}_{j(n)})^2 = \frac{1}{N}\sum_{n=1}^{N}((O_n - \overline{O}_j) - (\overline{P}_{j(n)} - \overline{O}_k))^2 \\
&= \overline{O}(1-\overline{O}) - \frac{1}{N}\sum_{j=1}^{J} N_j(\overline{O}_j - \overline{O})^2 + \frac{1}{N}\sum_{j=1}^{J} N_j(\overline{P}_{j(n)} - \overline{O}_j)^2 \qquad (4.12) \\
&= \text{Uncertainty} - \sum_{j=1}^{J} \text{Resolution}_j + \sum_{j=1}^{J} \text{Reliability}_j \\
&= \text{Uncertainty} - \text{Resolution} + \text{Reliability}.
\end{aligned}$$

The first term in the decomposition (4.12), over which the forecaster has no control, represents the sample variance of binary outcomes. It gives the lower bound of the Brier score that can be achieved for the given sample of observations, which is obviously a number between zero and 0.25. The uncertainty term is close to zero only when the event in the sample either almost always occurs or does not occur.

The resolution term in (4.12) shows how well the group partition based on the assignment of forecast probabilities is able to discriminate between zero and unity. This term depends on the forecast probabilities through the sorting of the events making up the bin partition of the sample. If the binned observations have substantially different relative frequencies than the overall sample, the resolution term is large, which is desirable, since it leads to reduction of the overall score.

Finally, the reliability term in (4.12) describes the ability of the forecasts to calibrate, i.e., assign meaningful probabilities to outcomes in the sample. The reliability term is represented by the weighted average of the squared differences between the forecast probabilities and the relative frequencies of the forecast event in each bin. For very reliable forecasts, the relative frequency corresponding to the jth bin is very close to the forecast probability corresponding to the same bin.

When assessing the quality of a forecasting system, it is desirable to compare it with the performance of a "fence-sitter", i.e., a forecaster always presenting the sample probability of event A as the forecast probability. For such a forecaster, both resolution and calibration terms are zero, and the Brier score is $\overline{O}(1-\overline{O})$. The *Brier relative quality score* (also called the *skill score*) of a given forecast measures its performance relative to the "fence-sitter" and is defined as

$$\mathrm{SS}(\hat{P}(A)) = 1 - \frac{\mathrm{BS}(\hat{P}(A))}{\overline{O}(1-\overline{O})}. \tag{4.13}$$

The quality score can be decomposed into the quality scores measuring separately the reso-

lution and calibration ability of the forecaster relative to the "fence-sitter" as follows

$$\begin{aligned} \mathrm{SS}(\widehat{P}(A)) &= 1 - \frac{\mathrm{BS}(\widehat{P}(A))}{\overline{O}(1-\overline{O})} = 1 - \frac{\text{Uncertainty} - \text{Resolution} - \text{Reliability}}{\overline{O}(1-\overline{O})} \\ &= \frac{\text{Resolution} - \text{Reliability}}{\text{Uncertainty}} = \frac{\sum_{j=1}^{J}(\text{Resolution}_j - \text{Reliability}_j)}{\text{Uncertainty}}. \end{aligned}$$

A subsample of probability forecasts corresponding to the jth bin exhibits positive skill in the sense of equation (4.13) if its resolution term is larger in absolute value than the reliability term.

The Murphy decomposition shows that both discrimination and calibration are important in the evaluation of forecasting performance. When we collect a new sample of covariate vectors $\mathbf{z}_{N+1}, \mathbf{z}_{N+2}, ..., \mathbf{z}_{N+M}$, for which we assess the probabilities of event A using the forecasting model $\widehat{P}(A; \mathbf{z})$. If the model is well calibrated, which is indicated by a high reliability score, then for all values of p between zero and one it should be true that among the observations for which the model assigns the probabilities close to p, event A will be actually observed in $100p\%$ of cases. On the other hand, the quality of discrimination measured by the resolution term characterizes the ability of the model to discriminate between different probability regimes. The "sharpness", or resolution of the model depends on the grouping procedure implied by the rule that assigns observations to the bins.

The *Wilks reliability diagram* (Wilks [126]) provides a convenient way to visualize the Murphy decomposition. It plots the cumulative frequencies of the binned observations against the forecast probabilities associated with these bins. Since for a very reliable forecast $\overline{P}_{j(n)} \approx \overline{O}_j$, the contribution of the reliability component is close to zero, then the plot on the Wilks diagram is close to the 45-degree line. The resolution term measures the mean square distance of the reliability plot to the sample frequency of the event $\overline{O}$. A forecast has high resolution when the dispersion of the grouped frequencies $\overline{O}_j$ around $\overline{O}$ is as large as possible. Conversely, a forecast has no resolution when the event occurs with the frequencies $\overline{O}_j \approx \overline{O}$ for all bins $j = 1, 2, ..., J$.

The Wilks reliability diagram is also useful for detection of biased and overconfident forecasting rules. In particular, if the reliability plot lies below the 45-degree line, the

predicted probability of the event is higher than the actual frequency, and the forecast is upward biased. Conversely, if the reliability plot lies above the 45-degree line, the predicted probability of the event is lower than the actual frequency, and forecast is downward biased. Overconfidence is detected when the reliability plot is located above the 45-degree line in the range of low forecast probabilities (indicating that the event occurs more frequently than predicted when the forecaster thought it was unlikely) and located below the 45-degree line in the range of high forecast probabilities (indicating that the event did not occur as frequently as predicted when the forecaster was almost sure about the event). Similarly, one can detect biased forecasters, who tend to report overly conservative forecasts most of the time.

4.6.3 Evaluation of Forecasts by the Method of Relative Cost Analysis

The Brier score and Wilks reliability diagram are powerful tools for evaluating and comparing the performance of alternative probability forecasts. However, the Brier score does not give an answer the question of whether the desired quality of forecast has been achieved by the given forecasting model $\widehat{P}(A;\mathbf{z})$. A useful diagnostic that can be aligned more closely with the value of forecast for a potential user is the *skill score*, which was introduced originally in the signal detection theory and now widely used in meteorology and decision sciences.

First, consider the case of a deterministic forecast. Over a large sample of observations, we can form the contingency table associated with a given forecasting model

	Forecast $\overline{A}$	Forecast A
$\overline{A}$ occurs	$\pi_{00} = \Pr\{\overline{A}$ & Forecast $\overline{A}\}$	$\pi_{01} = \Pr\{\overline{A}$ & Forecast $A\}$
A occurs	$\pi_{10} = \Pr\{A$ & Forecast $\overline{A}\}$	$\pi_{11} = \Pr\{A$ & Forecast $A\}$

Based on these values, the *hit rate* H and the *false alarm rate* F are defined as

$$H = \frac{\pi_{11}}{\pi_{10} + \pi_{11}} \quad \text{and} \quad F = \frac{\pi_{01}}{\pi_{00} + \pi_{01}}.$$

Similarly, the hit and false alarm rates for probability forecast $\widehat{P}(A;\mathbf{z})$ are defined as follows. Assume that event A occurs if the probability forecast $\widehat{P}(A;\mathbf{z})$ is greater or equal than some threshold π^*, and will not occur otherwise. Define the *hit rate* for a family of probability forecasts associated with alternative levels of threshold probability $\pi^* \in (0;1)$ as

$$H(\pi^*) = \frac{1}{\overline{O}} \times \frac{1}{N} \sum_{n=1}^{N} O_n \mathbf{1}\{\widehat{P}(A;\mathbf{z}_n) \geq \pi^*\}.$$

The *false alarm rate* for the family of probability forecasts is defined similarly as

$$F(\pi^*) = \frac{1}{1-\overline{O}} \times \frac{1}{N} \sum_{n=1}^{N} (1-O_n) \mathbf{1}\{\widehat{P}(A;\mathbf{z}_n) \geq \pi^*\}.$$

The relative operating characteristic curve (ROC; Swets [124]) is defined as a plot of $H(\pi^*)$ against $F(\pi^*)$ for $\pi^* \in (0;1)$. The forecaster's skill may be measured by the area under the ROC curve. A perfect forecast has this area equal to unity, while the area equals 0.5 for a statistical forecaster with the hit and false alarm rates proportional, respectively, to the frequencies of occurence and non-occurence of event A in the sample. The forecast is more valuable for the ROC curves that are more concave and bent to the upper-left corner of the diagram.

To show how the monetary measures can be assigned and used to evaluate the quality of forecasts, consider a risk-neutral decision maker (trader or, more generally, another potential user of forecasts) who takes some specific action depending on the likelihood assigned to event A by the forecasting model.[31] Assume the expenses associated with each combination of the trader's action/inaction and event occurence/non-occurence are given by the contingency table

	Don't take action	Take action
$\overline{A}$ occurs	$c_{00} = 0$	$c_{01} = C$
A occurs	$c_{10} = L$	$c_{11} = C$

[31] The risk-neutrality assumption is not important in the present context, if the trader's objective function can be modified appropriately to incorporate the measures of risk based on the higher-order moments (variance, skewness, etc.) However, our analysis does not extend to the case when the risk experienced by the trader cannot be expressed in terms on finite-order moments.

Assume that the decision maker would like to minimize his or her overall expected expenses.

If only the sample frequency $\overline{O}$ of event A is known, and the probability forecast is unavailable, the decision maker can choose only among two types of decision, either always or never take action. The expected expenses in both cases are, respectively,

$$\begin{aligned} E[\text{Cost} \mid \text{Take action}] &= C, \quad \text{and} \\ E[\text{Cost} \mid \text{Don't take action}] &= L\overline{O}. \end{aligned}$$

Therefore, the expected expense of the decision maker per unit of potential losses that could be experienced when no action is taken is expressed as follows

$$\begin{aligned} M &= \frac{E[\text{Cost}]}{L} = \frac{\pi_{10}L + (\pi_{01} + \pi_{11})C}{L} = \pi_{10} + \frac{C}{L}(\pi_{01} + \pi_{11}) \\ &= (1 - H)\overline{O} + F \cdot \frac{C}{L}(1 - \overline{O}) + H \cdot \frac{C}{L}\overline{O}. \end{aligned}$$

For a perfect deterministic forecast, $H = 1$ and $F = 0$, so that $M_{\text{perfect}} = \frac{C}{L}\overline{O}$. On the other hand, if the decision maker knows only the sample probability $\overline{O}$ of the event occurence, and nothing else, then M will be minimized if action is taken whenever $C/L < \overline{O}$. Therefore, given only the statistical knowledge of probability distribution, the expected expense is $M_{\text{stat}} = \min\left(\frac{C}{L}; \overline{O}\right)$.

Define the *information value* V of forecast as the reduction in the expected expense per unit loss M relative to M_{stat}, normalized by the maximum possible reduction $M_{\text{stat}} - M_{\text{perfect}}$ associated with the perfect forecast,

$$V = \frac{M_{\text{stat}} - M}{M_{\text{stat}} - M_{\text{perfect}}}.$$

It is obvious from this definition that V is always a number between zero and one.

Now suppose $H(\pi^*)$ and $F(\pi^*)$ are the hit and false alarm rates expressed as functions of the probability threshold π^*. For a given ratio of hedging cost C to the potential loss L if the action is not taken, the user's value of the forecast is

$$V_{\text{optimal}} = \max_{\pi^*} V(\pi^*), \tag{4.14}$$

where $V(\pi^*)$ is calculated as follows

$$V(\pi^*) = \frac{\min\left(\frac{C}{L}, \overline{O}\right) - F(\pi^*)\frac{C}{L}(1-\overline{O}) + H(\pi^*)\left(1-\frac{C}{L}\right)\overline{O} - \overline{O}}{\min\left(\frac{C}{L}, \overline{O}\right) - \frac{C}{L}\overline{O}}. \tag{4.15}$$

This is an absolute dollar value associated with the forecasting rule $\hat{P}(A; \mathbf{z})$, provided that the ratio of the dollar values C and L is known. Even though the exact implementation of this general recipe ultimately depends on the context of the specific problem and ideally should be tied to the particular hedging strategy that could be used by the decision maker, it is important to provide the trader with a full family of decision rules $\pi_{,} \geq \pi^*$, $\pi^* \in (0; 1)$, since the false alarm and the hit ratios ultimately depend on the threshold probability π^*, according to (4.14) and (4.15).

4.6.4 Forecasting the Probabilities of Buyer- and Seller-Initiated Transactions

Direction of next transaction on the market is an issue of ultimate importance for limit order traders who are intrinsically interested in fast execution of their limit orders at favorable prices. The prevalence of buyer- or seller-initiated transactions on the market may be intimately, but non-trivially related to the appreciation or depreciation of exchange rates as it might signal the informational advantage of the counterparties initiating the trades. Investigation of dynamic links between the order flow (which is the difference between the number of buyer- and seller-initiated transactions), appreciation or depreciation of transaction price, and various measures of limit order book liquidity is a dominant topic in the modern empirical microstructure literature (see the brief review in Chapter 1) and one of the main objectives of the present research. Therefore it would be natural to verify the ability of the model to issue the warnings about unusually high or low probabilities of transactions on the sell and buy side of the limit order book.

First, we evaluate the ability of the PCA factor model with five principal components (section 4.4), based on the estimates of risk indices $\tilde{Y}_r = \mathbf{U}'\boldsymbol{\gamma}_r$ reported in Tables 4.10, 4.11,

and 4.12, to provide a good in-sample fit between the realized and forecast probabilities of buyer- and seller-initiated transactions. As we have some freedom to select the number of event categories, we choose a coarser event classification scheme than the one with $S = 14$ observable risks used in the original model estimated in sections 4.2–4.4. Even though we started with $S = 14$ observable risks (marked by single and double stars in Table 4.1), their number was later reduced by pooling the events of types A1, A2, A3 in the "Sell Trade" category AA and the events of types B1, B2, B3 in the "Buy Trade" category BB, and by collapsing the events of types A4 and A5 into the "Ask Price Improvement" category AP+ and the events of types B4 and B5 in to the "Bid Price Improvement" category BP+. The composition of the event classes A6, B6, AC6, and BC6 corresponding, respectively, to "Ask Depth Improvement", "Bid Depth Improvement", "Ask Touch Cancellation", and "Bid Touch Cancellation" events were unchanged but the categories were renamed as AD+, BD+, A−, and B−, respectively. The modified event classification scheme with $S = 8$ observable categories is shown in Table ??.

Figure 4.17 gives an example of dynamic evolution of forecast probabilities for buyer- and seller-initiated transactions. The sample period shown on the graph is chosen to be identical to the one used to demonstrate the evolution of bid, ask, and transaction prices on Figure 4.9. The probability forecasts are based on formulas (4.7) and (4.6) as explained in the last paragraph under the assumption that no events occured at least for one second after the previous observable limit order book event, which explicitly takes into account the reaction time of the potential forecast user. The first seven minutes of the sample period captured on Figures 4.9 and 4.17 was the period of heavy buyer pressure accompanied by a rapid growth of market price by almost ten tick points. This was also the period when our competing risks model produces the probability forecasts which are much larger for buy than for sell transactions. The last portion of the sample covers the period when the price stabilizes just under the new level DEM 1.7575 per US dollar and is supported by a fairly strong "resistance" on the sell side. The sell and buy transaction probability forecasts for this period are approximately equal to each other on average, even though the fraction of

epochs terminated by seller-initiated transactions is slightly higher than predicted in this subperiod.

The quality of probability forecasts based on formulas (4.7) and (4.6) is evaluated using Wilks reliability diagrams reviewed in subsection 4.6.2. The graphs on Figure 4.18) show the reliability plots for the forecasts of buyer- and seller-initiated transactions made under assumption that no events occured at least for one second after the previous observable limit order book event. The first two graphs show the reliability plots for the probability forecasts of buyer-initiated transactions (graph 1) and seller-initiated transactions (graph 2) matched with the frequencies of buyer- and seller-initiated transactions observed over the estimation period of the first three days (31391 epochs). The graphs in the lower portion of Figure 4.18 show the reliability plots for the probability forecasts of buyer- and seller-initiated trades (graphs 3 and 4, respectively) when the forecast probabilities are matched with the corresponding transaction frequencies observed in the out-of-sample period covering the last two days (19385 epochs). Both sell and buy event forecasts offer high resolution with a broad range of covered probabilities. The forecasts also have good reliability properties as they do not reveal a strong tendency to deviate systematically from the main diagonal line that corresponds to perfect reliability. Even though the probability forecasts slightly underpredict the probabilities of buy and sell trades when the predicted probabilities are very small, this tendency towards overconfidence (underprediction of unlikely seller- and buyer-initiated transactions) is fairly weak. The overconfidence (underprediction) bias for the rare events is slightly stronger when the forecasts are evaluated out-of-sample, which can be easily detected from the deviations of left tails of reliability plots on the last two graphs from the perfect reliability line. This might be an issue of concern if the forecasts are ultimately used to measure the risk of transactions that might occur in the undesirable direction. However, even with this small caveat that must be taken into consideration by the ultimate users of forecast, the out-of-sample performance of the competing risks model turns out to be surprisingly good.

The high degree of persistence detected on the diagrams for cross-correlograms of PCA

factors (Figures 4.12 and 4.13) suggests that at least some of the factors are unlikely to change much over a relatively short time period that might be covering more than one epoch. Therefore, it may be interesting to compare the quality of forecasting rules that rely on the principal components derived from the competing risks model and the performance of some benchmark forecasting rules based on the current and lagged directional indicators. In our last exercise, we evaluate the quality of forecasts for the event that the next transaction in the limit order book will be initiated by buyer or seller within the next 30 seconds, or no trade will be recorded in the next 30 seconds since the time of forecast. The "naïve" benchmark used in our comparisons will be based on the trinomial logit regression of buy, sell, or no-trade indicator with the covariates given by the signs of ten most recent transactions (± 1 if transaction was initiated by seller/buyer, and zero if no transaction occurred). The only information used to predict the direction of next transaction in this simple "directional momentum" model is the direction of the last several trades. The performance of this simple forecasting model will be compared with the multinomial logit model with the index of covariates given by the first five PCA factors.[32]

The two diagrams in the left column on Figure 4.19 show the reliability plots for the forecast of event that the next transaction in limit order book will be buyer-initiated and occur in 30 seconds after the time of forecast. Similarly, the two diagrams in the right column on Figure 4.19 show the reliability plots for probability forecasts of event that the next transaction in limit order book will be seller-initiated and occur in 30 seconds since the time of forecast. The reliability plots in the first row are based on the trinomial logit regression with the covariates given by the five competing risks PCA factors. The reliability plots in the second row are based on the benchmark forecasting model based on the trinomial logit regression with the covariates given by signs of last 10 transactions (± 1 if transaction

[32]Since the efficient algorithm generating multistep forecasts in the competing risks framework is currently unavailable, and would require, in particular, the dynamic multistep forecasts of the covariate structure or the factor structure, we come up with a shortcut solution that ignores the dynamic properties of covariates but keeps intact the general structure of the model. The development of a truely dynamic forecasting model is left for future investigation.

was initiated by seller/buyer). All diagrams are produced from the sample covering the estimation period 6 am to 5 pm GMT on October 6–8, 1997.

Since the patterns of plots on the upper two diagrams on Figure 4.19 match very closely the main diagonal, we can conclude that the PCA factor model fits the data much better than the "naïve" benchmark based on the directional indicators. Moreover, the PCA factor model also has better discrimination properties as the range of probability forecasts based on this model is substantially wider. Apart from the slight downward bias of probability forecasts for small probabilities, the PCA factor model appears to provide a better fit to the empirical data in comparison to the alternative model. Now we check whether the PCA factor model is capable of giving the warnings about unusually high (or unusually low) probabilities that the next trade will be initiated by buyer or seller.

Figure 4.20 is analogous to Figure 4.19 in all respects, except that it shows the reliability plots of forecasts matched to the data from the out-of-sample period 6 am to 5 pm GMT on October 9 and October 10, 1997. Again, the PCA factor model delivers the forecasts that are more reliable and have much better resolution properties than the forecasts based on the alternative "momentum-based" model. Figure 4.20 also indicates a problem that seems to deteriorate with the forecast horizon.[33] In particular, the forecasts based on our model appear to be overconfident in the sense that the model underpredicts low probability events (left tails on the reliability plots are bent upward) and overpredicts high probability events (right tails on the reliability plots are bent downward), which may also serve as an evidence of rapid reaction of market participants to extremely high or low probability signals. As the quality of information contained in the PCA components quickly deteriorates as the forecast horizon increases, continuous monitoring of the relevant market information and updating the risk indices are crucial conditions for the success in this highly competitive segment of foreign exchange market.

[33]The reliability plots of one-minute-ahead probability forecasts (not reported here) have similar properties, except for the more substantial biases of these plots for high and low forecast probabilities.

4.7 Appendix: Graphs and Tables

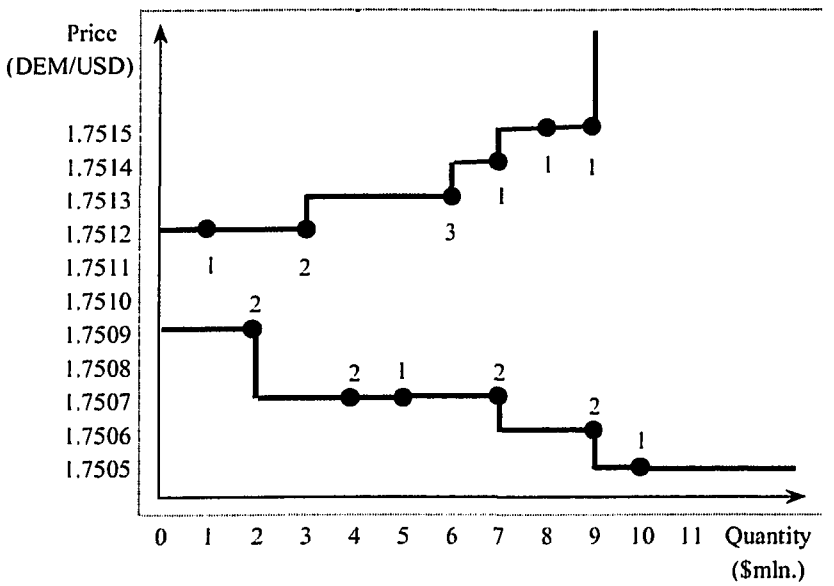


Figure 4.1: Example of supply and demand curves from Reuters D2000-2 dealing system

Figure 4.1 displays the state of the Reuters D2000-2 electronic limit order book at a particular moment in time. The ersatz supply and demand curves on the market for US dollars are represented by limit sell and buy orders waiting their execution. Two limit orders, for one and two million dollars, are available at the best market sell price of DEM 1.7512 per dollar. Additionally, there is one limit order for three millions at the ask price of DEM 1.7513, one limit order for one million at the ask price of DEM 1.7514, and two limit orders for one million each at the ask price of DEM 1.7515 per US dollar. On the bid side, there is one limit order to purchase two millions at the best market buy price of DEM 1.7509 per dollar, which is followed (in the order of priority) by three limit orders for two million, one million, and two million dollars at the bid price of DEM 1.7507, a limit order for two million at the bid price of DEM 1.7506, a limit order for one million at the bid of DEM 1.7505, and another large limit order at the same price (the size of this buy limit order is unclear from the graph). Note that traders observe only the best market buy and sell prices DEM 1.7509 and DEM 1.7512, along with the quantities $2 mln. and $3 mln., respectively, on their trading screens.

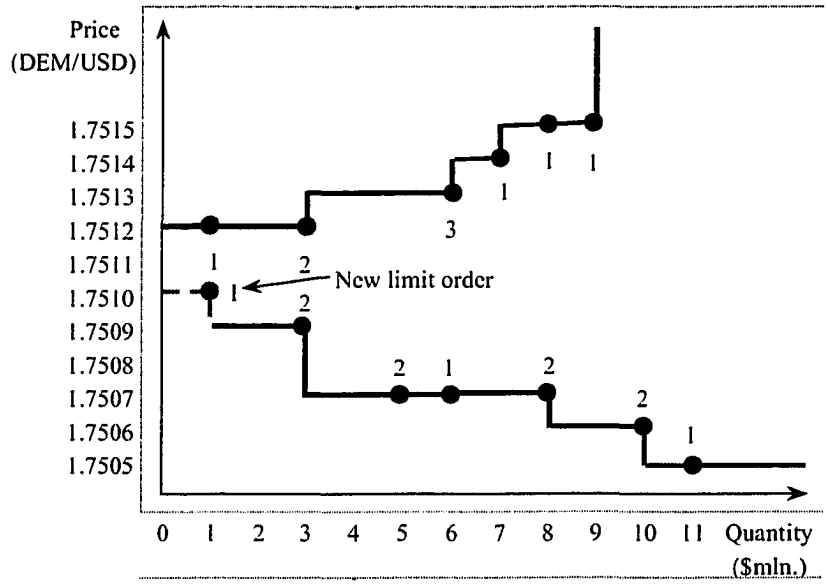


Figure 4.2: Supply and demand curves after the price improvement on demand side

Figure 4.2 displays the change in the state of the Reuters D2000-2 electronic limit order book shown on Figure 4.1 after arrival of a new limit order to purchase one million dollars at the bid price DEM 1.7510 per dollar. The arrival of new limit order leads to reduction of the market bid-ask spread, and shifts the prior ersatz demand curve to the right. The improved market liquidity associated with such an event validates the term "price improvement". However, the quantity (depth) available at the improved bid price DEM 1.7510 per dollar is smaller than the bid depth at the previous bid quote DEM 1.7509 per dollar on Figure 4.1.

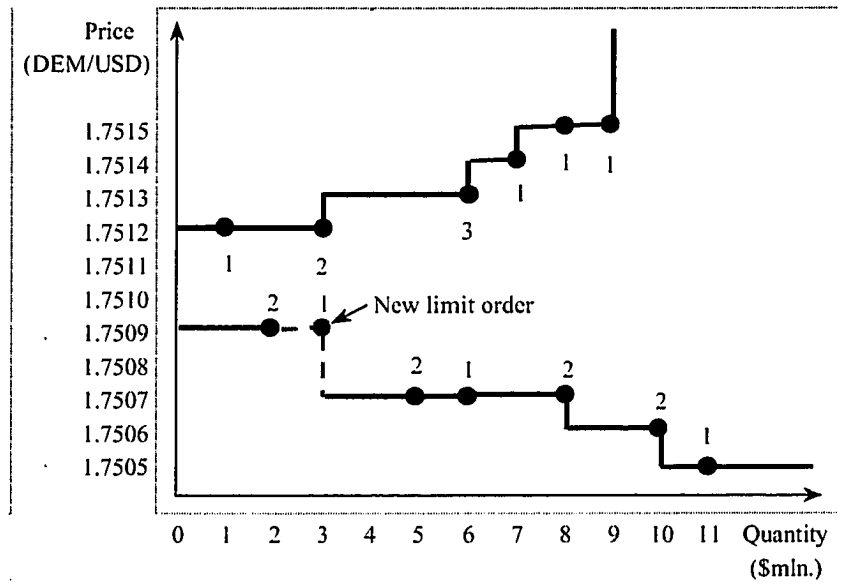


Figure 4.3: Supply and demand curves after the depth improvement on demand side

Figure 4.3 displays the change in the state of the Reuters D2000-2 electronic limit order book shown on Figure 4.1 after arrival of a new limit order to purchase one million dollars at the bid price DEM 1.7509 per dollar. The arrival of new limit order does not change the market bid-ask spread, but it shifts the portion of prior ersatz demand curve below the best bid price DEM 1.7509 per dollar to the right. The improved market depth on the bid side associated with such an event validates the term "depth improvement".

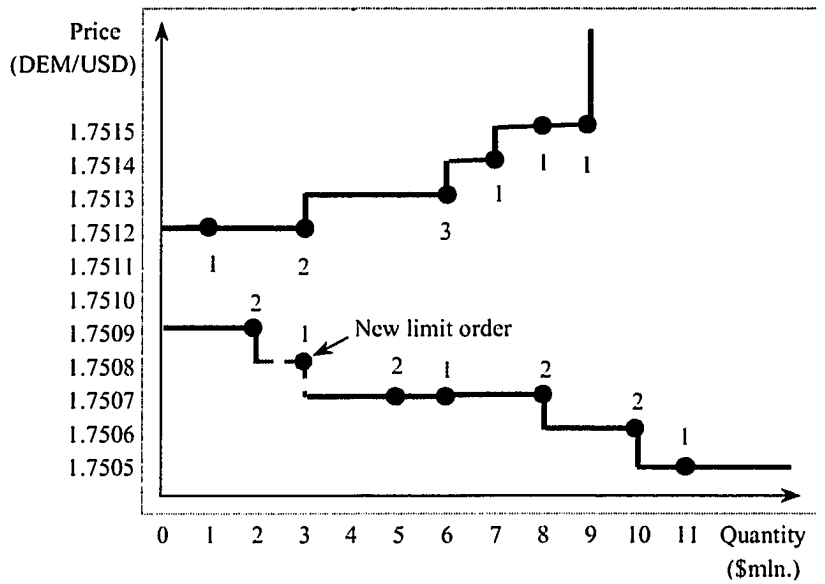


Figure 4.4: Supply and demand curves after the bid arrival one tick below the touch

Figure 4.4 displays the change in the state of the Reuters D2000-2 electronic limit order book shown on Figure 4.1 after arrival of a new limit order to purchase one million dollars at the bid price DEM 1.7508 per dollar, which is lower than the best bid price DEM 1.7509 available on the market. The arrival of new limit order does not change the public information on the Reuters D2000-2 trading screens, in particular, it does not affect the size of the bid-ask spread and the market depth at the touch (best bid quote). However, the market liquidity improves in a broader sense as the market depth one tick below the best bid quote increases.

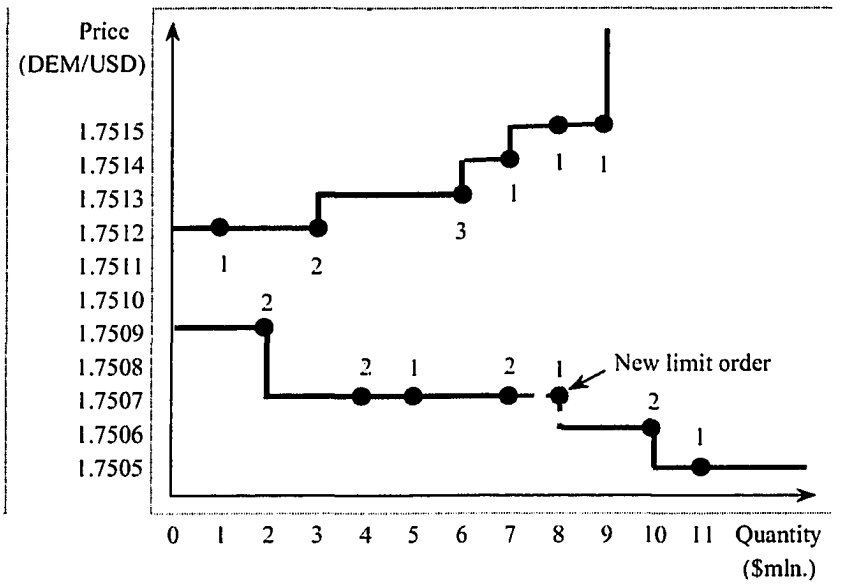


Figure 4.5: Supply and demand curves after the bid arrival two ticks below the touch

Figure 4.5 displays the change in the state of the Reuters D2000-2 electronic limit order book shown on Figure 4.1 after arrival of a new limit order to purchase one million dollars at the bid price DEM 1.7507 per dollar, which is two ticks below the best bid price DEM 1.7509 available on the market. The arrival of new limit order does not change the public information on the Reuters D2000-2 trading screens, since the market event is associated with improvement of market liquidity deep on the bid side of the limit order book.

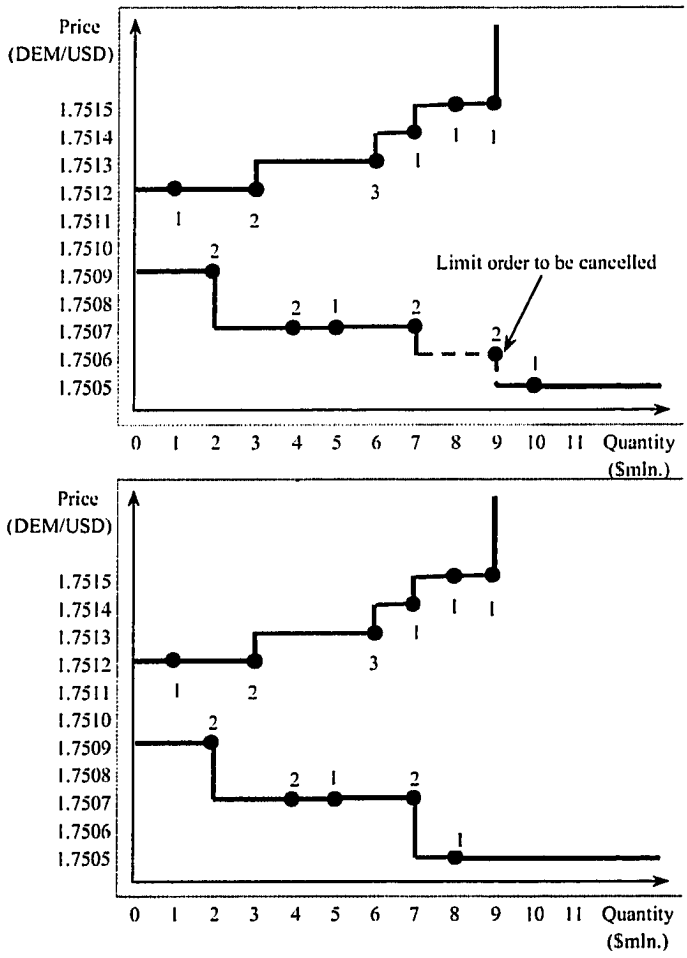


Figure 4.6: Effect of subsidiary limit order cancellation on the demand curve

Figure 4.6 displays the change in the state of the Reuters D2000-2 electronic limit order book shown on Figure 4.1 after cancellation of the subsidiary limit order to purchase two million dollars at the bid price DEM 1.7506 per dollar. The cancellation event does not affect public information on the Reuters D2000-2 trading screens, as it is associated with the deterioration of market liquidity deep inside the bid side of limit order book.

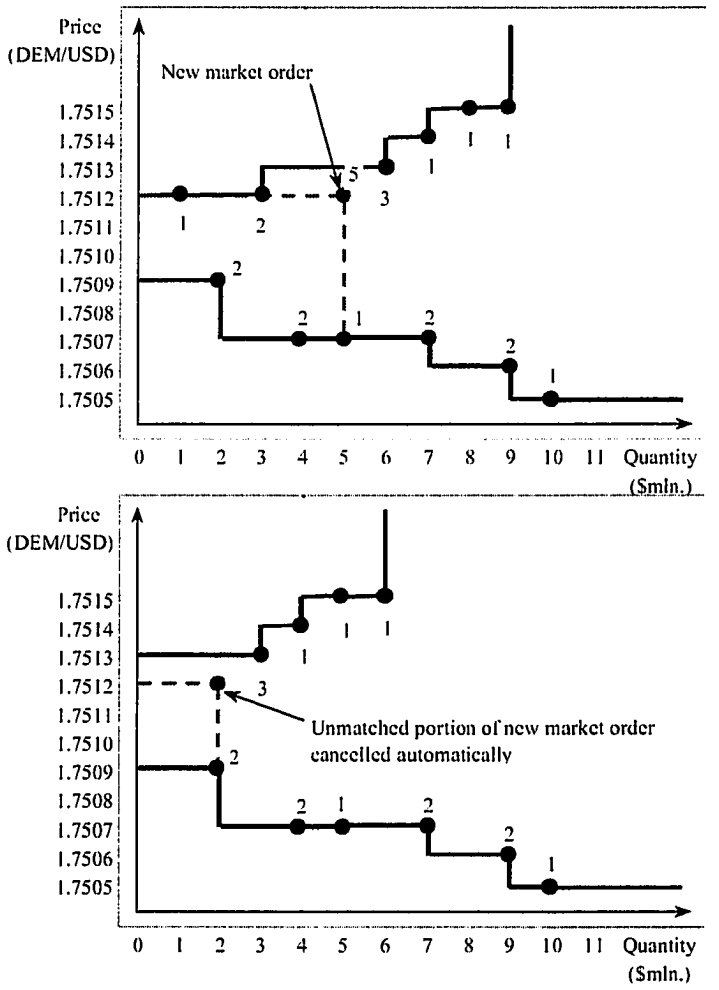


Figure 4.7: Effect of the large market buy order on the supply and demand curves

Figure 4.7 displays the change in the state of the Reuters D2000-2 electronic limit order book shown on Figure 4.1 after arrival of a market order to purchase five million dollars. which was submitted at the best ask market price of DEM 1.7512 per dollar. Since the quantity available at this price is only three million dollars. part of demand for liquidity created by the new market order arrival is not satisfied. The unmatched portion of the market order gets cancelled immediately. while the best ask price goes up one tick to DEM 1.7513 per dollar.

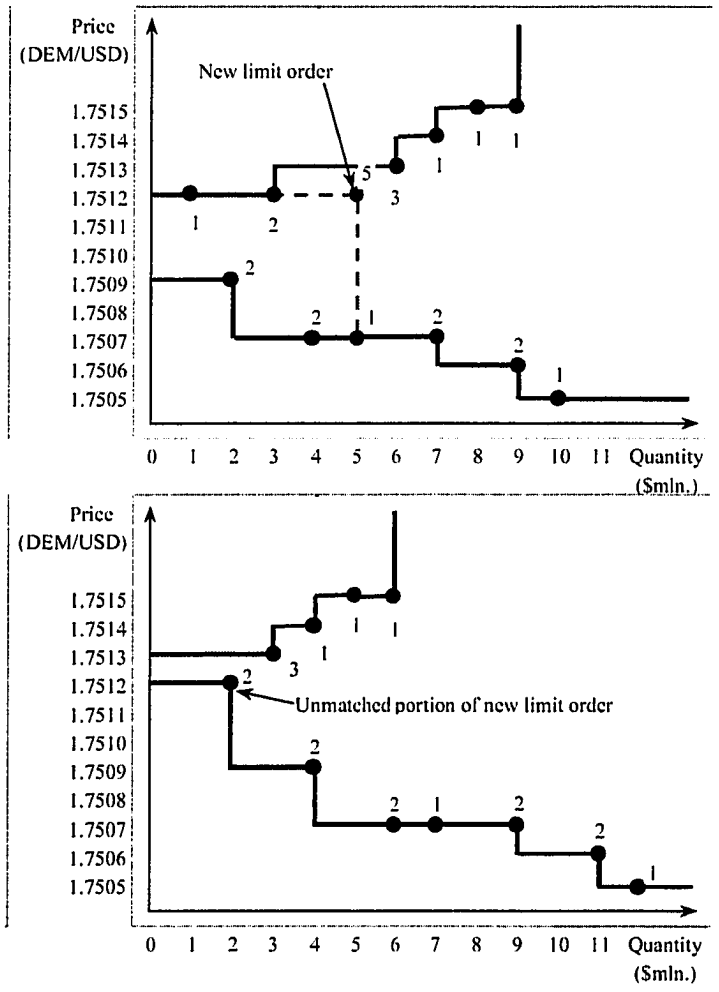


Figure 4.8: Effect of the large limit bid on the demand and supply curves

Figure 4.8 displays the change in the state of the Reuters D2000-2 electronic limit order book shown on Figure 4.1 after arrival of a limit order to purchase five million dollars at the limit order price which coincides with the best ask market price of DEM 1.7512 per dollar previously available on the market. The situation is analogous to the submission of buy market order (Figure 4.7), except that the unmatched portion of the arriving aggressive limit order remains on the limit order book, leading to the improvement of the best bid price by three ticks to DEM 1.7512 per dollar. The best bid price moves one tick up to DEM 1.7513 per dollar.

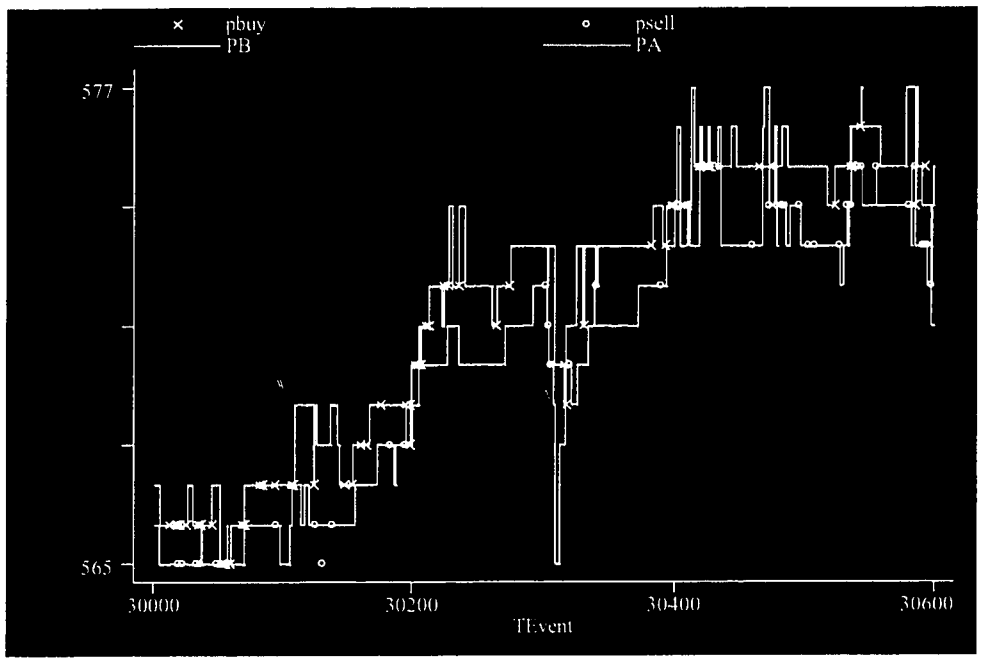


Figure 4.9: A small subsample of market bid and ask quotes and transactions in Reuters D2000-2 trading system

Figure 4.9 shows a small subsample of continuously sampled best market bid and ask quotes, as well as the times and prices of buyer- and seller-initiated transactions (marked by small white crosses and knots, respectively). All prices were obtained by matching limit and market orders from the original Reuters D2000-2 data set. The sampled time period covers the trading hours 8:20 to 8:30 a.m. GMT on Monday, October 6, 1997.

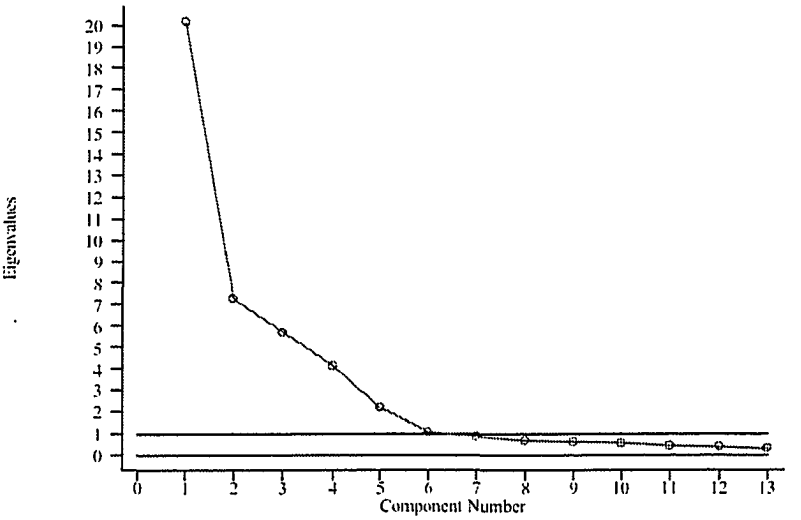


Figure 4.10: Eigenvalues of principal components for the competing risk indices

Figure 4.10 displays the plot of eigenvalues corresponding to the first $Q = 13$ PCA factors of the competing risk indices $r = 1, ..., R$. Since only five of these eigenvalues are unambiguously above the horizontal line $\lambda = 1$, which suggests PCA factors 1 through 5 can be treated as independent pervasive components driving the market dynamics. Even though factors 6 and 7 are only marginally significant, they are retained to prevent wrongful exclusion of additional, marginally significant factors.

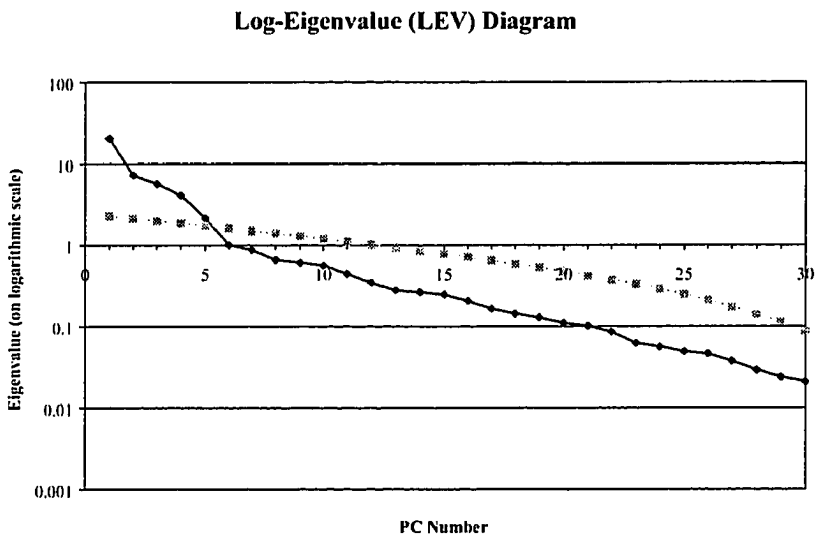


Figure 4.11: Log-eigenvalue (LEV) diagram for principal components of the competing risk indices

Figure 4.11 displays the LEV diagram, which is obtained by replotting Figure 4.10 on the logarithmic scale. The darker line plots the eigenvalues corresponding to the PCA factors of the competing risk indices $r = 1, ..., R$. Deviations of plot from the linear pattern on the left for low eigenvalues suggests that retention of $Q = 5$ PCA factors for further analysis would be appropriate. The choice of $Q = 5$ is confirmed by the comparison of the actual eigenvalues with the 95% confidence bounds shown on the graph in pink color that are obtained by a bootstrap procedure under the null hypothesis that all eigenvalues corresponding to the Qth and higher-order PCA factors are equal to each other.

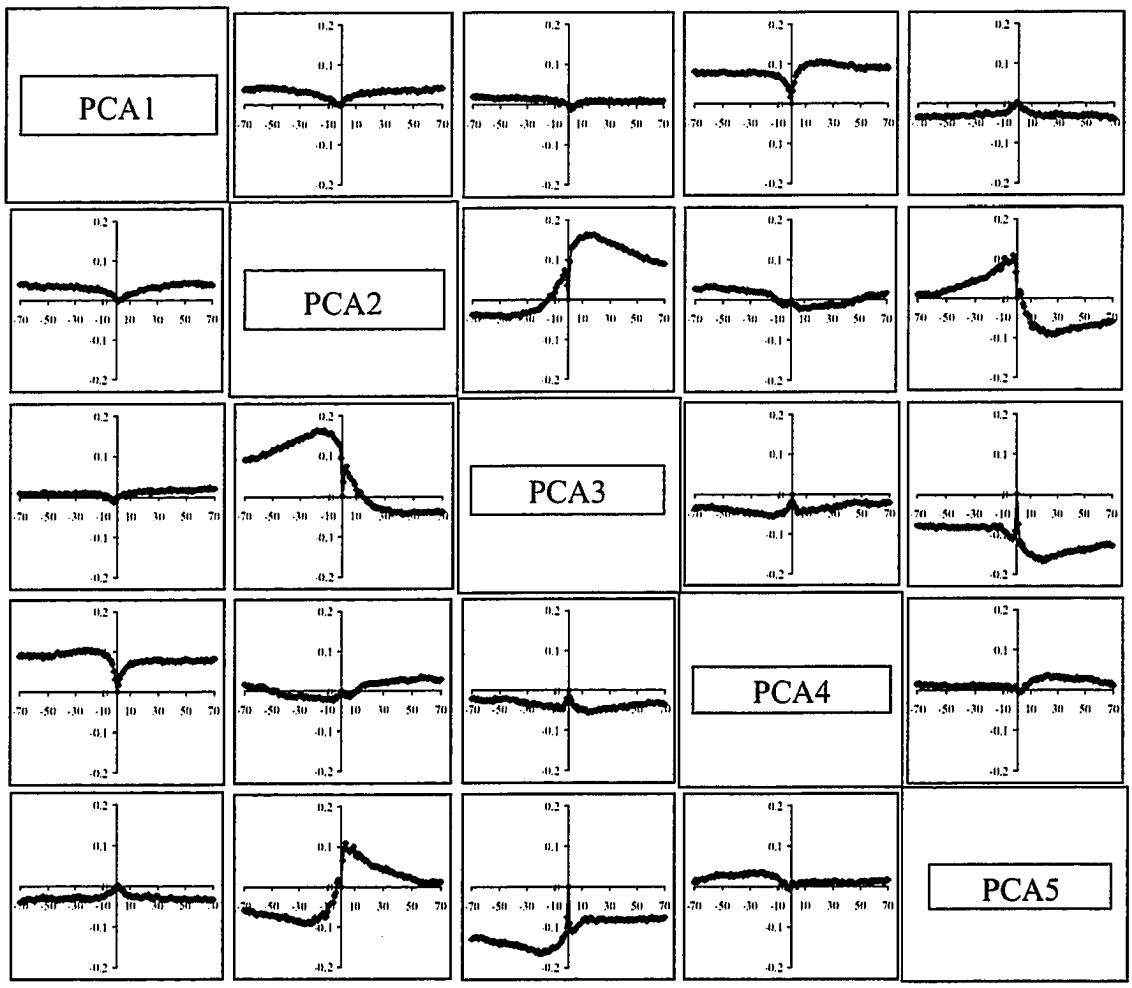


Figure 4.12: Cross-correlograms of PCA factors 1 to 5 driving the limit order book dynamics

Figure 4.12 displays the estimated off-diagonal terms of cross-correlograms of the first five PCA factors associated with the aggregate risks of limit order book events. Cell (i, j) of the chart contains the estimated cross-correlogram of factors i and j defined by the formula

$$R_{ij}(h) = \widehat{\mathrm{Corr}}(f_{i,n},\ f_{j,n+h}),$$

where h is the lead (forward shift) of factor j relative to factor i measured by the number of epochs (the definition of epoch is given in section 3.2). All calculations are performed for the values of h between -70 and 70 and based on the subsample covering the liquid trading hours (6 a.m. to 5 p.m. GMT) on October 6–8. 1997.

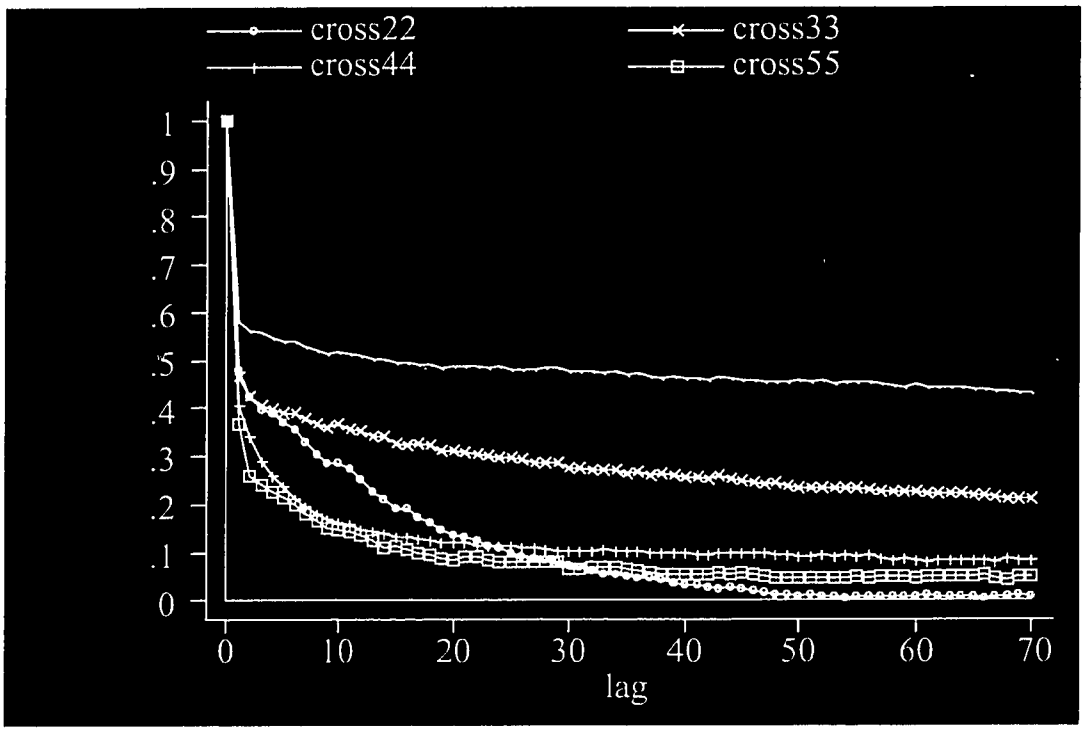


Figure 4.13: Autocorrelograms of the five PCA factors driving the limit order book dynamics

Figure 4.13 displays the estimated autocorrelation functions (autocorrelograms) for the first five PCA factors associated constructed from the risk indices of the aggregated limit order book activity. Autocorrelograms of factor 1 (shown by the unmarked solid line) and factor 3 (highlighted by "x" symbols) demonstrate slow rate of decline with the lag order, which is a clear evidence of long memory and potential nonstationarity. Autocorrelograms of factors 4 and 5 (highlighted by pluses and squares, respectively) also decline relatively slowly which serves as an evidence of long memory. The autocorrelogram of factor 2 (highlighted by circles) rapidly declines to zero with the lag order and becomes indistinguishable from zero at lag 50. All calculations are performed using the subsample covering only the liquid trading hours 6 am to 5 pm GMT on October 6-8. 1997.

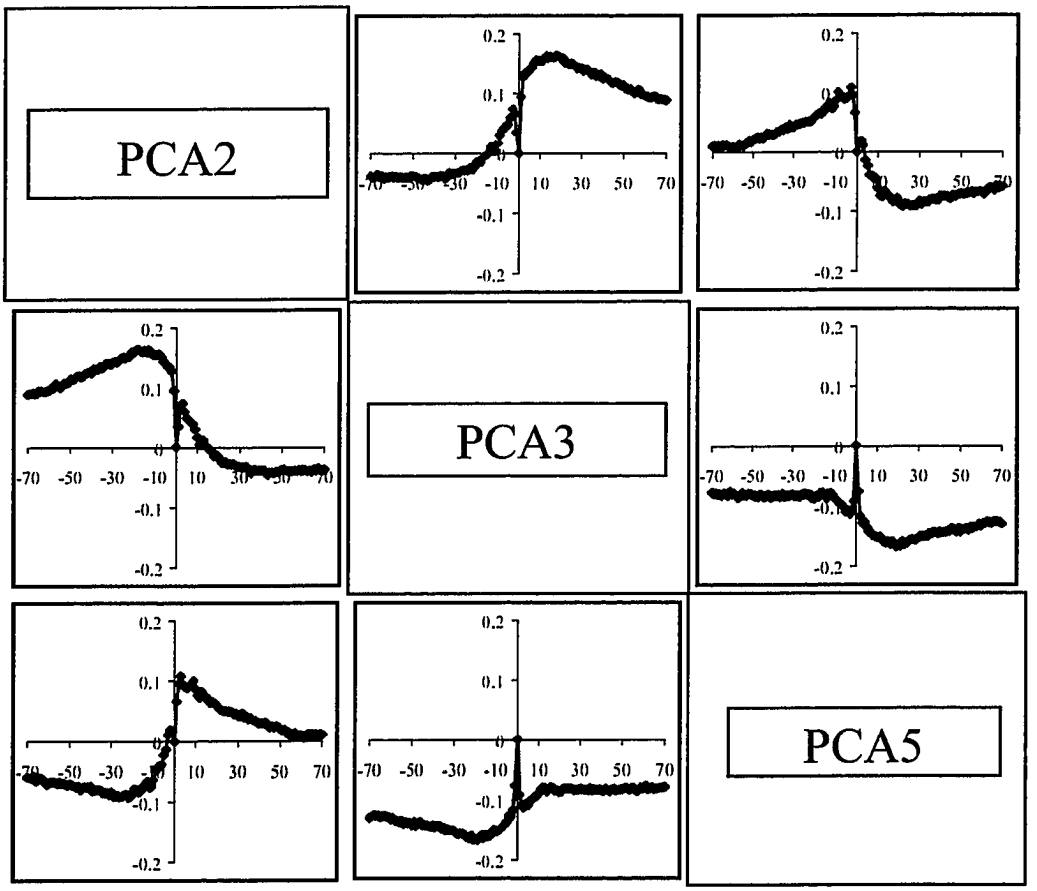


Figure 4.14: Cross-correlograms of activity imbalance, buyer pressure, and bull market momentum (PCA factors 2, 3 and 5)

Figure 4.14 highlights the portion of the estimated cross-correlogram 4.12 that displays the interaction of the second, third, and fifth PCA factors. These factors are identified in the text as the limit order book activity imbalance (factor 2), the buyer pressure (factor 3), and the bull market momentum (factor 5). Cell (i, j) of the chart contains the estimated cross-correlogram of factors i and j defined by the formula

$$R_{ij}(h) = \widehat{\mathrm{Corr}}(f_{i,n},\ f_{j,n+h}),$$

where h is the lead (forward shift) of factor j relative to factor i measured by the number of epochs (see definition of epoch in section 3.2). All calculations are performed for the values of h between -70 and 70 and based on the subsample covering the liquid trading hours (6 a.m. to 5 p.m. GMT) on October 6-8, 1997.

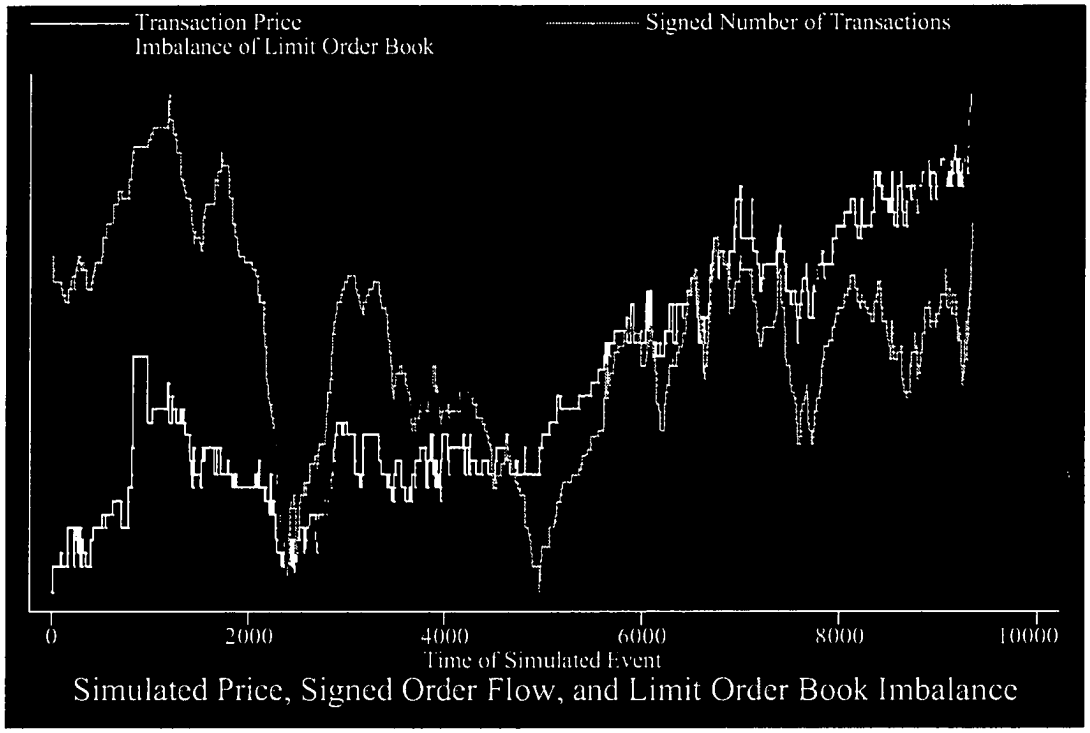


Figure 4.15: A representative simulated history of transaction prices, signed order flow, and imbalance of the limit order book

Figure 4.15 displays the simulated history of transaction prices (in a white pattern), the limit order book imbalance measured as the difference between the total dollar value available for sale and the total dollar value available for purchase (in a darker grey pattern), and the signed cumulative flow of trades (in a lighter grey pattern) over the period of 10.000 seconds. The simulations use the constrained smoothed-censored plan described in section 3.6 that was applied to the system of competing risks calibrated to the covariate coefficients reported in Tables 4.3–4.8.

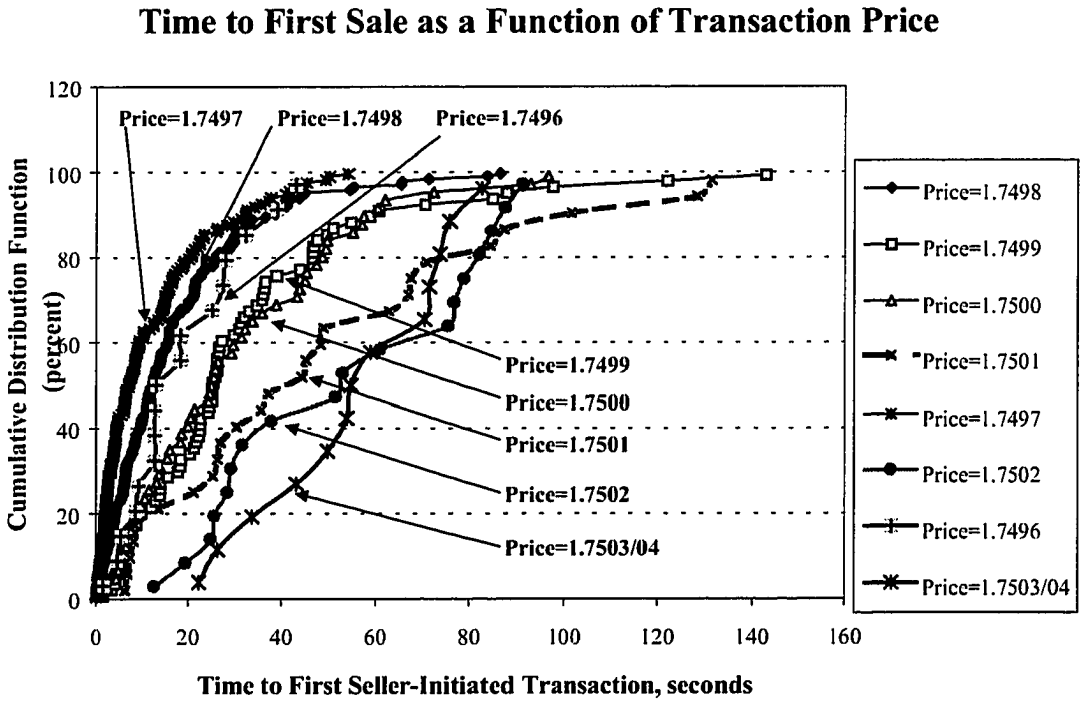


Figure 4.16: Cumulative distribution functions of time to first seller-initiated transaction as a function of initial state and transaction price

Figure 4.16 illustrates the asymmetry in the dependence of next seller-initiated transaction price on the initial conditions and the time elapsed since last sale. The alternative curves show the plots of simulated cumulative distribution functions for the time to first seller-initiated transaction given that this trade occurs at the specified price. The total number of simulations is 500. All simulations are performed for a fixed simulated history of trading (dominated by seller-initiated transactions) as described in subsection 4.5.2.

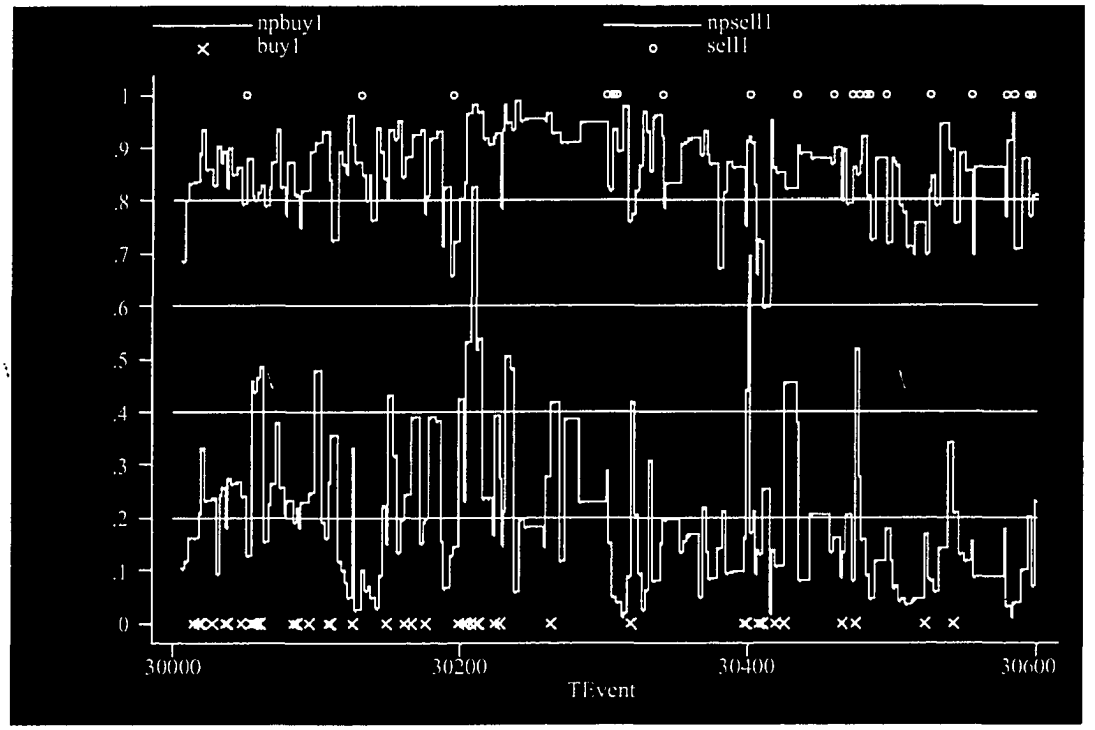


Figure 4.17: A sample of one-step-ahead forecast probabilities of buyer- and seller-initiated transactions

Figure 4.17 shows a small subsample of one-step forecast probabilities of buyer- and seller-initiated transactions made after at least one second elapsed since the previous event, as well as the times of actual buyer-initiated transactions (shown by crosses) and seller-initiated transactions (shown by knots). The forecasts are based on the version of PCA factor competing risks model with $S = 8$ types of observable risks and the covariate structure comprised by $Q = 5$ PCA factors as described in subsection 4.6.4. The sample period covers the episode 8:20 to 8:30 a.m. GMT on Monday, October 6, 1997, which is identical to the period used to produce the graph of the best market bid and ask quotes and transactions (Figure 4.9). The forecast probabilities of buyer-initiated trades are shown on the plot as the distance of lower solid line from the horizontal zero-probability line. The forecast probabilities of seller-initiated trades are shown as the distance of the upper solid line from the 100% probability horizontal line.

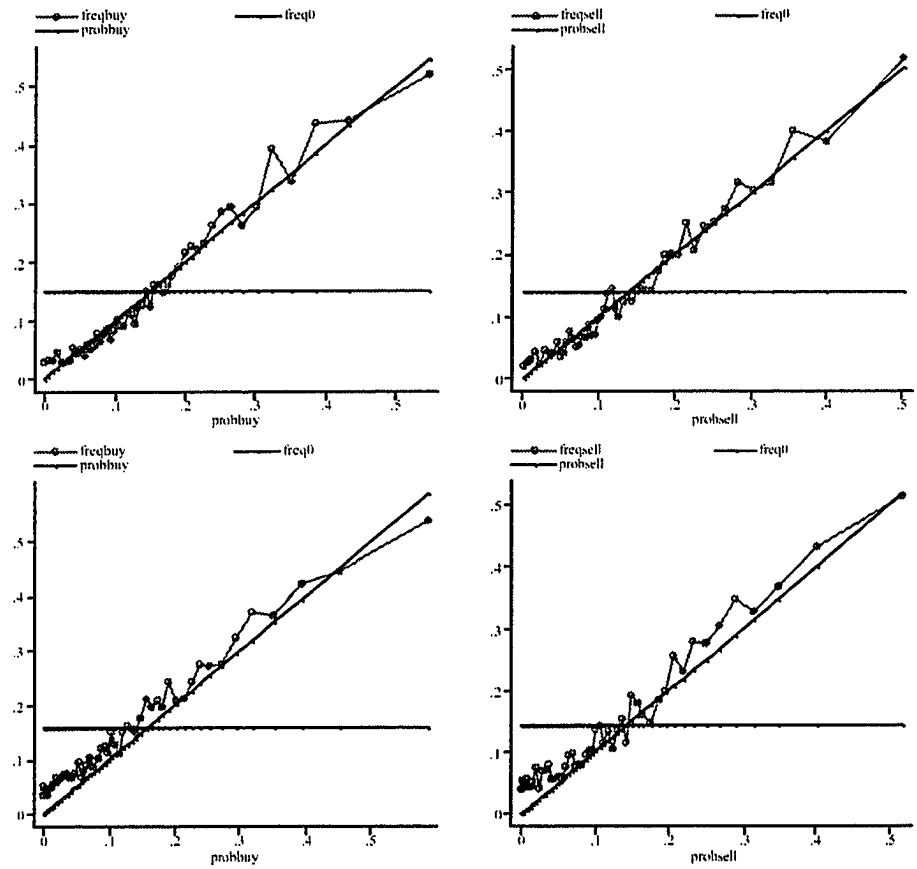


Figure 4.18: Reliability plots for one-step-ahead probability forecasts of buyer- and seller-initiated transactions

Figure 4.18 displays the reliability plots for one-step-ahead forecast probabilities of buyer- and seller-initiated transactions made after at least one second elapsed since the previous event. The forecasts are based on the version of PCA factor competing risks model with $S = 8$ types of observable risks and the covariate structure comprised by $Q = 5$ PCA factors as described in subsection 4.6.4. Two diagrams on the top are based on the liquid trading hours of the first three trading days (the model estimation period). The left and right diagrams on the top plot, respectively, the fractions of epochs terminated with buy and sell transactions against the forecast probabilities of such events. Two diagrams on the bottom are based on the liquid trading hours of the last two days (the out-of-sample period). The left and right diagrams on the bottom plot, respectively, the fractions of epochs terminated with buy and sell transactions against the forecast probabilities of such events.

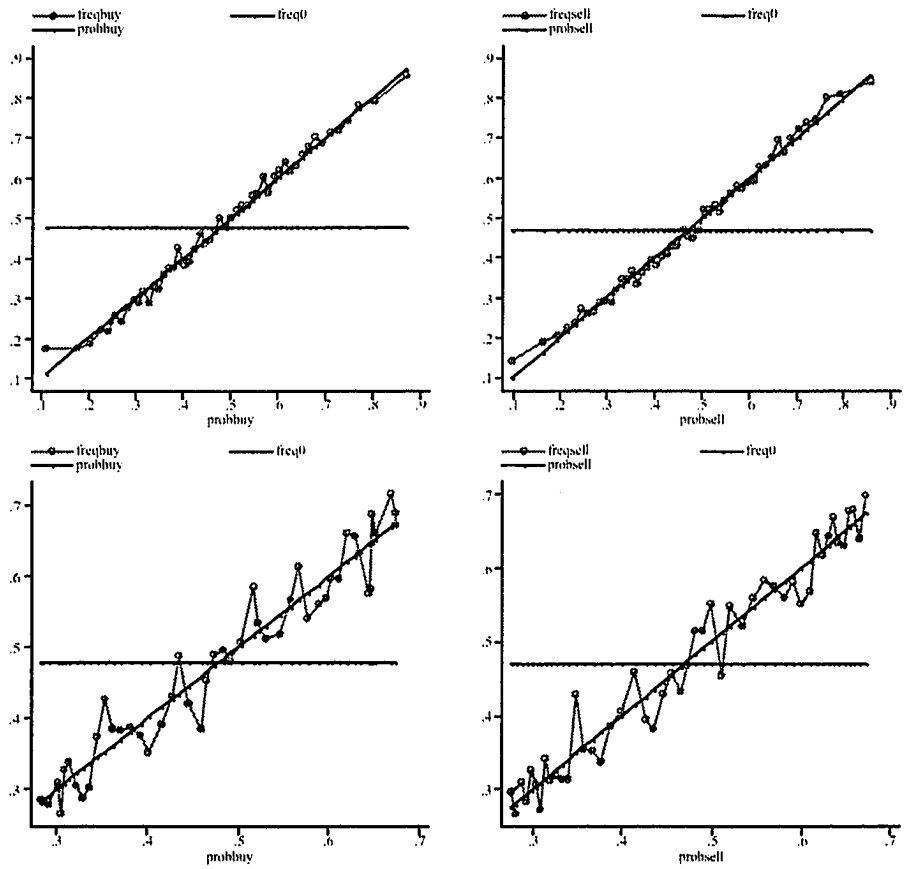


Figure 4.19: In-sample reliability plots for probability forecasts of the direction of next trade in 30 sec.

Two diagrams on the left-hand side of Figure 4.19 show reliability plots for probability forecasts of the event that the next transaction in the limit order book will be initiated by a buyer and will occur in 30 seconds since the time of forecast. Two diagrams on the right-hand side of Figure 4.19 show reliability plots for probability forecasts of the event that the next transaction in the limit order book will be initiated by a seller and will occur in 30 seconds since the time of forecast. The forecasts evaluated on the top section of the graph are produced by the trinomial logit regression with the covariates given by the five competing risks PCA factors. The forecasts evaluated on the lower section of the graph are produced by the benchmark forecasting model based on the trinomial logit regression with the covariates given by the signs of last 10 transactions (± 1 if the transaction was initiated by seller/buyer). All diagrams are based on the data from the estimation period 6 a.m. to 5 p.m. GMT on October 6–8. 1997.

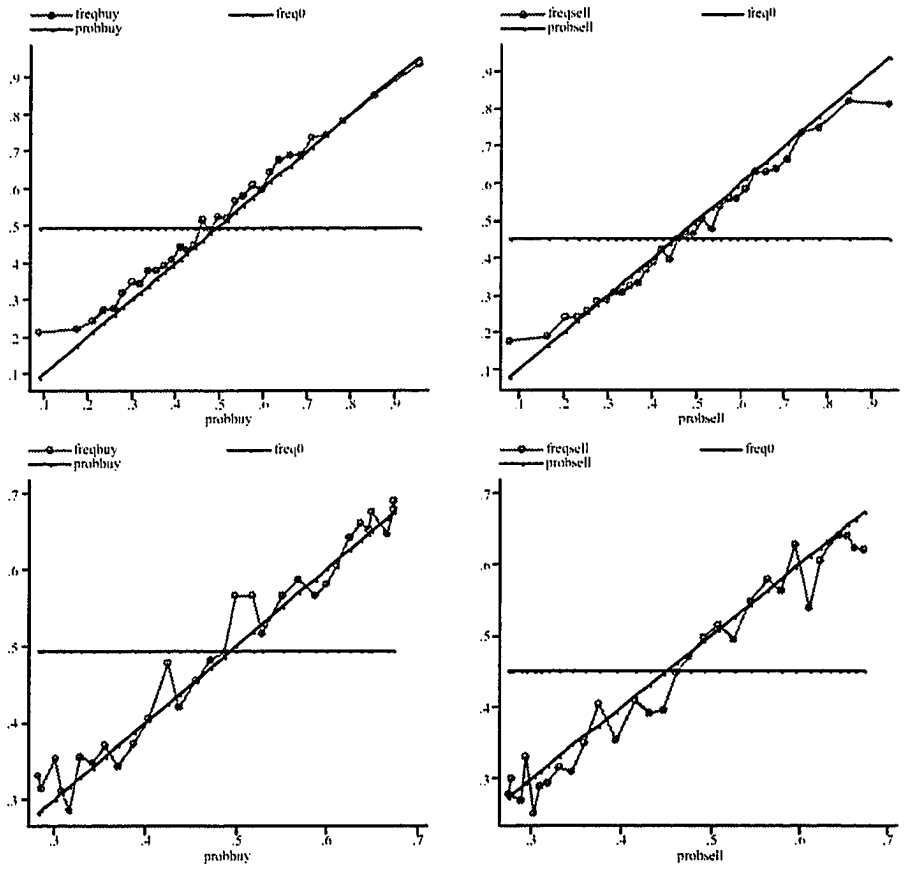


Figure 4.20: Out-of-sample reliability plots for probability forecasts of the direction of next trade in 30 sec.

Two diagrams on the left-hand side of Figure 4.20 show reliability plots for probability forecasts of the event that the next transaction in the limit order book will be initiated by a buyer and will occur in 30 seconds since the time of forecast. Two diagrams on the right-hand side of Figure 4.20 show reliability plots for probability forecasts of the event that the next transaction in the limit order book will be initiated by a seller and will occur in 30 seconds since the time of forecast. The forecasts evaluated on the top section of the graph are produced by the trinomial logit regression with the covariates given by the five competing risks PCA factors. The forecasts evaluated in the lower section of the graph are produced by the benchmark forecasting model based on the trinomial logit regression with the covariates given by the signs of last 10 transactions (± 1 if the transaction was initiated by seller/buyer). All diagrams are based on the data from the out-of-sample period covering the trading hours 6 a.m. to 5 p.m. GMT on October 9 and October 10, 1997.

Table 4.1: Classification of arrival and cancellation events on the sell side of limit order book

Risk (r)	Limit order price P^*	Price change	Quantity change
A1**	$P^* < P_{\text{bid}}$	$P_{\text{bid}} \searrow P_{\text{ask}} \searrow$	$Q_{\text{bid}} \downarrow\uparrow Q_{\text{ask}} \downarrow\uparrow$
A2**	Market sell order	$P_{\text{bid}} \searrow P_{\text{ask}}$ same	$Q_{\text{bid}} \downarrow$ if $\Delta P_{\text{bid}} = 0$
A3**	$P^* = P_{\text{bid}}$	$P_{\text{bid}} \searrow P_{\text{ask}} \searrow$	$Q_{\text{bid}} \downarrow\uparrow Q_{\text{ask}} \downarrow\uparrow$
A4*	$P^* - P_{\text{ask}} < -1$	$P_{\text{ask}} \downarrow$	$Q_{\text{ask}} \rightarrow Q^*$(lim.order size)
A5*	$P^* - P_{\text{ask}} = 1$	$P_{\text{ask}} \downarrow$	$Q_{\text{ask}} \rightarrow Q^*$(lim.order size)
A6*	$P^* - P_{\text{ask}} = 0$	No price effect	$Q_{\text{ask}} \uparrow$
A7	$P^* - P_{\text{ask}} = 1$	No price effect	No quantity effect
A8, A9, A10, and A11 defined similarly			
A12	$5 < P^* - P_{\text{ask}} \leq 10$	No price effect	No quantity effect
A13	$10 < P^* - P_{\text{ask}} \leq 20$	No price effect	No quantity effect
A14	$P^* - P_{\text{ask}} > 20$	No price effect	No quantity effect
AC6*	$P^* - P_{\text{ask}} = 0$	$P_{\text{ask}} \uparrow$	$Q_{\text{ask}} \downarrow$ if $\Delta P_{\text{ask}} = 0$
AC7	$P^* - P_{\text{ask}} = 1$	No price effect	No quantity effect
AC8, AC9, AC10, and AC11 defined similarly			
AC12	$5 < P^* - P_{\text{ask}} \leq 10$	No price effect	No quantity effect
AC13	$10 < P^* - P_{\text{ask}} \leq 20$	No price effect	No quantity effect
AC14	$P^* - P_{\text{ask}} > 20$	No price effect	No quantity effect

Table 4.2: Covariates in the Cox regressions of competing risks

Covariate	Description
Slippage	Current midquote minus last transaction price
Lt.Return	Last minus 2nd-to-last transaction price
$\text{Spread}_{>0}$	Best ask minus best bid quote, or zero, whatever is larger
$\text{Spread}^2_{>0}$	(Best ask minus best bid quote)^2, or zero, whatever is larger
ΔP_{ask}	Change of ask quote between last and 2nd-to-last events
$\Delta P_{ask,-1}$	Change of ask quote between 2nd- and 3rd-to-last events
ΔP_{bid}	Change of bid quote between last and 2nd-to-last events
$\Delta P_{bid,-1}$	Change of bid quote between 2nd- and 3rd-to-last events
$\log(Q_{ask})$	Log depth at best ask quote, or log \$10M, whatever is less
Q^+_{ask}	1 if log depth at ask equals \$10M, zero otherwise
$\Delta\log(Q_{ask})$	Change of $\log(Q_{ask})$ if $\Delta P_{ask}=0$, zero otherwise
$\Delta\log(Q_{ask,-1})$	Change of $\log(Q_{ask,-1})$ if $\Delta P_{ask}=0 \& \Delta P_{ask,-1}=0$, zero otherwise
$\log(Q_{bid})$	Log depth at best bid quote, or log \$10M, whatever is less
Q^+_{bid}	1 if log depth at bid equals \$10M, zero otherwise
$\Delta\log(Q_{bid})$	Change of $\log(Q_{bid})$ if $\Delta P_{bid}=0$, zero otherwise
$\Delta\log(Q_{bid,-1})$	Change of $\log(Q_{bid,-1})$ if $\Delta P_{bid}=0 \& \Delta P_{bid,-1}=0$, zero otherwise
Side	1 if last trade seller-initiated, -1 if buyer-initiated
Side_{-1}	1 if 2nd-to-last trade seller-initiated, -1 if buyer-initiated
$F_{0\text{-}5"}$	Signed number of trades 0 to 5 sec. prior to last event
$F_{5\text{-}10"}$, $F_{10\text{-}15"}$, $F_{15\text{-}30"}$, $F_{30\text{-}60"}$, $F_{1\text{-}2'}$, $F_{2\text{-}5'}$, $F_{5\text{-}15'}$ defined similarly	
$T_{0\text{-}5"}$	Number of trades 0 to 5 sec. prior to last event
$T_{5\text{-}10"}$, $T_{10\text{-}15"}$, $T_{15\text{-}30"}$, $T_{30\text{-}60"}$, $T_{1\text{-}2'}$, $T_{2\text{-}5'}$, $T_{5\text{-}15'}$ defined similarly	

Table 4.3: Estimated price and quantity coefficients for competing risks of seller-initiated events

Risk type (r) (1)	Slippage (2)	Spread (3)	ΔP_{ask} (4)	ΔP_{bid} (5)	Q_{ask} (6)	ΔQ_{ask} (7)	Q_{bid} (8)	ΔQ_{bid} (9)
Sell $< P_{\text{bid}}$	0.103	−0.754	0.437	0.172	−0.006	−0.479	−0.436	−0.381
Market sell	0.067	−0.525	0.135	−0.069	0.174	0.361	0.068	−1.161
Sell $= P_{\text{bid}}$	0.081	−0.670	0.222	0.091	0.140	0.041	0.005	−1.146
Sell $< P_{\text{ask}}-1$	0.062	0.451	0.089	−0.059	0.120	−0.013	0.053	−0.255
Sell at $P_{\text{ask}}-1$	0.017	0.029	0.048	−0.101	0.302	−0.235	0.005	−0.253
Sell at P_{ask}	−0.105	0.017	−0.051	−0.124	0.122	0.798	0.017	−0.100
Sell at $P_{\text{ask}}+1$	−0.176	−0.187	−0.130	−0.159	−0.101	0.060	0.058	0.025
Sell at $P_{\text{ask}}+2$	−0.126	−0.162	−0.142	−0.136	−0.080	0.081	0.125	0.003
Sell at $P_{\text{ask}}+3$	−0.138	−0.037	−0.120	−0.113	−0.026	0.037	0.080	−0.005
Sell at $P_{\text{ask}}+4$	−0.139	−0.153	−0.115	−0.012	−0.043	0.067	0.042	−0.009
Sell at $P_{\text{ask}}+5$	−0.208	−0.091	0.007	0.042	0.103	0.004	−0.045	0.051
Sell $\leq P_{\text{ask}}+10$	−0.135	−0.101	−0.080	−0.048	−0.081	0.013	0.025	−0.000
Sell $\leq P_{\text{ask}}+20$	−0.027	0.011	−0.098	−0.043	0.042	−0.028	0.009	−0.003
Cancel at P_{ask}	−0.149	−0.013	−0.014	−0.052	0.600	0.205	0.017	−0.228
Canc.at $P_{\text{ask}}+1$	−0.257	−0.128	0.039	−0.096	−0.046	−0.040	0.014	0.101
Canc.at $P_{\text{ask}}+2$	−0.299	−0.112	−0.102	0.050	0.029	−0.072	0.020	0.092
Canc.at $P_{\text{ask}}+3$	−0.158	−0.116	−0.078	−0.152	0.007	−0.058	−0.030	0.152
Canc.at $P_{\text{ask}}+4$	−0.212	−0.152	−0.153	0.021	−0.067	−0.015	0.015	0.127
Canc.at $P_{\text{ask}}+5$	−0.167	−0.159	−0.037	0.013	0.191	−0.170	0.029	0.145
Canc.$\leq P_{\text{ask}}+10$	−0.153	−0.016	−0.009	0.005	0.009	−0.047	0.048	0.091
Canc.$\leq P_{\text{ask}}+20$	−0.036	−0.088	−0.037	−0.070	−0.050	−0.007	−0.028	0.121

Table 4.4: Estimated lagged signed order flow coefficients for competing risks of seller-initiated events

Risk type (r) (1)	Side (2)	Side$_{-1}$ (3)	$F_{0\text{-}5''}$ (4)	$F_{5\text{-}10''}$ (5)	$F_{10\text{-}15''}$ (6)	$F_{15\text{-}30''}$ (7)	$F_{30\text{-}60''}$ (8)	$F_{1\text{-}2'}$ (9)
Sell $< P_{\text{bid}}$	0.311	0.288	−0.017	0.018	−0.028	−0.003	0.008	0.002
Market sell	0.319	0.170	0.044	0.013	−0.003	−0.004	0.002	−0.001
Sell $= P_{\text{bid}}$	0.238	0.104	0.029	−0.001	−0.024	−0.006	0.003	0.000
Sell $< P_{\text{ask}}-1$	0.216	0.075	0.018	−0.008	0.004	0.005	0.004	−0.000
Sell at $P_{\text{ask}}-1$	0.105	0.047	0.028	−0.019	−0.014	−0.009	0.000	−0.001
Sell at P_{ask}	0.022	−0.017	0.034	−0.013	−0.017	−0.008	−0.005	−0.004
Sell at $P_{\text{ask}}+1$	−0.021	−0.018	0.021	−0.022	−0.014	−0.010	−0.005	−0.006
Sell at $P_{\text{ask}}+2$	0.024	−0.017	−0.001	−0.022	−0.023	−0.008	−0.008	−0.002
Sell at $P_{\text{ask}}+3$	−0.057	−0.027	0.011	−0.034	−0.024	−0.009	−0.007	−0.005
Sell at $P_{\text{ask}}+4$	−0.085	0.035	0.001	−0.028	−0.034	−0.017	−0.008	−0.005
Sell at $P_{\text{ask}}+5$	0.011	−0.008	−0.001	−0.007	−0.030	−0.016	−0.006	−0.008
Sell $\leq P_{\text{ask}}+10$	−0.006	0.010	0.007	−0.025	−0.026	−0.013	−0.008	−0.005
Sell $\leq P_{\text{ask}}+20$	−0.066	−0.049	0.001	−0.033	−0.025	−0.021	−0.009	−0.010
Cancel at P_{ask}	0.042	0.085	0.014	0.022	0.007	−0.007	−0.002	−0.003
Canc.at $P_{\text{ask}}+1$	0.112	0.083	0.030	0.032	−0.009	−0.012	−0.003	−0.003
Canc.at $P_{\text{ask}}+2$	0.095	0.167	0.025	0.040	−0.005	−0.007	−0.004	−0.004
Canc.at $P_{\text{ask}}+3$	0.054	0.175	0.025	0.043	0.015	−0.002	−0.002	−0.007
Canc.at $P_{\text{ask}}+4$	0.138	0.097	0.039	0.042	0.015	−0.003	−0.005	−0.008
Canc.at $P_{\text{ask}}+5$	0.138	0.144	0.026	0.018	0.028	−0.008	0.001	−0.003
Canc.$\leq P_{\text{ask}}+10$	0.150	0.159	0.023	0.036	0.020	0.010	0.005	−0.001
Canc.$\leq P_{\text{ask}}+20$	0.076	0.041	0.020	0.005	0.028	0.013	0.013	0.007

Table 4.5: Estimated lagged trading activity coefficients for competing risks of seller-initiated events

Risk type (r) (1)	$T_{0\text{-}5''}$ (2)	$T_{5\text{-}10''}$ (3)	$T_{10\text{-}15''}$ (4)	$T_{15\text{-}30''}$ (5)	$T_{30\text{-}60''}$ (6)	$T_{1\text{-}2'}$ (7)	$T_{2\text{-}5'}$ (8)	$T_{5\text{-}15'}$ (9)
Sell $< P_{\text{bid}}$	0.069	0.009	0.037	0.004	0.004	0.006	−0.001	0.001
Market sell	0.019	0.008	0.018	0.006	0.004	0.003	0.001	0.001
Sell $= P_{\text{bid}}$	0.043	0.025	0.014	0.007	0.005	0.002	0.001	0.001
Sell $< P_{\text{ask}}-1$	0.037	0.011	0.008	−0.005	0.002	0.003	0.002	0.001
Sell at $P_{\text{ask}}-1$	0.014	0.022	0.008	0.001	0.007	0.003	0.002	0.001
Sell at P_{ask}	0.019	0.027	0.022	0.009	0.008	0.003	0.002	0.001
Sell at $P_{\text{ask}}+1$	0.020	0.034	0.016	0.011	0.007	0.003	0.002	0.001
Sell at $P_{\text{ask}}+2$	0.041	0.028	0.029	0.017	0.009	0.000	0.002	0.001
Sell at $P_{\text{ask}}+3$	0.024	0.034	0.024	0.022	0.012	0.005	0.001	0.001
Sell at $P_{\text{ask}}+4$	0.021	0.041	0.011	0.014	0.010	0.004	0.002	0.002
Sell at $P_{\text{ask}}+5$	0.035	0.033	0.033	0.033	0.006	0.005	0.001	0.001
Sell $\leq P_{\text{ask}}+10$	−0.002	0.039	0.032	0.026	0.013	0.007	0.001	0.001
Sell $\leq P_{\text{ask}}+20$	0.004	0.017	0.028	0.015	0.013	0.005	0.004	0.000
Cancel at P_{ask}	0.044	0.008	0.002	0.001	0.002	−0.000	0.000	0.001
Canc.at $P_{\text{ask}}+1$	0.027	0.003	0.005	0.004	0.005	0.005	0.002	0.001
Canc.at $P_{\text{ask}}+2$	0.036	0.011	0.018	0.008	0.006	0.004	0.002	0.001
Canc.at $P_{\text{ask}}+3$	0.035	0.010	0.015	0.015	0.005	−0.001	0.003	0.001
Canc.at $P_{\text{ask}}+4$	0.019	0.014	0.009	0.016	0.002	0.002	0.003	0.001
Canc.at $P_{\text{ask}}+5$	0.041	0.042	0.020	0.012	0.008	−0.001	0.003	0.000
Canc.$\leq P_{\text{ask}}+10$	0.031	0.018	0.005	0.008	0.004	0.003	0.001	0.002
Canc.$\leq P_{\text{ask}}+20$	0.029	0.021	0.016	0.012	0.002	−0.004	0.002	0.002

Table 4.6: Estimated price and quantity coefficients for competing risks of buyer-initiated events

Risk type (r) (1)	Slippage (2)	Spread (3)	ΔP_{bid} (4)	ΔP_{ask} (5)	Q_{bid} (6)	ΔQ_{bid} (7)	Q_{ask} (8)	ΔQ_{bid} (9)
Buy $> P_{\text{ask}}$	−0.269	−0.687	−0.224	−0.212	0.107	−0.518	−0.543	−0.503
Market buy	−0.132	−0.568	−0.133	0.087	0.194	0.552	0.053	−1.102
Buy $= P_{\text{ask}}$	−0.221	−0.760	−0.168	−0.050	0.199	−0.132	−0.032	−1.136
Buy $> P_{\text{bid}}+1$	−0.110	0.468	−0.102	0.037	0.175	−0.359	−0.002	−0.328
Buy at $P_{\text{bid}}+1$	−0.023	0.071	−0.097	0.110	0.234	−0.265	−0.003	−0.090
Buy at P_{bid}	0.047	0.061	0.066	0.080	0.113	0.709	0.030	−0.179
Buy at $P_{\text{bid}}-1$	0.066	−0.177	0.178	0.168	−0.092	0.062	0.069	0.017
Buy at $P_{\text{bid}}-2$	0.053	−0.161	0.211	0.135	−0.040	0.052	0.005	0.045
Buy at $P_{\text{bid}}-3$	0.053	−0.126	0.179	0.035	−0.054	0.071	−0.012	0.045
Buy at $P_{\text{bid}}-4$	0.058	−0.119	0.119	0.116	−0.094	0.068	0.085	0.013
Buy at $P_{\text{bid}}-5$	0.020	0.079	0.158	0.002	0.096	−0.043	0.002	0.028
Buy $\geq P_{\text{bid}}-10$	0.055	0.002	0.079	0.144	0.042	−0.009	−0.036	0.038
Buy $\geq P_{\text{bid}}-20$	0.043	−0.002	0.018	0.055	0.014	−0.036	0.021	0.008
Cancel at P_{bid}	0.077	−0.019	−0.023	0.035	0.600	0.111	0.003	−0.199
Canc.at $P_{\text{bid}}-1$	0.179	−0.106	−0.018	0.059	−0.043	−0.026	0.009	0.092
Canc.at $P_{\text{bid}}-2$	0.286	−0.134	0.066	0.029	−0.031	−0.072	−0.011	0.128
Canc.at $P_{\text{bid}}-3$	0.156	−0.088	0.118	0.047	−0.054	−0.085	−0.002	0.143
Canc.at $P_{\text{bid}}-4$	0.175	−0.057	0.140	0.014	0.042	−0.098	−0.131	0.182
Canc.at $P_{\text{bid}}-5$	0.110	0.219	0.139	−0.039	0.097	−0.132	−0.100	0.182
Canc.$\geq P_{\text{bid}}-10$	0.045	−0.033	0.072	0.036	0.048	−0.108	0.013	0.143
Canc.$\geq P_{\text{bid}}-20$	0.022	0.021	0.017	0.079	−0.060	−0.023	0.030	0.067

Table 4.7: Estimated lagged signed order flow coefficients for competing risks of buyer-initiated events

Risk type (r) (1)	Side (2)	Side$_{-1}$ (3)	$F_{0\text{-}5''}$ (4)	$F_{5\text{-}10''}$ (5)	$F_{10\text{-}15''}$ (6)	$F_{15\text{-}30''}$ (7)	$F_{30\text{-}60''}$ (8)	$F_{1\text{-}2'}$ (9)
Buy $> P_{\text{ask}}$	−0.042	−0.167	−0.041	−0.016	0.000	−0.010	−0.005	−0.003
Market buy	−0.246	−0.125	−0.053	−0.020	0.001	0.001	−0.003	−0.001
Buy $= P_{\text{ask}}$	−0.158	−0.103	−0.031	−0.002	0.009	0.004	−0.004	−0.003
Buy $> P_{\text{bid}}+1$	−0.157	0.040	−0.018	−0.001	0.011	−0.000	−0.007	−0.007
Buy at $P_{\text{bid}}+1$	−0.138	0.027	−0.028	0.010	0.021	0.007	−0.002	−0.001
Buy at P_{bid}	0.013	0.012	−0.025	0.012	0.021	0.009	−0.000	0.001
Buy at $P_{\text{bid}}-1$	0.056	0.030	−0.008	0.012	0.028	0.012	0.002	−0.001
Buy at $P_{\text{bid}}-2$	0.075	0.054	−0.004	0.018	0.020	0.003	0.003	0.000
Buy at $P_{\text{bid}}-3$	0.070	0.044	−0.007	0.017	0.024	0.003	0.002	0.001
Buy at $P_{\text{bid}}-4$	0.090	0.034	−0.006	0.015	0.019	0.007	0.004	−0.001
Buy at $P_{\text{bid}}-5$	0.041	0.096	0.002	0.026	0.009	0.001	0.006	−0.000
Buy $\geq P_{\text{bid}}-10$	0.086	0.060	0.001	0.015	0.020	0.005	0.009	0.004
Buy $\geq P_{\text{bid}}-20$	0.054	0.039	−0.020	0.011	0.017	0.013	0.003	0.006
Cancel at P_{bid}	−0.070	−0.069	−0.015	−0.025	−0.002	0.004	−0.002	0.002
Canc.at $P_{\text{bid}}-1$	−0.119	−0.023	−0.031	−0.029	−0.005	0.017	0.002	−0.001
Canc.at $P_{\text{bid}}-2$	−0.045	−0.132	−0.028	−0.024	0.005	0.008	0.000	−0.001
Canc.at $P_{\text{bid}}-3$	−0.102	−0.090	−0.049	−0.016	−0.017	0.007	0.005	0.001
Canc.at $P_{\text{bid}}-4$	−0.164	−0.120	−0.013	−0.026	−0.014	−0.000	0.000	0.002
Canc.at $P_{\text{bid}}-5$	−0.027	−0.144	−0.046	−0.051	−0.003	−0.001	0.001	−0.001
Canc.$\geq P_{\text{bid}}-10$	−0.098	−0.080	−0.046	−0.037	−0.032	−0.009	0.001	0.002
Canc.$\geq P_{\text{bid}}-20$	−0.073	−0.105	−0.011	−0.030	−0.035	−0.012	−0.012	−0.004

Table 4.8: Estimated lagged trading activity coefficients for competing risks of buyer-initiated events

Risk type (r) (1)	$T_{0\text{-}5"}$ (2)	$T_{5\text{-}10"}$ (3)	$T_{10\text{-}15"}$ (4)	$T_{15\text{-}30"}$ (5)	$T_{30\text{-}60"}$ (6)	$T_{1\text{-}2'}$ (7)	$T_{2\text{-}5'}$ (8)	$T_{5\text{-}15'}$ (9)
Buy $> P_{\text{ask}}$	0.075	0.022	0.036	0.015	0.006	0.001	0.003	0.000
Market buy	0.030	0.008	0.018	0.012	0.007	0.002	0.001	0.001
Buy $= P_{\text{ask}}$	0.050	0.017	0.017	0.009	0.010	0.001	0.001	0.000
Buy $> P_{\text{bid}}+1$	0.051	0.017	0.005	0.009	0.003	0.004	0.000	0.001
Buy at $P_{\text{bid}}+1$	0.017	0.014	−0.002	0.002	0.006	0.004	0.001	0.001
Buy at P_{bid}	0.025	0.017	0.017	0.007	0.006	0.002	0.001	0.001
Buy at $P_{\text{bid}}-1$	0.012	0.008	0.014	0.009	0.006	0.004	0.001	0.001
Buy at $P_{\text{bid}}-2$	0.026	0.023	0.030	0.015	0.006	−0.000	0.001	0.001
Buy at $P_{\text{bid}}-3$	0.031	0.014	0.012	0.019	0.008	−0.001	0.001	0.001
Buy at $P_{\text{bid}}-4$	0.012	0.008	0.020	0.018	0.009	0.004	0.002	0.001
Buy at $P_{\text{bid}}-5$	0.028	0.013	0.037	0.024	0.011	−0.004	0.002	0.001
Buy $\geq P_{\text{bid}}-10$	0.006	0.018	0.032	0.018	0.007	0.002	0.001	0.001
Buy $\geq P_{\text{bid}}-20$	0.014	−0.002	0.018	0.010	0.011	0.004	0.001	0.001
Cancel at P_{bid}	0.049	0.003	0.008	0.001	0.004	−0.001	0.000	0.000
Canc.at $P_{\text{bid}}-1$	0.031	−0.006	0.007	0.002	0.004	0.004	0.001	0.001
Canc.at $P_{\text{bid}}-2$	0.051	0.011	0.003	0.007	0.002	0.005	0.000	0.001
Canc.at $P_{\text{bid}}-3$	0.037	0.011	0.014	0.007	0.005	0.002	0.001	0.001
Canc.at $P_{\text{bid}}-4$	0.038	0.027	−0.006	0.010	0.010	0.002	0.002	0.001
Canc.at $P_{\text{bid}}-5$	0.056	0.030	0.014	0.011	0.013	0.008	−0.001	0.001
Canc.$\geq P_{\text{bid}}-10$	0.031	0.021	0.025	0.006	0.011	0.005	−0.000	0.001
Canc.$\geq P_{\text{bid}}-20$	0.027	0.029	0.017	0.016	0.007	0.008	0.001	0.001

Table 4.9: Representation of PCA factor indices in terms of observable characteristics of the limit order book and recent trading history

Covariate	Fact.1	Fact.2	Fact.3	Fact.4	Fact.5	Fact.6	Fact.7
Constant	−1.186	−0.287	−0.363	−2.496	0.227	−0.639	0.477
Slippage	−0.068	0.330	−0.230	−0.014	−0.280	0.042	−0.166
$\text{Spread}_{>0}$	−0.185	0.095	0.080	0.779	−0.058	−0.169	−0.076
ΔP_{ask}	0.028	0.129	−0.229	0.036	−0.103	0.150	−0.130
ΔP_{bid}	0.016	0.144	−0.221	−0.055	−0.152	−0.367	−0.101
$\log(Q_{\text{ask}})$	0.033	−0.131	−0.103	0.165	−0.367	0.712	0.124
$\Delta \log(Q_{\text{ask}})$	−0.004	−0.026	−0.051	0.193	−0.765	−0.380	0.459
$\log(Q_{\text{bid}})$	0.078	0.089	0.160	0.191	0.386	0.637	−0.318
$\Delta \log(Q_{\text{bid}})$	−0.058	0.042	0.012	−0.020	0.684	−0.753	−0.725
Side	0.034	−0.241	−0.223	0.025	−0.184	−0.086	−0.091
Side_{-1}	0.037	−0.211	−0.140	0.072	0.013	−0.015	−0.118
$F_{0-5''}$	−0.002	−0.077	−0.001	0.011	−0.038	−0.002	0.000
$F_{5-10''}$	−0.003	−0.048	−0.059	0.009	0.061	−0.013	0.002
$F_{10-15''}$	−0.001	−0.007	−0.059	0.019	0.066	0.000	−0.041
$F_{15-30''}$	−0.004	0.012	−0.031	0.003	0.027	0.003	−0.010
$F_{30-60''}$	−0.002	0.002	−0.022	0.003	0.011	−0.001	0.016
$F_{1-2'}$	−0.003	0.004	−0.013	−0.000	0.010	0.001	0.013
$T_{0-5''}$	0.059	0.016	−0.003	0.024	0.020	0.073	0.017
$T_{5-10''}$	0.038	−0.002	0.027	0.012	−0.022	−0.018	−0.011
$T_{10-15''}$	0.034	0.002	−0.002	0.006	0.002	−0.028	0.034
$T_{15-30''}$	0.027	0.002	0.004	0.010	0.004	−0.039	0.021
$T_{30-60''}$	0.016	0.004	0.004	0.008	−0.004	−0.011	0.002
$T_{1-2'}$	0.005	0.003	0.004	0.002	−0.007	0.002	0.000
$T_{2-5'}$	0.003	−0.001	0.001	0.002	−0.001	−0.003	−0.006
$T_{5-15'}$	0.002	0.000	−0.001	0.001	0.000	−0.000	−0.000

Table 4.10: Cox five-factor regressions for competing risks of sell order arrivals

Risk type (r)	Fact.1 (activity)	Fact.2 (imbalance)	Fact.3 (b.pressure)	Fact.4 (adv.selection)	Fact.5 (b.momentum)
A1: Sell below P_{bid}	0.988* (23.21)	−0.115* (−3.47)	−0.589* (−16.37)	−0.861* (−11.74)	−0.279* (−8.53)
A2: Market Sell	0.565* (40.04)	−0.463* (−32.42)	−0.423* (−33.19)	−0.275* (−16.55)	−0.325* (−22.63)
A3: Sell at P_{bid}	0.717* (43.34)	−0.245* (−14.23)	−0.391* (−24.78)	−0.495* (−21.15)	−0.329* (−18.62)
A4: Sell at $\leq P_{\text{ask}} - 1$	0.053* (2.45)	−0.078* (−3.00)	−0.114* (−4.76)	0.652* (26.61)	−0.140* (−5.25)
A5: Sell at $P_{\text{ask}} - 1$	0.340* (22.45)	−0.169* (−10.57)	−0.016 (−1.07)	0.168* (9.61)	−0.243* (−14.89)
A6: Sell at P_{ask}	0.412* (37.44)	−0.199* (−16.66)	−0.181* (17.22)	0.110* (9.38)	−0.132* (−10.30)
A7: Sell at $P_{\text{ask}} + 1$	0.537* (36.32)	−0.199* (−12.04)	0.329* (23.15)	−0.154* (−9.12)	0.091* (5.38)
A8: Sell at $P_{\text{ask}} + 2$	0.587* (29.20)	−0.117* (−4.89)	0.313* (15.66)	−0.108* (−4.56)	0.084* (3.71)
A9: Sell at $P_{\text{ask}} + 3$	0.621* (24.09)	−0.066 (−2.40)	0.362* (14.82)	−0.010 (−0.35)	−0.000 (−0.02)
A10: Sell at $P_{\text{ask}} + 4$	0.609* (16.56)	−0.063 (−1.66)	0.385* (11.24)	−0.092 (−2.24)	0.003 (0.07)
A11: Sell at $P_{\text{ask}} + 5$	0.624* (16.69)	−0.082 (−1.64)	0.334* (8.79)	0.015 (0.31)	−0.073 (−1.42)
A12: Sell at $\leq P_{\text{ask}} + 10$	0.599* (21.36)	−0.061 (−2.20)	0.336* (12.17)	0.019 (0.62)	−0.012 (−0.39)
A13: Sell at $\leq P_{\text{ask}} + 20$	0.431* (10.53)	−0.004 (−0.09)	0.394* (9.43)	0.096 (2.23)	−0.214* (−4.27)
A14: Sell above $P_{\text{ask}} + 20$	0.355* (8.17)	−0.038 (−0.89)	0.271* (6.82)	0.060 (1.30)	−0.160* (−3.69)

Table 4.11: Cox five-factor regressions for competing risks of buy order arrivals

Risk type (r)	Fact.1 (activity)	Fact.2 (imbalance)	Fact.3 (b.pressure)	Fact.4 (adv.selection)	Fact.5 (b.momentum)
B1: Buy above P_{ask}	1.013* (22.83)	−0.004 (−0.08)	0.495* (9.99)	−0.797* (−9.55)	0.585* (10.87)
B2: Market Buy	0.554* (39.36)	0.366* (29.45)	0.300* (21.42)	−0.486* (−30.06)	0.387* (26.26)
B3: Buy at P_{ask}	0.767* (46.34)	0.096* (5.73)	0.249* (14.64)	−0.650* (−27.28)	0.463* (24.83)
B4: Buy at $\geq P_{\text{bid}} + 1$	0.123* (5.59)	0.167* (6.44)	0.286* (12.52)	0.586* (22.19)	0.156* (5.16)
B5: Buy at $P_{\text{bid}} + 1$	0.297* (19.72)	0.176* (11.47)	0.053* (3.62)	0.159* (9.31)	0.184* (10.40)
B6: Buy at P_{bid}	0.364* (33.95)	0.177* (17.61)	−0.149* (−14.29)	0.112* (10.05)	0.107* (9.65)
B7: Buy at $P_{\text{bid}} - 1$	0.484* (33.31)	0.094* (7.58)	−0.316* (−20.39)	−0.129* (−7.64)	−0.048* (−3.24)
B8: Buy at $P_{\text{bid}} - 2$	0.536* (27.67)	0.067* (3.90)	−0.331* (−17.11)	−0.089* (−3.77)	−0.039 (−2.19)
B9: Buy at $P_{\text{bid}} - 3$	0.493* (20.07)	0.057 (2.35)	−0.292* (−11.86)	−0.082* (−2.72)	−0.004 (−0.18)
B10: Buy at $P_{\text{bid}} - 4$	0.527* (16.84)	0.074 (2.33)	−0.279* (−9.15)	−0.035 (−0.88)	−0.077* (−2.64)
B11: Buy at $P_{\text{bid}} - 5$	0.454* (13.99)	0.043 (1.27)	−0.232* (−6.74)	0.200* (5.34)	0.070 (1.92)
B12: Buy at $\geq P_{\text{bid}} - 10$	0.471* (20.73)	0.080* (3.55)	−0.323* (−13.46)	0.120* (4.57)	0.023 (1.10)
B13: Buy at $\geq P_{\text{bid}} - 20$	0.425* (12.82)	0.143* (4.54)	−0.242* (−7.27)	0.169* (4.78)	0.085* (2.76)
B14: Buy below $P_{\text{bid}} - 20$	0.334* (9.45)	0.121* (3.49)	−0.209* (−6.16)	0.210* (6.05)	0.133* (3.75)

Table 4.12: Cox five-factor regressions for competing risks of limit order cancellations

Risk type (r)	Fact.1 (activity)	Fact.2 (imbalance)	Fact.3 (b.pressure)	Fact.4 (adv.selection)	Fact.5 (b.momentum)
AC6: Canc.at P_{ask}	0.292* (25.56)	−0.351* (−24.44)	−0.057* (−4.91)	0.151* (12.55)	−0.171* (−11.40)
AC7: Canc.at $P_{\text{ask}} + 1$	0.446* (22.85)	−0.427* (−20.52)	0.096* (5.30)	−0.016 (−0.73)	0.080* (4.01)
AC8: Canc.at $P_{\text{ask}} + 2$	0.513* (20.98)	−0.511* (−17.50)	0.069* (2.88)	−0.002 (−0.06)	0.095* (3.48)
AC9: Canc.at $P_{\text{ask}} + 3$	0.553* (18.30)	−0.518* (−13.57)	−0.008 (−0.24)	0.017 (0.45)	0.093 (2.32)
AC10: Canc.at $P_{\text{ask}} + 4$	0.589* (14.29)	−0.549* (−11.10)	−0.004 (−0.10)	−0.090 (−1.76)	0.108 (2.20)
AC11: Canc.at $P_{\text{ask}} + 5$	0.581* (12.42)	−0.484* (−8.77)	−0.103 (−2.27)	0.028 (0.50)	−0.008 (−0.16)
AC12: Canc.at $\leq P_{\text{ask}} + 10$	0.468* (17.53)	−0.387* (−12.35)	−0.162* (−5.81)	0.176* (6.22)	0.118* (4.21)
AC13: Canc.at $\leq P_{\text{ask}} + 20$	0.444* (12.68)	−0.223* (−5.36)	−0.229* (−5.97)	0.074 (1.72)	0.182* (4.70)
AC14: Canc.above $P_{\text{ask}} + 20$	0.397* (9.88)	0.039 (0.95)	−0.179* (−4.21)	0.249* (6.02)	0.182* (4.33)
BC6: Canc.at P_{bid}	0.214* (18.67)	0.324* (29.02)	0.105* (8.96)	0.079* (6.87)	0.188* (14.53)
BC7: Canc.at $P_{\text{bid}} - 1$	0.365* (18.63)	0.259* (16.80)	−0.062* (−3.20)	−0.045 (−2.08)	0.025 (1.40)
BC8: Canc.at $P_{\text{bid}} - 2$	0.418* (18.18)	0.496* (20.75)	−0.123* (−4.88)	−0.155* (−5.68)	−0.105* (−4.26)
BC9: Canc.at $P_{\text{bid}} - 3$	0.384* (12.21)	0.536* (16.92)	−0.055 (−1.57)	−0.103* (−2.83)	−0.086* (−2.61)
BC10: Canc.at $P_{\text{bid}} - 4$	0.463* (10.85)	0.317* (10.99)	0.047 (1.13)	−0.043 (−0.89)	0.036 (1.03)
BC11: Canc.at $P_{\text{bid}} - 5$	0.506* (12.43)	0.559* (13.65)	0.138* (3.56)	0.124* (2.84)	−0.000 (−0.01)
BC12: Canc.at $\geq P_{\text{bid}} - 10$	0.439* (16.60)	0.360* (18.27)	0.188* (7.24)	0.036 (1.27)	0.011 (0.52)
BC13: Canc.at $\geq P_{\text{bid}} - 20$	0.426* (11.50)	0.233* (9.09)	0.252* (8.03)	0.131* (3.69)	−0.158* (−5.37)
BC14: Canc.below $P_{\text{bid}} - 20$	0.372* (9.32)	0.096* (2.56)	0.291* (7.88)	0.172* (4.30)	−0.101 (−2.44)

Table 4.13: Eigenvalues and cumulative contribution of principal components to the competing risk indices

	1	2	3	4	5	6	7	8	9	10
Eigenvalue	20.16	7.26	5.70	4.12	2.15	1.01	0.88	0.66	0.61	0.56
Cumulative	0.438	0.596	0.720	0.810	0.857	0.879	0.898	0.912	0.925	0.937

Table 4.14: Modified classification of observable events in the forecasting model

Risk category (s)	Description of event	Risk types (r) included
AA	Seller-initiated transaction	Risks A1–A3 in Table 4.1
AP+	Ask price improvement	Risks A4–A5 in Table 4.1
AD+	Ask depth improvement	Risk A6 in Table 4.1
A−	Ask touch cancellation	Risk AC6 in Table 4.1
BB	Buyer-initiated transaction	Risks B1–B3 in Table 4.1
BP+	Bid price improvement	Risks B4–B5 in Table 4.1
BD+	Bid depth improvement	Risk B6 in Table 4.1
B−	Bid touch cancellation	Risks BC6 in Table 4.1

Chapter 5

Conclusion and Future Research

The analysis of the time between transactions, quotes, and other market events remains a popular topic of the empirical research of financial market microstructure. After the publication of papers by Engle and Russell ([41], [42]), a large number of extensions and modifications have appeared in the literature. One line research that may be entertained in the future takes on the possibility of using more flexible forms of the underlying distributions that might be able to capture unobserved heterogeneity. The unobserved heterogeneity can be introduced into the model via a random frailty parameter acting multiplicatively on the hazard rates of competing risks (Parner [113]). The frailty multiplier accommodating the clustering of the market activity may be interpreted as a latent factor which is unobserved by the econometrician. The frailty parameter then can be modeled as a gamma random variable and estimated jointly with the competing risks by the semiparametric maximum likelihood technique (Nielsen *et al.* [107]).

A closely related set of papers aim to extend the proposed framework to capture the serial dependence in the inter-event durations with a more flexible and parsimonious model. The serial dependence in activity levels that is not captured by recent transaction activity variables can be modeled by one or several dynamic latent factors perturbed multiplicatively by innovations with positive support (Bauwens and Veredas, [11]). Here the challenge is

in developing a computationally feasible estimation procedure that deals efficiently with censored observations.

Another line of research tackles the challenge of extending the existing univariate duration models to the multivariate case, which opens up the possibility of the full-scale empirical analysis of individual assets subject to multiple risks, as well as their portfolios. The competing risks approach, which is by design a multivariate methodology, represents one step towards this goal but still leaves open many theoretical and practical questions. In conclusion, we outline two practical applications of this methodology that can be made towards this goal.

One natural application of the competing risks technique involves the empirical analysis of thinly or irregularly traded financial instruments (Spierdijk *et al.* [123]) and emerging market securities. In combination with more traditional pricing models, the methodology developed in this dissertation may also be applied to the analysis of mortgage-backed securities and credit derivatives that might be simultaneously affected by complex combinations of qualitative risks (Duffie and Singleton [35]). For instance, credit rating agencies collecting information on credit histories of firms and individuals observe the times of their loan applications, debt refinancings, and bankruptcies, as well as the joint evolution of the market conditions and key characteristics of firms and individuals. Even though at any moment of time the economic entities (firms or individuals) can be at risk of experiencing any of these events, it is sufficient for only one of these events to occur in order to change the state of the world, leading to drastic realignments of the risk profiles after event. Until recently, almost all empirical work in the area of credit risk management was of exploratory nature, so the focus should be on the development of new econometric models and rigorous quantitative techniques for model evaluation and forecasting.

Bibliography

[1] Acar, E., "A Comparative Study of Foreign Exchange Quotes: Indicative versus Market Quotes." *Working Paper*, Citibank (1999).

[2] Addison, J.T. and P. Portugal, "Some Specification Issues in Unemployment Duration Analysis." *Labour Economics*, **5** (1998), 53–66.

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

[4] Admati, A. and P. Pfleiderer, "Divide and Conquer: A Theory of Intraday and Day-of-the-Week Mean Effects." *Review of Financial Studies*, **2** (1989), 189–224.

[5] Aït-Sahalia, Y. and P.A. Mykland, "The Effects of Random and Discrete Sampling When Estimating Continuous-Time Diffusions." *Econometrica*, **71** (2003), 483–549.

[6] Al-Suhaibani, M. and L. Kryzanowski, "The Information Content of Orders on the Saudi Stock Market." *Journal of Financial Research*, **23** (2000), 145–156.

[7] Andersen, P.K. and R.D. Gill, "Cox's Regression Model for Counting Processes: A Large Sample Study." *Annals of Statistics*, **10** (1982), 1100–20.

[8] Andersen, P.K., Ø. Borgan, R.D. Gill, and N. Kieding, *Statistical Models Based on Counting Processes.* Springer-Verlag, New York (1993).

[9] Bachelier, L., "Theory of Speculation." In Cootner, P. (Ed.), *The Random Character of Stock Market Prices.* MIT Press, Cambridge, MA (1964).

[10] Bauwens, L. and P. Giot, "The Logarithmic ACD Model: An Application to the Bid-Ask Quote Process of Three NYSE Stocks." *Annales d'Economie et de Statistique*, **60** (2000), 117–149.

[11] Bauwens, L. and D. Veredas, "The Stochastic Conditional Duration Model: A Latent Variable Model for the Analysis of Financial Durations." *Discussion Paper*, CORE, Université catholique de Louvain (1999).

[12] Biais, B., P. Hillion, and C. Spatt, "An Empirical Analysis of the Limit-Order Book and the Order Flow in the Paris Bourse." *Journal of Finance*, **50** (1995), 1655–89.

[13] Bisière, C. and T. Kamionka, "Timing of Orders, Order Aggressiveness and the Order Book in the Paris Bourse." *Working Paper*, University of Toulouse (1999).

[14] Bjønnes, G.H. and D. Rime, "FX Trading...Live! Dealer Behavior and Trading Systems in Foreign Exchange Markets." *Working Paper*, University of Oslo (2000).

[15] Breslow, N.E., "Covariance Analysis of Censored Survival Data." *Biometrics*, **30** (1974), 89–99.

[16] Brier, G.W., "Verification of Forecasts Expressed in Terms of Probability." *Monthly Weather Review*, **78** (1950), 1–3.

[17] Brown, P., N. Thomson, and D. Walsh, "Characteristics of the Order Flow Through an Electronic Open Limit Order Book." *Journal of International Financial Markets, Institutions and Money*, **9** (1999), 335–357.

[18] Burr, D., "A Comparison of Certain Bootstrap Confidence Intervals in the Cox Model." *Journal of the American Statistical Association*, **89** (1994), 1290–1302.

[19] Campbell, J.Y., A.W. Lo, and A.C. MacKinlay, *The Econometrics of Financial Markets*. Princeton University Press, Princeton, NJ (1997).

[20] Chakravarty, S. and C. Holden, "An Integrated Model of Market and Limit Orders." *Journal of Financial Intermediation*, **4** (1995), 213–241.

[21] Cheung, Y.-W. and M. Chinn, "Foreign Exchange Traders in Hong Kong, Tokyo, and Singapore: A Survey Study." *Working Paper*, University of California at Santa Cruz (1999).

[22] Cheung, Y.-W. and C.Y.-P. Wong, "A Survey of Market Practitioners: Views on Exchange Rate Dynamics." *Journal of International Economics*, **51** (2000), 401–419.

[23] Cleveland, W.S., "Robust Locally Weighted Regression and Smoothing Scatterplots." *Journal of the American Statistical Association*, **74** (1979), 829–836.

[24] Coppejans, M. and I. Domowitz, "Screen Information, Trader Activity, and Bid-Ask Spreads in a Limit Order Market." *Working Paper*, ITG Inc. (1999).

[25] Cowles, A., "Can Stock Market Forecasters Forecast?" *Econometrica*, **1** (1933), 309–324.

[26] Cox, D.R., "Regression Models and Life Tables." *Journal of the Royal Statistical Society, Series B*, **34** (1972), 187–220.

[27] Cox, D.R., "Partial Likelihood." *Biometrika*, **62** (1975), 269–276.

[28] Cox, D.R. and D. Oakes, *Analysis of Survival Data*. Chapman and Hall, London, UK (1984).

[29] Dacorogna, M.M., U.A. Müller, C. Jost, O.V. Pictet, R.B. Olsen, and J.R. Ward, "Heterogeneous Real-Time Trading Strategies in the Foreign Exchange Market," *European Journal of Finance*, **1** (1995), 243–253.

[30] Dacorogna, M.M., R. Gençay, U.A. Müller, R.B. Olsen, and O.V. Pictet, *An Introduction to High-Frequency Finance*. Academic Press, San Diego, CA (2001).

[31] Daníelsson, J. and R. Payne, "Measuring and Explaining Liquidity on an Electronic Limit Order Book: Evidence from Reuters D2000-2." *Working Paper*, Financial Markets Group, London School of Economics (2001).

[32] Daníelsson, J. and R. Payne, "Real Trading Patterns and Prices in Spot Foreign Exchange Markets." *Journal of International Money and Finance*, **21** (2002), 203–222.

[33] Davis, R.A., T.H. Rydberg, N. Shephard, and S.B. Streett, "The Cbin Model for Counts: Testing for Common Features in the Speed of Trading, Quote Changes, Limit and Market Order Arrivals." *Working Paper* (1999).

[34] Diamond, D.W. and R.E. Verrecchia, "Information Aggregation in a Noisy Rational Expectation Model." *Journal of Financial Economics*, **16** (1987), 277–311.

[35] Duffie, D. and K.J. Singleton, *Credit Risk: Pricing, Measurement, and Management.* Princeton University Press, Princeton, New Jersey (2003).

[36] Dufour, A. and R. Engle, "Time and the Price Impact of a Trade." *Journal of Finance*, **55** (2000), 2467–98.

[37] Easley, D. and M. O'Hara, "Time and the Process of Security Price Adjustment." *Journal of Finance*, **47** (1992), 577–606.

[38] Efron, B., "Censored Data and the Bootstrap." *Journal of the American Statistical Association*, **76** (1981), 312–319.

[39] Engle, R.F., "The Econometrics of Ultra-High-Frequency Data." *Econometrica*, **68** (2000), 1–22.

[40] Engle, R. and A. Lunde, "Trades and Quotes, a Bivariate Point Process." *Discussion Paper*, University of California at San Diego (1999).

[41] Engle, R.F. and J.R. Russell, "Forecasting the Frequency of Changes in Quoted Foreign Exchange Prices with the Autoregressive Conditional Duration Model." *Journal of Empirical Finance*, **4** (1997), 187–212.

[42] Engle, R.F. and J.R. Russell, "Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data." *Econometrica*, **66** (1998), 1127–1162.

[43] Evans, M.D.D. and R.K. Lyons, "Order Flow and Exchange Rate Dynamics." *NBER Working Paper* #7317, National Bureau of Economic Research, Cambridge, MA (1999).

[44] Fama, E., "Efficient Capital Markets: A Review of Theory and Empirical Work." *Journal of Finance*, **25** (1970), 383–417.

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

[46] Foucault, T., O. Kadan, and E. Kandel, "Limit Order Book as a Market for Liquidity." *Working Paper.*

[47] Frontier, S., "Étude de la decroissance des valeurs propres dans une analyze en composantes principales: comparison avec le modèle de baton brisé." *Journal of Experimental Marine Biology and Ecology*, **25** (1976), 67–75.

[48] Geman, H. and T. Ané, "Stochastic Subordination." *Risk*, **9** (1996), 145–149.

[49] Ghysels, E. and J. Jasiak, "GARCH for Irregularly Spaced Financial Data: The ACD-GARCH Model." *Studies in Nonlinear Dynamics and Econometrics*, **2**, 133–149.

[50] Glosten, L., "Is the Electronic Order Book Inevitable?" *Journal of Finance*, **49** (1994), 1127–61.

[51] González-Manteiga, W., R. Cao, and J.S. Marron, "Bootstrap Selection of the Smoothing Parameter in Nonparametric Hazard Rate Estimation." *Journal of the American Statistical Association*, **91** (1996), 1130–1140.

[52] Goodhart, C.A.E. and M. O'Hara, "High Frequency Data in Financial Markets: Issues and Applications." *Journal of Empirical Finance*, **4** (1997), 73–114.

[53] Goodhart, C.A.E. and R. Payne, "Microstructural Dynamics in a Foreign Exchange Electronic Broking System." *Journal of International Money and Finance*, **15** (1996), 829–852.

[54] Goodhart, C., T. Ito, and R. Payne, "One Day in June 1993: A Study of the Working of the Reuters 2000-2 Electronic Foreign Exchange Trading System." In Frankel, J., G. Galli, and A. Giovannini (Eds.), *The Microstructure of Foreign Exchange Markets.* The University of Chicago Press (1996), 107–179.

[55] Gouriéroux, C., J. Jasiak, and G. Le Fol, "Intra-Day Market Activity." *Journal of Financial Markets*, **2** (1999), 193–216.

[56] Grammig, J. and M. Wellner, "Modeling the Interdependence of Volatility and Inter-Transaction Duration Processes." *Journal of Econometrics*, **106** (2002), 369–400.

[57] Greene, W.H., *Econometric Analysis*, 5th edition. Prentice Hall, Upper Saddle River, NJ (2002).

[58] Griffiths, M.D., B.F. Smith, D.A.S. Turnbull, and R.W. White, "The Costs and Determinants of Order Aggressiveness." *Journal of Financial Economics*, **56** (2000), 65–88.

[59] Hamao, Y. and J. Hasbrouck, "Securities Trading in the Absence of Dealers: Trades and Quotes on the Tokyo Stock Exchange." *Review of Financial Studies*, **8** (1995), 849–878.

[60] Han, A. and J.A. Hausman, "Flexible Parametric Estimation of Duration and Competing Risk Models." *Journal of Applied Econometrics*, **5** (1990), 1–28.

[61] Handa, P. and R.A. Schwartz, "Limit Order Trading." *Journal of Finance*, **51** (1996), 1835–61.

[62] Hansen, L. and J. Scheinkman, "Back to the Future: Generating Moment Implications for Continuous Time Markov Processes." *Econometrica*, **63** (1995), 767–804.

[63] Härdle, W., *Applied Nonparametric Regression.* Econometric Society Monographs No.19, Cambridge University Press, UK (1990).

[64] Harris, L., *Trading and Exchanges: Market Microstructure for Practitioners.* Oxford University Press, UK (2003).

[65] Harris, L. and J. Hasbrouck, "Market vs. Limit Orders: The SuperDOT Evidence on Order Submission Strategy." *Journal of Financial and Quantitative Analysis*, **31** (1996), 213–231.

[66] Hasbrouck, J., "The Summary Informativeness of Stock Trades: An Econometric Analysis." *Review of Financial Studies*, **4** (1991), 571–595.

[67] Hasbrouck, J., "Stalking the 'Efficient Price' in Market Microstructure Specifications: An Overview." *Journal of Financial Markets*, **5** (2002), 329–339.

[68] Hasbrouck, J. and T.S.Y. Ho, "Order Arrival, Quote Behavior and the Return Generating Process." *Journal of Finance*, **42** (1987), 1035–48.

[69] Hasbrouck, J. and G. Saar, "Limit Orders and Volatility in a Hybrid Market: The Island ECN." *Working Paper*, New York University.

[70] Hausman, J.A., A.W. Lo, and A.C. MacKinlay, "An Ordered Probit Analysis of Transaction Stock Prices." *Journal of Financial Economics*, **31** (1992), 319–379.

[71] Hautsch, N., "Analyzing the Time between Trades with a Gamma Compounded Hazard Model: An Application to LIFFE Bund Future Transactions." *Working Paper*, University of Konstanz, Germany (1999).

[72] Hillman, R. and M. Salmon, "Intrinsic Stationarity: Investigating Predictability in Real Time Forex Transactions." *Working Paper*, City University Business School, London, UK (2001).

[73] Hjort, N.L., "Semiparametric Estimation of Parametric Hazard Rates." In *Survival Analysis: State of the Art*, Eds. J.P. Klein and P.K. Goel. Kluwer Academic Publishers, Dordrecht, Netherlands (1992), 211–236.

[74] Hollifield, B., R.A. Miller, and P. Sandås. "An Empirical Analysis of Limit Order Markets." *Working Paper*, Carnegie Mellon University, Pittsburgh, PA (1999).

[75] Huang, R.D. and H.R. Stoll, "The Components of the Bid-Ask Spread: A General Approach." *Review of Financial Studies*, **10** (1997), 995–1034.

[76] Jackson, D.A., "Stopping Rules in Principal Components Analysis: A Comparison of Heuristical and Statistical Approaches." *Ecology*, **74** (1993), 2201–14.

[77] Jasiak, J., "Persistence of Intertrade Durations." *Finance*, **19** (1998), 166–195.

[78] Jolliffe, I.T.. *Principal Component Analysis* (2nd Ed.) Springer-Verlag, New York (2002).

[79] Jones, C., G. Kaul, and M. Lipson, "Transactions, Volume and Volatility." *Review of Financial Studies*, **7** (1994), 631–651.

[80] Jones, M.C., "Simple Boundary Correction for Kernel Density Estimation." *Statistics and Computing*, **3** (1993), 135–146.

[81] Kalbfleisch, J.D. and R.L. Prentice, "Marginal Likelihoods Based on Cox's Regression and Life Model." *Biometrika*, **60** (1973), 267–278.

[82] Kalbfleisch, J.D. and R.L. Prentice. *The Statistical Analysis of Failure Time Data.* John Wiley & Sons, New York (1980).

[83] Killeen, W., R. Lyons, and M. Moore, "Fixed Versus Floating Exchange Rates: Lessons from Order Flow." *Working Paper*, University of California at Berkeley (2000).

[84] Lancaster, T., *The Econometric Analysis of Transition Data.* Cambridge University Press, Cambridge, U.K. (1990).

[85] Lequeux, P., "Real Trading Volume and Price Action in the Foreign Exchange Markets." In Knight, J. and S. Satchell (Eds.), *Forecasting Volatility in the Financial Markets.* Butterworth–Heinemann Press, Oxford, UK (1998), 117–136.

[86] Levich, R.M. and L.R. Thomas, "The Significance of Technical Trading-Rule Profits in the Foreign Exchange Market: A Bootstrap Approach," *Journal of International Money and Finance*, **12** (1993), 451–474.

[87] Lin, D. Y. and L. J. Wei, "The Robust Inference for the Cox Proportional Hazards Model." *Journal of American Statistical Association*, **84** (1989), 1074–78.

[88] Lo, A.W., "Maximum Likelihood Estimation of Generalized Itô Processes with Discretely Sampled Data." *Econometric Theory*, **4** (1988), 231–247.

[89] Lo, A.W., A.C. MacKinlay, and J. Zhang, "Econometric Models of Limit-Order Executions." *Journal of Financial Economics*, **65** (2002), 31–71.

[90] Lo, A., H. Mamaysky, and J. Wang, "Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation." *Journal of Finance*, **55** (2000), 1705–65.

[91] Lunde, A., A. Timmermann, and D. Blake, "The Hazards of Mutual Fund Underperformance: A Cox Regression Analysis." *Journal of Empirical Finance*, **6** (1999), 121–152.

[92] Luo, J., "Market Conditions, Order Flow and Exchange Rate Determination." *Working Paper*, London School of Economics, Financial Market Group (2001).

[93] Lyons, R., "Tests of Microstructural Hypotheses in the Foreign Exchange Market." *Journal of Financial Economics*, **39** (1995), 321–351.

[94] Lyons, R., "Comment on Goodhart *et al.* [54]." In Frankel, J., G. Galli, and A. Giovannini (Eds.), *The Macrostructure of Foreign Exchange Markets*. The University of Chicago Press (1996), 180–182.

[95] Lyons, R., "Foreign Exchange Volume: Sound and Fury Signifying Nothing?" In Frankel, J., G. Galli, and A. Giovannini (Eds.), *The Macrostructure of Foreign Exchange Markets*. The University of Chicago Press (1996), 183–201.

[96] Lyons, R., *The Microstructure Approach to Exchange Rates*. MIT Press, Cambridge, MA (2001).

[97] Lyons, R.K., "New Perspective on FX Markets: Order-Flow Analysis." *International Finance*, **4** (2001), 303–320.

[98] Marron, J.S., "Bootstrap Bandwidth Selection." In Le Page, P. and L. Billard (Eds.), *Exploring the Limits of Bootstrap*. John Wiley and Sons (1992), 249–262.

[99] Marsh and Rock, "The Transaction Process and Rational Stock Price Dynamics." *Working Paper*, University of California at Berkeley (1986).

[100] McLachlan, G. and D. Peel, *Finite Mixture Models*. John Wiley & Sons, New York (2000).

[101] Meese, R. and K. Rogoff, "Empirical Exchange Rate Models of the Seventies, Do They Fit Out of Sample?" *Journal of International Economics*, **14** (1983), 3–24.

[102] Meyer, B.D., "Unemployment Insurance and Unemployment Spells." *Econometrica*, **58** (1990), 757–782.

[103] Müller, H.-G. and J.-L. Wang, "Hazard Rate Estimation Under Random Censoring with Varying Kernels and Bandwidths." *Biometrics*, **50** (1994), 61–76.

[104] Murphy, A.H., "A New Vector Partition of the Probability Score." *Journal of Applied Meteorology*, **12** (1973), 595–600.

[105] Niederhoffer, V. and M. Osborne, "Market Making and Reversal on the Stock Exchange." *Journal of the American Statistical Association*, **61** (1966), 897–916.

[106] Nielsen, J.P., "Marker Dependent Kernel Hazard Estimation from Local Linear Estimation." *Scandinavian Actuarial Journal*, **2** (1998), 113–124.

[107] Nielsen, G. G., R. D. Gill, P. K. Andersen, and T. I. A. Sørensen, "A Counting Process Approach to Maximum Likelihood Estimation in Frailty Models." *Scandinavian Journal of Statistics*, **19** (1992), 25–43.

[108] Nielsen, J.P. and O.B. Linton, "Kernel Estimation in a Nonparametric Marker Dependent Hazard Model." *Annals of Statistics*, **23** (1995), 1735–1748.

[109] O'Hara, M., *Market Microstructure Theory*. Blackwell Publishers, Cambridge, U.K. (1995).

[110] Osler, C.L., "Currency Orders and Exchange Rate Dynamics: Explaining the Success of Technical Analysis." *Working Paper*, Federal Reserve Bank of New York (2001).

[111] Osler, C.L., "Stop-Loss Orders and Price Cascades in Currency Markets." *Staff Report* #150, Federal Reserve Bank of New York (2002).

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

[113] Parner, E., "Asymptotic Theory for the Correlated Gamma-Frailty Model." *Annals of Statistics*, **26** (1998), 183–214.

[114] Phillips, P.C.B., "The Problem of Identification in Finite Parameter Continuous Time Models." *Journal of Econometrics*, **1** (1973), 351–362.

[115] Ramlau-Hansen, H., "Smoothing Counting Process Intensities by Means of Kernel Functions." *Annals of Statistics*, **11** (1983), 453–466.

[116] Ridder, G. and I. Tunali, "Stratified Partial Likelihood Estimation." *Journal of Econometrics*, **92** (1999), 193–232.

[117] Royston, P. and M.K.B. Parmar, "Flexible Parametric Proportional-Hazards and Proportional-Odds Models for Censored Survival Data, with Application to Prognostic Modelling and Estimation of Treatment Effects." *Statistics in Medicine*, **21** (2002), 2175–97.

[118] Russell, J.R., "Econometric Modeling of Multivariate Irregularly-Spaced High-Frequency Data." *Working Paper*, Graduate School of Business, University of Chicago (1999).

[119] Russell, J.R., R. Tsay, and M. Zhang, "A Nonlinear Autoregressive Conditional Duration Model with Applications to Financial Transactions Data." *Journal of Econometrics*, **104** (2001), 179–207.

[120] Singpurwalla, N.D., "Survival in Dynamic Environments." *Statistical Science*, **10** (1995), 86–103.

[121] Snyder, D. and M. Miller, *Random Point Processes in Time and Space.* Springer-Verlag, New York (1991).

[122] Spierdijk, L., T.E. Nijman, and A.H.O. van Soest, "Modeling Comovements in Trading Intensities to Distinguish Sector and Stock Specific News." *Discussion Paper*, CentER, Tilburg University (2002).

[123] Spierdijk, L., T.E. Nijman, and A.H.O. van Soest, "The Price Impact of Trades in Illiquid Stocks in Periods of High and Low Market Activity." *Discussion Paper*, CentER, Tilburg University (2002).

[124] Swets, J.A., "The Relative Operating Characteristic in Psychology." *Science*, **182** (1973), 990–1000.

[125] Wand, M.P., J.S. Marron, and D. Ruppert, "Transformations in Density Estimation." *Journal of the American Statistical Association*, **86** (1991), 343–353.

[126] Wilks, D.S., *Statistical Methods in the Atmospheric Sciences.* Academic Press, San Diego, CA (1995).

[127] Yates, J.F., "External Correspondence: Decompositions of the Mean Probability Score." *Organizational Behavior and Human Performance*, **30** (1982), 132–156.

[128] Zhou, B., "Forecasting Foreign Exchange Rates Subject to Devolatilization." In Dunis, C. (Ed.), *Forecasting Financial Markets: Exchange Rates, Interest Rates and Asset Management.* John Wiley & Sons (1996), 51–67.