Trading activity and liquidity supply in a pure limit order book market: An empirical analysis using a multivariate count data model.

Grammig, Joachin and Heinen, Andreas and Rengifo, Erick

CORE Discussion paper

August 2004

Trading activity and liquidity supply in a pure limit order book market

An empirical analysis using a multivariate count data model *

Joachim Grammig† **Andréas Heinen**‡ **Erick Rengifo**§

August 2004

JEL Classification codes: C32, C35, G10.

Keywords: Market microstructure; Liquidity; Trading activity; Multivariate count process.

*We would like to thank Luc Bauwens, Helena Beltran, Rob Engle, Thierry Foucault, Pierre Giot, Christian Gouriéroux, Joel Hasbrouck, Bruce Lehmann, Craig Lewis and the participants of the Economics and Econometrics of Market Microstructure Summer School 2004 at University of Konstanz for helpful discussions and suggestions. Uwe Schweickert's expertise in the Xetra system provided an invaluable asset for our project. We also wish to thank Helena Beltran for sharing her computer code, which was of great help for our computations. The usual disclaimers apply.
This work is supported by the European Community's Human Potential Programme under contract HPRN-CT-2002-00232 Microstructure of Financial Markets in Europe.

†Eberhard Karls University of Tübingen.

‡Department of Statistics and Econometrics, University Carlos III, Madrid and Center of Operations Research and Econometrics, Catholic University of Louvain.

§Corresponding author. Center of Operations Research and Econometrics, Catholic University of Louvain, 34 Voie du Roman Pays, 1348 Louvain-la-Neuve, Belgium. Tel: (+32) 10.47.43.58, fax: (+32) 10.47.43.01, e-mail: rengifo@core.ucl.ac.be

Abstract

In this paper we perform an empirical analysis of the trading process in a pure limit order book market, the Xetra system which operates at various European exchanges. We study how liquidity supply and demand as well as price volatility affect future trading activity and market resiliency, and discuss the results in the light of predictions implied by theoretical models of financial market microstructure. Parameter estimation and hypotheses testing is conducted using a new econometric methodology designed for the analysis of multivariate count processes.

1 Introduction

The most important stock markets of continental Europe are organized as electronic open limit order book markets.[1] Unlike traditional stock markets, most prominently the New York Stock Exchange, no specialist is responsible for managing liquidity supply and demand. Whether or not a trader asking for immediate execution of an order has to incur a volume dependent price adjustment depends on the state of the open limit order book, which consists of previously submitted, non executed buy and sell orders. The arrival of new information induces traders to cancel, revise and (re)submit limit and market orders which implies that the open limit order book is permanently in flux. The resiliency of such a market design is crucial both for the operator of the trading venue and the agents participating in the trading process. Microstructure theory has put forth a variety of hypotheses about how information events affect liquidity supply and demand in open limit order book markets. The availability of detailed transaction data makes it possible to test these predictions, assess market resiliency, and draw conclusions for market design.

This paper uses data from the Xetra system, a pure limit order book market which operates at several exchanges in Continental Europe, to test hypotheses and empirically assess predictions of microstructure models. We identify liquidity and informational factors describing the state of the limit order book and show how these factors, as well as volatility and liquidity demand, affect future trading activity and market resiliency. For these purpose we use a new econometric methodology, a dynamic model for multivariate time series of counts introduced by Heinen and Rengifo (2003).

This is not the first paper that deals with those issues. Related work has focussed on whether a trader chooses a market or limit order, and how market conditions affect these choices (Biais, Hillion, and Spatt (1995), Griffiths, Smith, Turnbull, and White (2000), Ranaldo (2003) Cao, Hansch, and X. (2004) and Pascual and Veredas (2004)). Sandas (2001) uses Swedish order book data and estimates a version of the celebrated Glosten (1994) limit order book model. and analyze the limit order book dynamics at the Australian and the Spanish Stock Exchange, respectively. Degryse, de Jong, Ravenswaaij, and Wuyts (2003) and Gomber, Schweickert, and Theissen (2004) analyze the resiliency of a pure limit order market by investigating the order flow around aggressive orders using data from Paris Bourse and the German Stock exchange, respectively. The present paper links and contributes to the literature in the following ways. As in Biais, Hillion, and Spatt (1995) we study in detail the trading process in an electronic limit order market. Following their approach we categorize limit orders according to their aggressiveness and study the interdependence of order submission, execution and cancelation processes. Additionally, we distinguish less aggressive limit orders in terms of their relative position in the limit order book with respect to the best quotes. We show that this constitutes an improvement over the categories proposed in Biais, Hillion, and Spatt (1995) as the analysis of the disaggregated order categories provides new insights into the trading process. The detailed analysis is possible since we can exploit the information of a complete record of submission/cancelation/execution events (referred to as "market events") of different types of orders over a three month period. The market events we are particularly interested in are market order entries, limit and market order submissions and cancelations. Using these data and implementing the trading rules of the electronic market, we are able to reconstruct the prevailing order book at any point in time. No hidden orders were allowed during the sample period which implies that market participants and econometricians have an unobstructed (ex post) view of the entire order book.

The main empirical results can be summarized as follows. As predicted by theoretical models of financial market microstructure (Foucault (1999), Handa and Schwartz (1996)) we find that larger spreads reduce the relative importance of market order trading compared to limit order submissions. Consistent with Parlour's (1998) theoretical model, depth at the best quotes stimulates the submission of aggressive limit orders on the same side of the market, as limit order traders strive for price priority. On the other hand, larger depth on the opposite side of the market reduces the aggressiveness of own-side limit orders.

We follow Beltran-Lopez, Giot, and Grammig (2004) and employ the Principal Components Analysis (PCA) methodology to extract latent factors which can explain a considerable fraction of the variation of market liquidity. Consistent with hypotheses derived from the theoretical analyses in Foucault (1999) and Handa, Schwartz, and Tiwari (2003) we find that the first two extracted principal components, associated with a latent "liquidity' and an "informational" factor, can predict future trading activity. If the informational factor indicates a "bad news" state, aggressive limit and market sell order trading increases while buyer activity decreases. In line with theoretical predictions we also find that order aggressiveness is reduced and cancelation activity increases when price volatility is high. Evidence for market resiliency in this automated auction market is provided by the result that an increase in liquidity demand induces an increase in limit order submission activity. Furthermore, we show that cancelations do matter in the sense that they carry information for predicting future market activity and liquidity supply.

The methodological challenge when modeling financial transactions data is the irregular spacing of the multivariate time series data (see Hasbrouck (1999) for a useful discussion). The count data methodology employed in the present paper avoids the caveats of discrete choice models (see e.g. Ranaldo (2003)), in which time series aspects cannot adequately be taken into account, and the drawbacks of financial duration models for which it is difficult to formulate multivariate specifications (see e.g. Bauwens and Hautsch (2003), Engle and Lunde (2003) and Russell (1999)).

The remainder of the paper is organized as follows. Section 2 describes the market structure. Section 3 presents the data and Section 4 explains the econometric methodology. Section 5 discusses the empirical results. Section 6 concludes and provides an outlook for future research.

2 Market Structure

We use data from the automated auction system Xetra. After its introduction at the Frankfurt Stock Exchange (FSE) in 1997, Xetra has become the main trading venue for

German blue chip stocks. The Xetra system is also the trading platform of the Dublin and Vienna stock exchanges as well as the European Energy exchange. The Xetra system operates as a pure electronic order book market. The computerized trading protocol keeps track of the entries, cancelations, revisions, executions and expirations of market and limit orders. For blue chip stocks there are no dedicated market makers, like the specialists at the New York Stock Exchange (NYSE) or the Japanese saitori. For some small capitalized stocks listed in Xetra there may exist so-called Designated Sponsors - typically large banks - who are requiered to provide a minimum liquidity level by simultaneously submitting competing buy and sell limit orders.

Xetra/FSE does face some local competition for order flow. The FSE maintains a parallel floor trading system, which bears some similarities with the NYSE. Furthermore, like in the US, some regional exchanges participate in the hunt for liquidity. However, due to the success of the Xetra system, the FSE floor, previously the main trading venue for German blue chip stocks, as well as the regional exchanges became less important. The same holds true for the regional exchanges. Initially, Xetra trading hours at the FSE extended from 8.30 a.m to 5.00 p.m. CET. From September 20, 1999 the trading hours were shifted to 9.00 a.m to 5.30 p.m. CET. The trading day begins and ends with call auctions and is interrupted by another call auction which is conducted at 12.00 p.m. CET. The regular, continuous trading process is organized as a double auction mechanism with automatic matching of orders based on price and time priority.[2] Five other Xetra features should be noted.

- Assets are denominated in euros, with a decimal system, which implies a small minimum tick size (1 euro-cent).

- Unlike at Paris Bourse, market orders exceeding the volume at the best quote are allowed to "walk up the book". At Paris Bourse the volume of a market order in excess of the depth at the best quote is converted into a limit order at that price entering the opposite side order book. However, in Xetra, market orders are guaranteed immediate full execution, at the cost of incurring a higher price impact on the trades.

- Dual capacity trading is allowed, i.e. traders can act on behalf of customers (agent) or as principal on behalf of the same institution (proprietary).

- Until March 2001 no block trading facility (like the upstairs market at the NYSE) was available.

- Before 2002, and during the time interval from which our data is taken, only round lot order sizes could be filled during continuous trading hours. A Xetra round lot was defined as a multiple of 100 shares. Execution of odd-lot parts of an order - this is an integer valued fraction of one hundred shares - was possible only during call auctions.

Besides these technical details, the trading design entails some features which render our sample of Xetra data (described in the next section) particularly appropriate for our empirical analysis. First, the Xetra system displays not only best quotes, but the contents of the whole limit order book. This is a considerable difference compared to other systems like the Paris Bourse's CAC system, where only the five best orders are displayed. Second, hidden limit orders (or iceberg orders) were not known until a recent change in the Xetra trading rules that permitted them.[3] As a result, the transparency of liquidity supply offered by the system was quite unprecedented. On the other hand, Xetra trading is completely anonymous, i.e. the Xetra order book does not reveal the identity of the traders submitting market or limit orders.[4]

3 Data

The dataset used for our study contains complete information about Xetra market events, that is all entries, cancelations, revisions, expirations, partial-fills and full-fills of market and limit orders that occurred between August 2, 1999 and October 29, 1999. Due to the considerable amount of data and processing time, we had to restrict the number of assets. Market events were extracted for three blue chip stocks, Daimler Chrysler (DCX), Deutsche

Telekom (DTE) and SAP. At the end of the sample period their combined market capitalization represented 30.4 percent of the German blue chip index DAX 30. The three blue-chip stocks under study are traded at several important exchanges. Daimler-Chrysler shares are traded at the NYSE, the London Stock Exchange (LSE), the Swiss Stock Exchange, Euronext, the Tokyo Stock Exchange (TSE) and at most German regional exchanges. SAP is traded at the NYSE and at the Swiss Stock Exchange. Deutsche Telekom is traded at the NYSE and at the TSE. The stocks are also traded on the FSE floor trading system, but this accounts for less than 5% of daily trading volume in those shares. Trading volume at the NYSE accounts for about 20% of daily trading volume in those stocks. As the prices for the three stocks remained above 30 euros during the sample period, the tick size of 0.01 euros is less than 0.05% of the stock price. Hence, we should not observe any impact of the minimum tick size on prices or liquidity. Starting from the initial state of the order book, we track each change in the order book implied by entry, partial or full fill, cancelation and expiration of market and limit orders and perform a real time reconstruction of the order books. For this purpose we implement the rules of the Xetra trading protocol outlined in Deutsche Börse AG (1999) in the reconstruction program. From the resulting real-time sequences of order books, snapshots were taken at one minute intervals during continuous trading hours.

Following Biais, Hillion, and Spatt (1995) we classify market and limit orders in terms of aggressiveness:

- Category 1: Large market orders, orders that walk up or down the book (BMO-agg and SMO-agg).
- Category 2: Market orders, orders that consume all the volume available at the best quote (BMO-inter and SMO-inter).
- Category 3: Small market orders, orders that consume part of the depth at the best quote (BMO-small and SMO-small).

- Category 4: Aggressive limit orders, orders submitted inside the best quotes (BLO-inside and SLO-inside).
- Category 5: Limit orders submitted at the best quote (BMO-at and SMO-at).
- Category 6: Limit orders submitted outside the best quotes, orders that are below (above) the bid (ask). (BMO-outside and SMO-outside).
- Category 7: Cancelations. (BCANC and SCANC)

Moreover, we break up categories 6 and 7 according to their relative position in terms of the number of quotes away from the best quote:

- Limit Orders submitted within the first two quotes away from the best quotes (BLO-outside-1-2 and SLO-outside-1-2).
- Limit Orders submitted within the third and fifth quotes away from the best quotes (BLO-outside-3-5 and SLO-outside-3-5).
- Limit Orders submitted outside the best quotes beyond the fifth quote from the inside market (BLO-outside-5+ and SLO-outside-5+).
- Cancelations of standing limit orders at, or one or two quotes away from the best quotes (BCANC-0-2 and SCANC-0-2).
- Cancelations of standing limit orders between the third and the fifth quotes away from the best quotes (BCANC-3-5 and SCANC-3-5).
- Cancelations of standing limit orders beyond the fifth quote away from the best quotes (BCANC-5+ and SCANC-5+).

For our empirical analysis we then count the submission/cancelation events in the different categories during each one minute interval of the sample. The resulting multivariate sequence of counts provides the input for the econometric model described in the next section.

To avoid dealing with the change in trading times, and given the large number of observations, we restrict the whole sample to observations between August 20 to September 20, 1999. The data therefore contain information about 21 trading days with 510 one-minute intervals per day giving a total of 10730 one minute intervals. Due to space limitations we only report the results for Daimler-Chrysler (DCX).[5] Sample statistics are presented in Table (1) where the main characteristics of the data can be appreciated. The large number of marketable limit orders (MLO) compared to "true" market orders is remarkable. A MLO is a limit order which is submitted at a price which makes it immediately executable. In this respect it is indistinguishable from a "true" market order. However, MLOs differ from market orders in that the submitter specifies a limit of how much the order can walk up the book. Hence, a MLO might be immediately, but not necessarily completely filled. The non-executed volume of the MLO then enters the book.[6] In our empirical analysis we therefore treat the either completely or partially filled parts of an MLO just like a market order. When, for the sake of brevity, we refer in the following to "market orders" what we precisely mean is "true market orders and completely/partially filled marketable limit orders". The number of buy (sell) limit orders is 3.35 (4.7) times larger than the number of market orders. As one can see from table 1, the sample means of the counts series are very small and all series are overdispersed (the sample variance is greater than the sample mean). This has implications for the appropriate statistical specification.

Table (2) presents the descriptive statistics for Daimler-Chrysler (DCX) in which the limit orders submitted outside the best quotes have been further disaggregated according to their relative position to the inside market, as well as descriptive statistics on cancelations, also categorized relative to the best quotes.

[Please insert Table 1 around here]

[Please insert Table 2 around here]

Figure (1) presents two-day auto- and cross-correlograms of the aggregated count series for Daimler-Chrysler (DCX). We consider buy market orders (BMO), sell market orders

(SMO), buy limit orders (BLO), sell limit orders (SLO), buy cancelations (BCANC) and sell cancelations (SCANC). Observing the autocorrelations one can see that all series of counts show persistence. A visual inspection of the cross correlations between market buys and market sells reveals that they are almost symmetric. This implies that the tendency of market buys at time t to follow market sells at time $t-k$ is almost the same as the tendency of market sells to follow market buys. This indicates that the informational effects, found by Hasbrouck (1999), are not detectable in our data.

[Please insert Figure 1 around here]

Figure (2) depicts the intraday seasonality in the series of market event counts. Neither buy nor sell market order counts reflect the often reported U-shape of intra-day financial series. There is a small increase in the number of counts at about 2.30 p.m. CET which most likely corresponds to the NYSE opening. The number of buy limit orders is large early in the morning, but decays quite fast. Limit orders at both sides of the book behave similarly in that we observe an increase in trading activity in the afternoon at the same time as the market order activity increases. We observe a similar diurnal pattern in the cancelation series.

[Please insert Figure 2 around here]

4 Methodology

In order to model the dynamics of the multivariate series of counts of order submissions and cancelations within one minute intervals, we adopt the Multivariate Autoregressive Conditional Double Poisson (MDACP) modeling framework introduced by Heinen and Rengifo (2003). In the following we briefly sketch the econometric specification and the estimation strategy. A more detailed exposition can be found in the appendix.

Collecting the one-minute-interval submission and cancelation counts at time t in a $K-$dimensional vector $N_t = (N_{1,t}, N_{2,t}, \ldots, N_{K,t})'$, the MDACP sets up a VARMA-type system for the conditional mean vector $E[N_t|\mathcal{F}_{t-1}] \equiv \mu_t = (\mu_{1,t}, \ldots, \mu_{K,t})'$,

$$\mu_t = \omega + \sum_{j=1}^{p} A_j N_{t-j} + \sum_{j=1}^{q} B_j \mu_{t-j}, \tag{1}$$

where ω, A_j and B_j are parameter vectors and matrices, respectively. One might ask, why not estimate a Gaussian Vector Autoregression to keep the econometrics simple? The reason is that, as indicated by the descriptive statistics, the submission and cancelation count series exhibit very small means. This renders the assumption of a continuous, symmetric distribution clearly inappropriate. The discreteness of the data definitely has to be accounted for. Another feature of the data complicates the formulation of an appropriate statistical model. The descriptive analysis shows that most of the one-minute count sequences are overdispersed, i.e. the empirical variance is greater than the mean. It can be shown that the autoregressive specification (1) already generates some overdispersion, but to tie together the two main features of the data, autocorrelation and (unconditional) overdispersion, seems to be a restrictive modeling strategy. To provide the necessary flexibility we employ the Double Poisson distribution (DP) introduced by Efron (1986). The advantage of the Double Poisson compared to the Poisson distribution is that it can be under- and overdispersed, depending on whether a dispersion parameter (ϕ) is larger or smaller than one.

Accordingly, we assume that the distribution of the $i-$th count series $N_{i,t}$, conditional on the information set $\mathcal{F}_{t-1}$, is the Double Poisson

$$N_{i,t}|\mathcal{F}_{t-1} \sim DP(\mu_{i,t}\ , \phi_i)\ ,\ \forall i = 1, \ldots, K. \tag{2}$$

where ϕ_i is the dispersion parameter associated with the i-th count series. Transferring Efron's (1986) results it is easy to show that the conditional variance of the count $N_{i,t}$ is given by

$$V[N_{i,t}|\mathcal{F}_{t-1}] = \sigma^2_{i,t} = \frac{\mu_{t,i}}{\phi_i}. \tag{3}$$

Besides the VARMA dynamics in equation (1) we allow a vector of predetermined variables observed at $t-1$, and collected in a vector X_{t-1}, to impact on the conditional mean $E(N_{i,t}|\mathcal{F}_{t-1})$ of the one-minute submission/cancelation count. The predetermined variables are derived from models of market microstructure and include liquidity and informational indicators that can be extracted from the order book information and transaction data (e.g. inside spread, depth and volatility). Furthermore, to account for intra-day seasonality (or "diurnality") of the count sequences, we include a trigonometric spline function in the conditional mean equation. This method has been advocated and successfully applied by Andersen and Bollerslev (1997) to account for diurnality in volatility models. Including both predetermined variables and accounting for seasonality as outlined above, the conditional distribution of $N_{i,t}$ in equation (2) becomes

$$N_{i,t}|\mathcal{F}_{t-1} \sim DP(\mu^*_{t,i}\ , \phi_i)\ ,\ \forall i = 1, \ldots, K. \tag{4}$$

where

$$\mu^*_{t,i} = \mu_{t,i} \exp\left(X'_{t-1}\gamma_i + \sum_{p=1,2}\left(\psi_{c,p}\cos\frac{2\pi p\, Re[t,N]}{N} + \psi_{s,p}\sin\frac{2\pi p\, Re[t,N]}{N}\right)\right) \tag{5}$$

The first term in the exponent accounts for the effect of the predetermined variables X_{t-1} on the conditional mean, where γ_i is a parameter vector. The second term is the trigonometric spline function, where $Re[t, N]$ is the remainder of the integer division of t by N, the number of one-minute periods in a trading session. $\psi_{c,p}$ and $\psi_{c,q}$ are parameters to be estimated.

We employ a multivariate Gaussian copula to account for contemporaneous cross-correlation in the count sequences. The appendix shows how this facilitates writing down the likelihood function. Adopting the two step method outlined by Patton (2002), the parameters can straightforwardly be estimated by Maximum Likelihood. The appendix describes the details of the two step estimation procedure and provides further information about the use of copulas to account for contemporaneous dependencies between the count sequences.

Specification tests can be conducted based on the usual likelihood statistics, but conveniently also by analyzing the properties of the "Pearson residuals", which are defined as $\epsilon_{i,t} = \frac{N_{i,t}-\mu_{i,t}}{\sigma_{i,t}}$. When a model is correctly specified, the estimated Pearson residuals should have an empirical variance close to one and exhibit no significant autocorrelation. The appendix discusses an additional specification test based on probability integral transforms as suggested in Diebold, Gunther, and Tay (1998).

5 Empirical Results

5.1 Parameter estimates and specification tests

Estimation and test results are reported in tables 4, 5, 6, 7, 8 and 9. Table 4 contains the results for an MDACP model with six endogenous count variables: buy market orders (BMO), buy limit orders (BLO), sell market orders (SMO), sell limit orders (SLO), buy order cancelations (BCANC) and sell order cancelations (SCANC). This specification (henceforth referred to as the aggregated model) will already be useful to test several predictions of theoretical microstructure. Tables 5 (bid side) and 6 (ask side) report the estimation results for a disaggregated MDACP system, where order counts are classified, according to aggressiveness, into the six categories described in Section 3. Table 7 presents the results of a bivariate MDACP model for buy and sell market orders in which lagged cancelation counts enter as predetermined variables. Table 8 reports the results of an MDACP model which focuses on the counts of the three limit order categories (LO-inside, LO-at, and LO-outside) and that also uses lagged cancelation counts as predetermined variables. To obtain the results reported in table 9 we estimated an MDACP model which is based on a finer categorization of limit orders outside the best bid.

In all tables we report the parameters of the autoregressive parameters (β), the parameters on the lagged counts (α), the parameters which determine the impact of the predetermined variables on the expected number of counts (γ), and the estimated dispersion parameters (ϕ). Significant (at 5 %) parameter estimates are printed in boldface. The last rows of the estimation results tables report the empirical variance of the Pearson residuals. Because of space limitations we refrain from presenting the parameter estimates for the seasonality model. Instead, we report the p-value of the Wald statistic ($W(\psi's = 0)$) for a test of the joint significance of the seasonality parameters. Under the null hypothesis the test statistic is distributed Chi-square with four degrees of freedom. Except for two cases, the Wald statistic is highly significant, underlining the necessity to account for diurnality in the count sequences. We have outlined above that a correctly specified model implies that the Pearson residuals have variance close to one and exhibit no significant autocorrelation. Inspecting the estimated variances of the Pearson residuals in the results tables and the sample autocorrelogram of the Pearson residuals (aggregated system) in figure 3 we find no evidence for specification problems.[7] Following the suggestions of Diebold, Gunther, and Tay (1998) we also employed graphical tools to check for uniformity and serial dependence in the probability integral transform (PIT) sequences. The visual inspections did not point to specification problems, as the Q-Q plots almost coincide with the 45-degree line and the empirical autocorrelograms of the PIT sequences do not indicate serial correlation.[8]

[Please insert 3 around here]

The estimation results indicate a clear rejection of the Poisson assumption as all estimated dispersion coefficients are significantly different from one. The distributions are either over- or underdispersed, supporting the use of the Double Poisson distribution.

Table 3 reports the estimated contemporaneous correlation matrix of the quantile vector q_t implied by the aggregated MDACP system. The appendix shows that this correlation measure the part of the contemporaneous and lagged cross-correlation which does not go through the time-varying mean. With a single exception, all correlations are positive and especially the own-side correlations of limit order submissions and cancelations are considerable. This indicates that an increase in trading activity generally involves all types of market events, but that the same side dependence is stronger. The result that market sell and buy order events are negatively correlated is quite expected.

[Please insert Table 3 around here]

5.2 Discussion

5.2.1 Liquidity supply, volatility and order submission activity

Inside Spread and depth, and trading activity

Theoretical models put forth by Handa and Schwartz (1996) and Foucault (1999) hypothesize that large spreads reduce the proportion of market orders relative to limit orders in the total order flow. The explanation is that a larger spread implies a higher price of immediacy. This makes market orders less attractive than limit orders which receive a higher premium for providing liquidity. Griffiths, Smith, Turnbull, and White (2000) and Ranaldo (2003) have provided empirical evidence for these predictions. The estimation results for the aggregated MDACP system (table 4) indicate that an increase of the inside spread exerts a negative effect on all six order categories and cancelations, thus inducing a general slowdown in trading activity. In line with theory, the impact on market orders is considerably stronger than the effect on limit orders. The estimation results for the disaggregated system (tables 5 and 6) lead to the same conclusions. The empirical analysis thus confirms the theoretical prediction that the proportion of market orders decreases and the proportion of limit orders increases when large spreads prevail.

[Please insert Table 4 around here]

[Please insert Table 5 around here]

[Please insert Table 6 around here]

In the models proposed by Parlour (1998) and Handa, Schwartz, and Tiwari (2003) the volume (depth) at the best quotes is related to the execution probability of limit orders at the respective side of the book, which in turn affects trading activity. More precisely, it is predicted that when the execution probability of a limit order is low, traders on the respective side of the market act more aggressively when striving for price-time priority. A large volume at the best quote (at the bid side, say) will induce bid-side traders to act aggressively by submitting more market orders or limit orders inside the best quotes. On the other hand, when the depth at the **opposite** side of the market is large, **own side** order aggressiveness is expected to decrease. This is a mechanical consequence of the previous result. Coming back to the example, large volume at the bid-side, which induces bid-side traders to submit more aggressive buy limit orders, increases the probability of execution of limit orders at the ask side relative to market orders, thereby decreasing aggressiveness on the opposite side. The empirical evidence for these hypotheses obtained from the estimation of the aggregated MDACP system is mixed. Table 4 shows that volume at the best quotes (denoted BIDVOL and ASKVOL) exerts a positive effect on all components of the order flow. Larger volume at the best quotes does not only have a positive effect on own side trading activity, but also on the opposite side. While the own side effect is in line with the theoretical predictions outlined above, the opposite side effect is clearly not. The estimation results of the disaggregated MDACP system presented in tables 5 and 6 are more in accordance with the theoretical predictions. As hypothesized, the empirical results confirm that traders on the respective side of the market act more aggressively when the volume at the best quote is large. For example, when depth at the bid is large, traders are expected to submit more buy limit orders inside the best quotes. As predicted, volume at the bid exerts a positive effect on the expected number of buy market orders of the most aggressive categories. The ask side results are quite similar. The opposite side effects are now also in accordance with the theoretical predictions. For example, an increase of the volume at the best bid decreases the expected number of most aggressive sell market orders. In other words, own-side order aggressiveness tends to decrease when opposite-side depth at the best quote increases, as hypothesized.

Beyond the inside market: Liquidity and informational factors, and trading activity

Beltran-Lopez, Giot, and Grammig (2004) propose to employ Principal Components Analysis (PCA) for the analysis of commonalities in the limit order book. We adopt their approach to analyze the impact of the order book state beyond the inside market on trading activity. The basic idea is to compute the hypothetical unit price of a market order of volume v if it were executed immediately against the time t order book. Dividing the unit price by the best quote prevailing at time t yields the relative price impact. In our application the relative price impact is computed for v=3,000 to 40,000 with $1,000$ shares increments. PCA is then employed to summarize the information using a small number of factors (principal components) which are, by construction, uncorrelated.[9] As in Beltran-Lopez, Giot, and Grammig (2004) we find that three factors for each side of the order book suffice to account for the variation in the price impact. The first principal component has nearly constant loadings for all volumes v. An increase of this factor, given the positivity of the factor weights, implies that the book is depleted and thus the percentage relative price impact increases.[10] The second factor is negatively related to the price impacts at small volumes, with factor loadings increasing monotonically with v. In other words, an increase in the second factor induces the slope of the price impact curve to become steeper. A steep slope of the book indicates that limit order traders are more cautious and want to protect themselves against information based trading by submitting less aggressive limit orders. The second principal component can therefore be interpreted as an "informational" factor.

The extracted principal components can conveniently be used to test hypotheses found in the theoretical papers and previous empirical findings. Tables 5 and and 6 present results of an MDACP model where the first principal component from each market side is used as an explanatory variable. Biais, Hillion, and Spatt (1995) find that investors provide liquidity to the market when it is valuable and consume liquidity when it is plentiful. Our empirical results support this finding. When the first factors (in the results tables denoted as

SFACT1 and BFACT1) increase, the own-side aggressiveness increases, in that traders use more aggressive limit orders to replenish the book. Order aggressiveness on the opposite market side is also increased as the favorable price impacts stimulate the submission of opposite-side market orders. In that sense, this factor is related with the liquidity provision and consumption in the markets: liquidity is offered when it is needed and consumed when the book is filled.

Hall, Hautsch, and Mcculloch (2003) point out that this liquidity effect, which stimulates overall trading activity, has to be distinguished from an informational effect for which the theoretical predictions are quite different. In the theoretical models of Foucault (1999) and Handa, Schwartz, and Tiwari (2003) an imbalance in the order book with a steep buy side and flat sell side order book indicates a bad news state in which prospective buy side traders act cautiously, by submitting buy limit orders away from the best bid, while sellers are expected to submit market orders and aggressive sell limit orders. To test this hypothesis, we construct a convenient indicator by taking the difference of the absolute values of the bid and ask side informational factors extracted by the PCA. This indicator (in the results tables denoted DIFFSLOPE) is positive when the ask side of the limit order book is relatively flat and the bid side of the book relatively steep (thus indicating a bad news state). The disaggregated MDACP specification uses this "bad news" indicator as an explanatory variable. The estimation results reported in tables 5 and 6 show that the bad news indicator induces the buy side to become less aggressive, while the sell side acts more aggressively, which is in accordance with the theoretical predictions.

Volatility and order submission activity

Foucault's (1999) theoretical model implies that when volatility increases, limit order traders ask for a higher compensation for the risk of being picked off, i.e. being executed when the market has moved against them. Based on these considerations, Ranaldo (2003) formulates the hypothesis that higher volatility induces less aggressive order submissions. The empirical evidence regarding this hypothesis is found in Bae, Jang, and Park (2003) and Danielson and Payne (2001) that find that traders place more limit orders respect to market orders when volatility is high, in the same direction Griffiths, Smith, Turnbull, and White (2000) and Ranaldo (2003) report less aggressive trades when temporary volatility increases. In order to test the hypothesis in the MDACP framework, we measure volatility as the standard deviation of the midquote returns during the last 5 minutes and include this indicator as a predetermined variable (in the results tables denoted VOLAT). The estimation results for the aggregated system in table 4 are in accordance with the predictions expected from theory. Volatility exerts a negative impact on the most aggressive orders (market orders) and a positive impact on less aggressive orders (limit orders). Moreover, volatility impacts positively on cancelations on both sides of the book, which is in line with the prediction that as volatility increases, traders cancel their positions more frequently to avoid being picked off.

The estimation results for the disaggregated system in tables 5 and 6 reconfirm these conclusions and provide a more detailed view. Volatility affects negatively and significantly the submission intensity of most aggressive market orders (category one) and negatively, but not significantly, the orders of categories two and three. Furthermore, volatility exerts a positive impact on limit orders at or outside the best quotes, but has a negative effect on limit orders inside the best quotes. These results again confirm the theoretical prediction that order aggressiveness decreases when volatility increases.

5.2.2 Order submission dynamics, cancelations, and market resiliency

The VARMA structure of the MDACP model provides a convenient framework to analyze autoregressive dynamics of order submissions and cancelations in an automated auction market in the spirit of the papers by Biais, Hillion, and Spatt (1995) and Bisière and Kamionka (2000). In the following we will exploit this feature in an empirical assessment of market resiliency, particularly with regard to cancelation events.

The estimation results for the aggregated MDACP system in table 4 show that lagged buy (sell) market order counts exert a positive and significant effect on the expected number of sell (buy) limit orders. In other words, when liquidity is consumed by market orders, liquidity suppliers (voluntarily) enter into the market, and new (competitive) limit orders are submitted which replenish the limit order book. These results indicate market resiliency despite the absence of designated market makers. Estimation results for the disaggregated MDACP system lead to the same conclusion as lagged market orders impact positively on all opposite-side limit order categories.[11]

So far, the theoretical literature did not devote a great deal of attention to the role of limit order cancelations in explaining future trading activity. This is surprising, as it seems natural to hypothesize that cancelation events, especially when occurring near the inside market, carry informational content. The estimation results for the aggregated system already provide some empirical evidence for the informational significance of cancelations: Table 4 shows that by affecting own-side expected market order submissions negatively, but by exerting a positive impact on own-side limit order submissions, cancelations tend to reduce own-side order aggressiveness.

[Please insert Table 7 around here]

[Please insert Table 8 around here]

More detailed empirical analyses provide further evidence for the informational content of cancelation events. First, we estimate a bivariate MDACP model for buy and sell market orders set up to study the effect of cancelation events on market order submissions. The results are reported in table 7. Secondly, we estimate an MDACP model designed to assess the effect of cancelations on limit order submissions (see table 8). For both models we categorize the position of the canceled limit order counts relative to the best quotes. The estimation results evidence that, as hypothesized, cancelations close to the inside market are informationally the most important events. These "aggressive" cancelations exert a negative impact on the expected number of own-side market order submissions. Furthermore, they decrease the expected number of most aggressive limit orders (those submitted inside and at the best quotes). However, aggressive cancelations also exert a positive impact on limit order submissions *outside* the best quotes. This leads to the conclusion that the aggressive cancelations induce limit order traders to act more cautious and to demand higher liquidity premia.

[Please insert Table 9 around here]

The estimation results reported in table 9 provide more detailed insights. Here we estimated an MDACP model where the limit order submission category "outside the best quote" is further disaggregated. The results show that aggressive cancelations induce a higher limit order submission activity close to (yet not inside or at) the best quotes. We conclude that, although cancelations negatively affect liquidity quality (by the negative effect on limit orders which provide the best liquidity quality), it is only reduced, and not erased. This again indicates the market resiliency property of this automated auction market.

6 Conclusions and future work

This paper has presented an empirical analysis of the trading process in an automated auction market. For these purpose we use a new econometric methodology, a dynamic model for multivariate time series of counts introduced by Heinen and Rengifo (2003). This econometric methodology is tailor-made to account for the various dimensions of the trading process. Compared with alternative empirical strategies which tackle the natural irregular spacing of the transactions data by formulating a duration model or marked point process our approach delivers results that are much easier to communicate. We have tested several hypotheses put forth by market microstructure. The results that we have obtained using the new methodology both confirm previous findings and offer new insights:

- We have found empirical support for hypothesis that larger spreads reduce the relative importance of market order trading compared to limit order submission activity. Furthermore, we have confirmed the hypothesis that increasing depth at the best quotes stimulates the submission of aggressive limit order at the same side of the market while larger depth on the opposite side of the market reduces the aggressiveness of own-side limit orders. Consistent with theoretical predictions we have found that order aggressiveness is reduced when volatility is high.

- Using a principal components analysis of the order book we have obtained the result that one of the extracted factors, identified as the "informational factor", proved to be very successful in predicting the future order submission process. As predicted both by theory and intuition, we found that if the informational factor indicates "bad news" the number of aggressive sell limit and market orders increases while buyer activity decreases.

- The results indicate the important role that cancelations play for predicting future order submission activity. More precisely, we have found that cancelations of aggressive limit orders (standing orders close to the best quotes) generally reduce the trading activity. However, those "aggressive" cancelations increase the submission activity within the first five quotes, again indicating market resiliency.

There are a number of directions for further research along the lines presented here. A potential extension could use the econometric methodology to study cross-security or cross-trading venues differences. For example, it is tempting to conduct a comparative analysis of the trading activity for assets with different ownership and/or market capitalization. Will the encouraging results regarding market resiliency still hold for small caps? Another idea is to compare trading venues which offer different degrees of pre-trade transparency. Will we still obtain those empirical confirmations of theoretical predictions outlined above if the trading process is less transparent, for example if hidden orders are allowed? The presence of hidden orders disguises part of the liquidity, i.e. blurs "informational" component discussed above.

An extension to issues in international finance is another interesting research direction. For example, Daimler Chrysler is both listed at the NYSE and Xetra/FSE (and other international exchanges, too). As a matter of fact, the DCX globally registered share is traded simultaneously during overlapping trading hours of these international exchanges. In a comparative study one could analyze how the different degrees of pre-trade transparency at Xetra/FSE and NYSE affects the trading process and price discovery for those cross listed stocks. The method presented here is straightforwardly extended to multiple markets, thus offering the possibility to study linkages of international stock markets on much more detailed level.

7 Appendix: Details on the specification and estimation of the MDACP model

We account for contemporaneous dependence in the sequences of one minute counts of order submissions and cancelations by employing a multivariate Gaussian copula. There are a couple of methodological complications associated with this strategy that we will outline below. We start by refreshing some basic results about the use of copulas, before we discuss the complications we have to deal with when working with count data, and present more details about our estimation and specification strategy.

Sklar (1959) showed that the joint distribution of K random variables can be decomposed into the K marginal distributions and an object referred to as copula which accounts for dependence between the variables. More precisely, let $H(y_1, \ldots, y_K)$ denote a continuous K-variate cumulative distribution function (cdf) with univariate marginal cdfs $F_i(y_i)$, $i = 1, \ldots, K$. Sklar (1959), shows that there exists a function C, the copula, mapping from $[0,1]^K$ into $[0,1]$, such that

$$H(y_1, \ldots, y_K) = C(F_1(y_1), \ldots, F_K(y_K)) \; . \tag{6}$$

The joint density function can be written as the product of the marginal densities $f_i(y_i)$ and an object referred to as "copula density",

$$\begin{aligned} \frac{\partial H(y_1, \ldots, y_K)}{\partial y_1 \ldots \partial y_K} &= \prod_{i=1}^{K} f_i(y_i) \frac{\partial C(F_1(y_1), \ldots, F_K(y_K))}{\partial F_1(y_1) \ldots \partial F_K(y_K)} \\ &= \prod_{i=1}^{K} f_i(y_i) \; c(z_1, \ldots, z_K) \; , \end{aligned} \tag{7}$$

$c(z_1, \ldots, z_K)$ is the copula density, where $z_i = F_i(y_i)$, for $i = 1, \ldots, K$. It is a well known result that the distribution of z_i (referred to as probability integral transforms) is $U(0,1)$ if the marginal distribution F_i is a) correctly specified and b) continuous.[12] We can then write

$$C(z_1, \ldots, z_K) = H(F_1^{-1}(z_1), \ldots, F_K^{-1}(z_K)) \; . \tag{8}$$

Copulas provide a convenient way to generate a valid joint distribution from known marginal distributions. In the following we will sketch how this idea is used for the present paper.

There exist many choices for copulas in the bivariate case, but the number of multivariate copulas is limited. We work with the most prominent candidate, the Gaussian copula (see Sklar (1959)). Assuming correct specification of the marginal distributions, this copula can be written:

$$C(z_1, \ldots, z_K; \Sigma) = \Phi^K(\Phi^{-1}(z_1), \ldots, \Phi^{-1}(z_K); \Sigma) \; , \tag{9}$$

Φ^K denotes the cdf of a K-dimensional standard normally distributed random vector. $\Phi^{-1}()$ is the quantile function of the (univariate) standard normal distribution. Σ denotes the variance covariance matrix of the random vector $q = (q_1, \ldots, q_K)'$, where $q_i = \Phi^{-1}(z_i)$. The corresponding copula density is given by

$$c(z_1, \ldots, z_K; \Sigma) = \mid \Sigma \mid^{-1/2} \exp\left(\frac{1}{2}(q'(I_K - \Sigma^{-1})q\right) \; . \tag{10}$$

The present paper deals with discrete marginal distributions for count data. This complicates the use of copulas to account for (contemporaneous) dependence in the count sequences. First, a copula is uniquely defined only for continuous marginal distributions. In the discrete case this is no longer true. Second, the result that the probability integral transforms z_i are $U(0,1)$ does not hold for discrete random variables. To circumvent these problems we resort to the continuous extension argument put forth by Denuit and Lambert (2002).[13] The basic idea is to create a continuous random variable by adding to the discrete count an independent continuous random variable with support on the $[0,1]$ interval, and with a strictly increasing distribution function. The obvious choice is a $U(0,1)$ random variable.

Parameter estimation of the MDACP model is conducted as follows. Combining the assumption of Double Poisson distributed count sequences $N_{i,t}$ and a multivariate Gaussian copula to account for contemporaneous dependence, the joint density of the vector sequence of counts N_t, conditional on pre-sample values, can be written as

$$h(N_{1,t},\ldots,N_{K,t},\theta,\Sigma)=\prod_{i=1}^{K} f_{DP}(N_{i,t};\mu^*_{i,t},\phi_i)\cdot c(q_t;\Sigma)\ , \tag{11}$$

and $f_{DP}(N_{i,t};\mu^*_{i,t},\phi_i)$ the Double Poisson density as a function of the observation $N_{i,t}$, the conditional mean $\mu^*_{i,t}$ (defined as in equation (5) in the main text), and the dispersion parameter ϕ_i. $c()$ is the Multivariate Gaussian copula density and the vector θ collects all model parameters. As above, we define $q_t = (\Phi^{-1}(z_{1,t}),\ldots,\Phi^{-1}(z_{K,t}))'$, where $z_{i,t}$ denote the probability integral transforms of the continuous extension of the original count data,

$$z_{i,t}=F^*(N^*_{i,t}), \tag{12}$$

where $F^*()$ denotes the cdf of the continuous extension of the count data,

$$N^*_{i,t}=N_{i,t}+(U_{i,t}-1)\ , \tag{13}$$

where $U_{i,t}$ denotes a $U(0,1)$ random variable. Adapting the results of Denuit and Lambert (2002) we can use the relation

$$F^*(N^*_{i,t})=F_{DP}(N_{i,t}-1;\mu^*_{i,t},\phi_i)+f_{DP}(N_{i,t};\mu^*_{i,t},\phi_i)\cdot U_{i,t}, \tag{14}$$

where $F_{DP}()$ denotes the cdf of the Double Poisson, to compute the $z_{i,t}$ series.

Taking the log of the joint density (equation (11)) we obtain the conditional log likelihood function,

$$logL=\sum_{t=1}^{T}\sum_{i=1}^{K} log(f_{DP}(N_{i,t};\mu^*_{i,t},\phi_i))+log(c(q_t;\Sigma)), \tag{15}$$

where T denotes the number of observations.

To estimate the model parameters we adopt a two step procedure that was proposed by Patton (2002). In the first step we maximize the first part of the log-likelihood (15) which, written in detail, is given by

$$\sum_{t=1}^{T}\sum_{i=1}^{K}\left(\frac{1}{2}log\phi_i - \mu_{i,t}^{*}\phi_i - N_{i,t} + (1-\phi_i)log(N_{i,t}^{N_{i,t}}) + N_{i,t}\phi_i(1+log\mu_{i,t}^{*}) - log(N_{i,t}!)\right). \tag{16}$$

Since we employ the multivariate Gaussian copula, the second estimation step does not require any numerical optimization. The maximum likelihood estimate of the variance covariance matrix Σ is simply the sample variance covariance matrix,

$$\hat{\Sigma} = \frac{1}{T}\sum_{t=1}^{T}\hat{q}_t\hat{q}_t^{'} . \tag{17}$$

Besides the methods for specification testing outlined in the main text one can also use the sequence of probability integral transforms as discussed by Diebold, Gunther, and Tay (1998). A correct specification of the marginal density is crucial. Any mistake would invalidate the use of copulas. If a model is correctly specified, however, the sequence of probability integral transforms $\{z_{i,t}\}$ is iid $U(0,1)$. This suggests a convenient way to test the specification of the marginal distribution, as tests for iid uniformity are readily available.

References

Andersen, T. G., and T. Bollerslev, 1997, Intraday periodicity and volatility persistence in financial markets, *Journal of Empirical Finance* 4, 115–158.

Bae, K.-H., H. Jang, and K.S. Park, 2003, Trader's choice between limit and market orders: Evidence from nyse stocks, *Journal of Financial Markets* 6, 517–538.

Bauwens, L., and P. Giot, 2001, *Econometric Modelling of Stock Market Intraday Activity* (Kluwer Academic Publishers).

Bauwens, L., and N. Hautsch, 2003, Dynamic latent factor models for intensity processes, Working paper 2003/103, Center of Operactions Research and Econometrics, Catholic University of Louvain.

Beltran-Lopez, H., P. Giot, and J. Grammig, 2004, Commonalities in the order book, Mimeo, CORE, Universities of Louvain, Namur, and Tübingen.

Biais, B., P. Hillion, and C. Spatt, 1995, An empirical analysis of the limit order book and the order flow in the Paris Bourse, *Journal of Finance* 52, 655–681.

——— , 1999, Price discovery and learning during the preopening period in the paris bourse, *Journal of Political Economy* 107, 1218–1248.

Bisière, C., and T. Kamionka, 2000, Timing of orders, orders aggressiveneess and the order book at the paris bourse, *Annales d'Economie et de Statistique* 56, 43–72.

Cao, C, O. Hansch, and Wang X., 2004, The informational content of an open limit order book, EFA 2004 Working Paper, Penn State University and Southern Illinois University.

Danielson, J., and R. Payne, 2001, Measuring and explaining liquidity on an electronic limit order book: evidence from reuters d200-2, London School of Economics, Working paper.

Degryse, H., F. de Jong, M. Ravenswaaij, and G. Wuyts, 2003, Aggressive orders and the resiliency of a limit order market, Working Paper, K.U. Leuven, Department of Economics.

Denuit, M., and P. Lambert, 2002, Constraints on concordance measures in bivariate discrete data, Discussion Paper 02-12, Institut de Statistique, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.

Diebold, Francis X., Todd A. Gunther, and Anthony S. Tay, 1998, Evaluating density forecasts with applications to financial risk management, *International Economic Review* 39, 863–883.

Efron, Bradley, 1986, Double exponential families and their use in generalized linear regression, *Journal of the American Statistical Association* 81, 709–721.

Engle, R., and A. Lunde, 2003, Trades and quotes: a bivariate point process, *Journal of Financial Econometrics* 1, 159–188.

Foucault, T., 1999, Order flow composition and trading costs in a dynamic limit order market, *Journal of Financial Markets* 2, 99–134.

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

Gomber, P., U. Schweickert, and E. Theissen, 2004, Zooming in on liquidity, EFA 2004 Working Paper, University of Bonn.

Griffiths, M., B. Smith, A. Turnbull, and R. White, 2000, The costs and determinants of aggressiveness, *Journal of Financial Economics* pp. 65–88.

Hall, A., N. Hautsch, and J. Mcculloch, 2003, Estimating the intensity of buy and sell arrivals in a limit order book market, Mimeo, University of Technology, Sidney and University of Konstanz.

Handa, P., R. Schwartz, and A. Tiwari, 2003, Quote setting and price formation in an order driven market, *Journal of Financial Markets* 6, 461–489.

Handa, Puneet, and Robert A. Schwartz, 1996, Limit order trading, *The Journal of Finance* 51, 1835–1861.

Hasbrouck, Joel, 1999, Trading fast and slow: Security markets events in real time, mimeo.

Heinen, A., and E. Rengifo, 2003, Multivariate modeling of time series count data: An autoregressive conditional poisson model, CORE Discussion Paper 2003/25.

Joe, Harry, 1997, *Multivariate Models and Dependence Concepts* (Chapman and Hall: London).

Machado, J., and J.M.C. Santos Silva, 2003, Quantiles for counts, Working Paper, ISEG, Universidade Técnica de Lisboa.

Nelsen, Roger B, 1999, *An introduction to Copulas* (Springer: New York).

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

Pascual, R., and D. Veredas, 2004, What pieces of limit order book information do are informative? an empirical analysis of a pure order-driven market, Mimeo. Dept. of Business. Universidad de las Islas Baleares.

Patton, Andrew, 2002, Skewness, asymmetric dependence, and portfolios, mimeo.

Ranaldo, A, 2003, Order aggressiveness in limit order book markets, *Journal of Financial Markets* 7, 53–74.

Russell, J., 1999, The econometric modeling of multivariate irregularly-spaced high-frequency data, Working Paper Tilburg University and CentER.

Sandas, P., 2001, Adverse selection and competitive market making: Empirical evidence from a limit order market, *Review of Financial Studies* 14, 705–734.

Sklar, A, 1959, Fonctions de répartitions à n dimensions et leurs marges, *Public Institute of Statistics of the Univeristy of Paris* 8, 229–231.

Notes

[1]The largest of these markets is Euronext, the joint venture of the Amsterdam, Brussels and Paris stock exchanges, with a trading volume of 890 billion euro (in stocks) during the first two quarters of 2004, followed by the German stock exchange/Xetra (490 billion euro) and the Swiss SWX/Virt-X trading platform (170 billion euro). Trading volume at the London Stock Exchange during this period amounted to 660 billion euro.

[2]Bauwens and Giot (2001) provide a complete description of an order book market and Biais, Hillion, and Spatt (1999) describe the opening auction mechanism employed in an order book market and corresponding trading strategies.

[3]Biais, Hillion, and Spatt (1995) show that the possibility of hiding part of the volume of a limit order leads to all sorts of specific trading behavior, for example submitting orders to "test" the depth at the best quote for hidden volume.

[4]Further information about the organization of the Xetra trading process and a description of the trading rules that applied to our sample period is provided in Deutsche Börse AG (1999).

[5]The results obtained with the other two assets confirm the findings. These results are available upon request.

[6]MLOs therefore share some properties with Paris Bourse market orders.

[7]To conserve space and since the results are qualitatively identical we do not present the autocorrelograms of all models

[8]To conserve space, we do not display the Q-Q plots and the autocorrelograms. These are available upon request.

[9]Prior to the PCA the data is standardized by substracting time-of-day specific means

and dividing by standard deviation.

[10]We thank Pierre Giot for having pointed out this.

[11] The estimation results of the aggregated and disaggregated MDACP systems (tables 4, 5 and 6) also provide empirical evidence for the "diagonal effect" identified by Biais, Hillion, and Spatt (1995). The diagonal effect describes the stylized fact that the probability of observing a market event (a market order submission, say), given that the most recent market event was of the same type, is higher than the unconditional probability. The statistically and economically significant effect of the lagged counts on the expected number of counts of the same order category is consistent with the presence of a diagonal effect.

[12]A detailed treatment of copulas can be found in Joe (1997) and Nelsen (1999).

[13] Machado and Santos Silva (2003) use this idea in order to work out the theoretical properties of a quantile estimator for discrete data.

Table 1: **Descriptive statistics for market event one-minute counts**

	Obs.	**Mean**	**Std. Dev.**	**Disp.**	**Max.**	**Q(60)**
BUY ORDERS	**52712**	**4.91**	**4.37**	**3.89**	**68**	**37817**
Category 1 (BMO-agg)	**3494**	**0.33**	**0.71**	**1.53**	**22**	**7888**
- True Market Orders	898	0.08	0.36	1.51	18	872
- Marketable Limit Orders	2596	0.24	0.56	1.28	6	7397
Category 2 (BMO-inter)	**3369**	**0.31**	**0.64**	**1.32**	**6**	**1629**
- True Market Orders	18	0.01	0.04	1.00	1	64
- Marketable Limit Orders	3351	0.31	0.64	1.33	6	1627
Category 3 (BMO-small)	**5250**	**0.49**	**0.81**	**1.33**	**7**	**11106**
- True Market Orders	2564	0.24	0.54	1.22	6	8990
- Marketable Limit Orders	2686	0.25	0.55	1.23	5	1344
Buy Market Orders (BMO)	**12113**	**1.13**	**1.46**	**1.89**	**29**	**22759**
Category 4 (BLO-inside)	18312	1.71	1.85	2.00	17	21309
Category 5 (BLO-at)	11411	1.06	1.33	1.68	18	14313
Category 6 (BLO-outside)	10876	1.01	1.28	1.62	11	8657
Buy Limit Orders (BLO)	**40599**	**3.78**	**3.35**	**2.96**	**39**	**33304**
Cancelations (BCANC)	**20534**	**1.91**	**2.03**	**2.15**	**18**	**13623**
SELL ORDERS	**43163**	**4.02**	**3.92**	**3.82**	**38**	**20498**
Category 1 (SMO-agg)	**2263**	**0.21**	**0.53**	**1.36**	**6**	**1442**
- True Market Orders	524	0.05	0.23	1.12	3	472
- Marketable Limit Orders	1739	0.16	0.45	1.25	5	1125
Category 2 (SMO-inter)	**3077**	**0.29**	**0.63**	**1.38**	**8**	**2602**
- True Market Orders	94	0.01	0.11	1.33	5	305
- Marketable Limit Orders	2983	0.28	0.62	1.36	7	2551
Category 3 (SMO-small)	**2241**	**0.21**	**0.52**	**1.32**	**10**	**833**
- True Market Orders	892	0.08	0.31	1.14	5	362
- Marketable Limit Orders	1349	0.13	0.40	1.24	8	426
Sell Market Orders (SMO)	**7581**	**0.71**	**1.15**	**1.86**	**15**	**5331**
Category 4 (SLO-inside)	15012	1.34	1.68	2.00	13	11184
Category 5 (SLO-at)	10166	0.95	1.30	1.78	23	8660
Category 6 (SLO-outside)	10404	0.97	1.25	1.62	11	6738
Sell Limit Orders (SLO)	**35582**	**3.32**	**3.14**	**2.97**	**38**	**21272**
Cancelations (SCANC)	**20010**	**1.86**	**2.09**	**2.34**	**29**	**11379**

This table presents the descriptive statistics of the one-minute Daimler-Chrysler count series of market events. Category 1 orders are market orders that walk up the book. Category 2 orders are market orders which consume all (but not more than) the volume available at the best quote. Category 3 orders are market orders that consume part of the depth at the best quote. Category 4 orders are aggressive limit orders, i.e. orders submitted inside the best quotes. Category 5 orders are limit orders submitted at the best quote. Category 6 orders are limit orders outside the best quotes, i.e. below (above) the bid (ask). Q(60) reports the Ljung-Box Q-statistic computed with 60 lagged autocorrelations. The Disp. column reports the ratio of sample variance to sample mean.

Table 2: **Descriptive statistics for disaggregated market event one-minute counts (least aggressive limit orders and cancelations)**

	Obs	Mean	Std. Dev.	Disp.	Max.	Q(60)
Category 6 buy orders (BLO-outside)	**10876**	**1.01**	**1.28**	**1.62**	**11**	**8657**
- BLO-outside-1-2	4322	0.40	0.75	1.41	7	2513
- BLO-outside-3-5	3702	0.35	0.66	1.28	7	3375
- BLO-outside-5+	2852	0.27	0.61	1.40	9	1929
Buy cancelations: (BCANC)	**20534**	**1.70**	**1.86**	**2.04**	**16**	**12715**
- BCANC-0-2	8518	0.79	1.12	1.58	9	4748.5
- BCANC-3-5	6306	0.59	0.90	1.39	8	4350
- BCANC-5+	5710	0.53	0.96	1.72	11	4948
Category 6 sell orders (SLO-outside)	**10404**	**0.97**	**1.25**	**1.62**	**11**	**6738**
- SLO-outside-1-2	4286	0.40	0.77	1.49	8	2642
- SLO-outside-3-5	3479	0.32	0.62	1.19	5	2023
- SLO-outside-5+	2639	0.25	0.58	1.36	7	1723
Sell cancelations: (SCANC)	**20010**	**1.66**	**1.94**	**2.27**	**29**	**9799**
- SCANC-0-2	8139	0.76	1.09	1.58	9	4714.6
- SCANC-3-5	6219	0.58	0.89	1.36	9	3689
- SCANC-5+	5652	0.53	0.11	2.33	26	4227

This table presents the descriptive statistics of the one-minute Daimler-Chrysler count series of market events. Buy and sell orders of category 6 (limit orders submitted outside the best quotes) have been disaggregated according to their relative position to the best quotes. BLO-outside-1-2 and SLO-outside-1-2 count the number of buy and sell orders submitted one or two quotes away from the best quotes. The categories BLO-outside-3-5, SLO-outside-3-5, BLO-outside-5+ and SLO-outside-5+ are defined accordingly. A disaggregation of the buy and sell cancelations (BCANC and SCANC) is conducted accordingly: BCANC-0-2 and SCANC-0-2 denote cancelations of standing limit orders at, or one or two quotes away from the best quotes. The categories BCANC-3-5, SCANC-3-5, BCANC-5+ and SCANC-5+ are defined accordingly. Q(60) reports the Ljung-Box Q-statistic computed with 60 autocorrelations

Table 3: **Contemporaneous dependence of market events**

	BMO	BLO	SMO	SLO	BCANC	SCANC
BMO	1.000					
BLO	0.100	1.000				
SMO	-0.025	0.187	1.000			
SLO	0.171	0.195	0.159	1.000		
BCANC	0.177	0.574	0.103	0.193	1.000	
SCANC	0.119	0.202	0.196	0.574	0.191	1.000

For the MDACP approach we use a Gaussian copula to account for contemporaneous dependence in the market event count sequences. This implies that the degree of contemporaneous dependence can be measured by computing the correlation matrix of the quantile vector $q_t = (\Phi^{-1}(z_{1,t}), \ldots, \Phi^{-1}(z_{K,t}))'$, where Φ^{-1} denotes the quantile function of the standard normal distribution, K the number of count series, and $z_{i,t}$ the sequence of probability integral transforms of the ith continuous extension of the count data series (see appendix for details). The table reports this estimated correlation matrix for an aggregated MDACP system which uses buy and sell market orders (BMO and SMO), limit orders (BLO and SLO) as well as cancelation counts (BCANC and SCANC) as dependent variables.

Table 4: **Estimation results for an aggregated MDACP system**

	BMO	BLO	SMO	SLO	BCANC	SCANC
ω	**0.032**	**0.270**	**0.052**	**0.348**	**0.115**	**0.159**
α_{BMO}	**0.086**	**0.051**	0.002	**0.168**	-0.019	-0.017
α_{BLO}	**0.023**	**0.178**	0.004	**0.041**	**0.102**	0.012
α_{SMO}	0.011	**0.266**	**0.087**	**0.048**	-0.001	0.015
α_{SLO}	-0.006	-0.008	**0.034**	**0.180**	**-0.015**	**0.121**
α_{BCANC}	**-0.016**	0.023	-0.001	**0.027**	**0.104**	0.009
α_{SCANC}	0.004	**0.064**	**-0.011**	**0.054**	**0.044**	**0.086**
β	**0.848**	**0.647**	**0.709**	**0.541**	**0.660**	**0.600**
γ_{SPREAD}	**-1.019**	-0.134	**-1.736**	**-0.341**	-0.300	-0.391
γ_{BIDVOL}	6.32E-6	3.96E-6	**1.015E-5**	**5.06E-6**	6.48E-7	6.57E-6
γ_{ASKVOL}	**7.26E-6**	2.95E-6	**1.65E-5**	**6.01E-6**	4.10E-6	4.12E-6
γ_{VOLAT}	**-0.643**	0.160	-0.492	**0.380**	**1.475**	**0.573**
ϕ	**0.713**	**0.525**	**0.752**	**0.511**	**0.614**	**0.604**
$W(\psi's = 0)$	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Var(ϵ_t)	1.02	1.00	1.05	1.01	1.00	1.03
Log likelihood	-14661.9	-23535.2	-12049.8	-22717.2	-18410.6	-18254.2

The table reports Maximum Likelihood estimates of a MDACP system. The dependent variables are the one-minute counts of buy market orders (BMO), buy limit orders (BLO), sell market orders (SMO), sell limit orders (SLO), buy cancelations (BCANC) and sell cancelations (SCANC). The mean equations are specified as

$$\mu^*_{t,i} = \mu_{t,i} \exp\left(X_{t-1}\gamma_i + \textstyle\sum_{p=1,2} \psi_{c,p} \cos \frac{2\pi p\, Re[t,N]}{N} + \psi_{s,p} \sin \frac{2\pi p\, Re[t,N]}{N}\right),$$

where $\mu_{t,i} = \omega_i + \sum_{j=1}^{6} \alpha_{i,j} N_{t-1,j} + \beta_i \mu_{t-1,i}$, for $t = 1, \ldots, 10731$. $Re[t, N]$ denotes the remainder of the integer division of t by N, the number of periods in a trading session. X_{t-1} collects the vector of predetermined variables, the inside spread (SPREAD), the volume at the best bid (BIDVOL), the volume at the best ask (ASKVOL) and volatility measured by the standard deviation of the last 5 minutes midquote returns (VOLAT). ϕ is the dispersion parameter of the Double Poisson. Parameters significant at the 5% level are printed boldface. For ϕ the null hypothesis is that the parameter is one, for β and the α parameters the null is that the true parameter is zero. The $W(\psi's = 0)$ row reports the p-values for a Wald-test of the hypothesis that the seasonality parameters $\psi_{s,1}$, $\psi_{s,2}$, $\psi_{c,1}$ and $\psi_{c,2}$ are jointly zero. $Var(\varepsilon_t)$ is the variance of the Pearson residual which should be close to one for a correctly specified model.

Table 5: **Estimation results for a disaggregated MDACP system - Bid side**

	BMO-agg	BMO-inter	BMO-small	BLO-inside	BLO-at	BLO-outside
ω	**0.008**	**0.050**	**0.007**	**0.150**	**0.045**	**0.060**
$\alpha_{\text{BMO-agg}}$	**0.056**	0.009	**0.011**	**0.062**	0.019	-0.004
$\alpha_{\text{BMO-inter}}$	0.008	**0.047**	-3.13E-4	-0.011	0.001	-0.003
$\alpha_{\text{BMO-small}}$	**0.022**	0.011	**0.035**	**0.066**	0.011	0.001
$\alpha_{\text{BLO-inside}}$	**0.009**	**0.038**	0.003	**0.110**	**0.018**	0.011
$\alpha_{\text{BLO-at}}$	**0.011**	**0.011**	**0.006**	**0.052**	**0.091**	**0.031**
$\alpha_{\text{BLO-outside}}$	-0.001	0.001	-0.001	**0.024**	**0.058**	**0.138**
$\alpha_{\text{SMO-agg}}$	0.001	-0.014	**-0.011**	**0.092**	**0.158**	**0.110**
$\alpha_{\text{SMO-inter}}$	**0.017**	0.002	-0.004	**0.104**	**0.070**	**0.051**
$\alpha_{\text{SMO-small}}$	0.001	0.015	-0.003	0.020	**0.034**	0.010
$\alpha_{\text{SLO-inside}}$	-0.003	0.004	0.004	**0.027**	-0.001	0.011
$\alpha_{\text{SLO-at}}$	-0.001	-0.003	0.001	-0.015	**0.017**	0.012
$\alpha_{\text{SLO-outside}}$	-0.003	0.007	-0.002	-0.003	-0.001	**0.015**
β	**0.807**	**0.579**	**0.921**	**0.656**	**0.721**	**0.689**
γ_{BFACT1}	2.24E-4	**0.010**	0.004	**0.005**	**0.010**	**-0.009**
γ_{BFACT3}	**0.039**	0.011	0.001	0.018	0.013	-0.014
γ_{SFACT1}	**0.018**	0.004	**0.010**	**0.005**	**0.005**	-0.003
γ_{SFACT3}	**-0.008**	-0.004	-0.006	-0.012	0.003	0.005
$\gamma_{\text{DIFFSLOPE}}$	0.009	**-0.023**	0.002	**-0.016**	-0.008	**0.014**
γ_{SPREAD}	0.150	**-2.549**	**-0.809**	0.086	-0.500	-0.404
γ_{BIDVOL}	9.80E-6	**2.20E-5**	**-1.70E-5**	**1.46E-5**	**-1.82E-5**	2.70E-6
γ_{ASKVOL}	-3.33E-6	**-3.21E-5**	**2.87E-5**	2.65E-6	2.66E-6	-2.38E-7
γ_{VOLAT}	**-0.622**	-0.591	-0.263	**-0.153**	**0.174**	0.257
ϕ	**1.177**	**1.146**	**1.004**	**0.652**	**0.790**	**0.762**
$W(\psi's = 0)$	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Var(ϵ_t)	1.06	1.04	1.02	0.98	1.00	1.01
Log likelihood	-7318.7	-7425.4	-9486.6	-17499.5	-14001.5	-13891.0

The table reports Maximum Likelihood estimates of a MDACP system. The dependent variables are the one-minute counts of category 1 orders (BMO-agg/SMO-agg, market orders that walk up or down the book), category 2 orders (BMO-inter/SMO-inter, market orders which consume all volume available at the best quote), category 3 orders (BMO-small/SMO-small, market orders which consume part of the depth at the best quote), category 4 orders (BLO-inside/SLO-inside, limit orders submitted inside the best quotes), category 5 orders (BLO-at/SLO-at, limit orders submitted at the best quote), category 6 orders (BLO-outside/SMO-outside, limit orders outside the best quotes) and cancelations (BCANC/SCANC). The table reports the results of the bid side equations. The mean equations are specified as:

$$\mu^*_{t,i} = \mu_{t,i} \exp\left(X_{t-1}\gamma_i + \textstyle\sum_{p=1,2}(\psi_{c,p} \cos \frac{2\pi p\, Re[t,N]}{N} + \psi_{s,p} \sin \frac{2\pi p\, Re[t,N]}{N})\right),$$

where $\mu_{t,i} = \omega_i + \sum_{j=1}^{12} \alpha_{i,j} N_{t-1,j} + \beta_i \mu_{t-1,i}$, for $t = 1, \ldots, 10731$. $Re[t,N]$ is the remainder of the integer division of t by N, the number of periods in a trading session. X_{t-1} collects the vector of predetermined variables, the inside spread (SPREAD), the volume at the best bid (BIDVOL), the volume at the best ask (ASKVOL) and volatility measured by the standard deviation of the last 5 minutes midquote returns (VOLAT). BFACT1 (SFACT1) denotes the first factor (liquidity factor) extracted by PCA at the bid (ask) side, DIFFSLOPE is the difference of the absolute values of the second factors (informational factor). BFACT3 and SFACT3 denote the third factor for the bid (ask) side. ϕ is the dispersion parameter of the Double Poisson. Parameters significant at the 5% level are printed boldface. For ϕ the null hypothesis is that the parameter is one, for β and the α parameters the null is that the true parameter is zero.The $W(\psi's = 0)$ row reports the p-values for a Wald-test of the hypothesis that the seasonality parameters $\psi_{s,1}$, $\psi_{s,2}$, $\psi_{c,1}$ and $\psi_{c,2}$ are jointly zero. $Var(\varepsilon_t)$ is the variance of the Pearson residual which should be close to one for a correctly specified model

Table 6: **Estimation results for a disaggregated MDACP system - Ask side**

	SMO-agg	SMO-inter	SMO-small	SLO-inside	SLO-at	SLO-outside
ω	**0.017**	**0.027**	**0.024**	**0.151**	**0.057**	**0.121**
$\alpha_{\text{BMO-agg}}$	**-0.014**	-0.008	-0.002	0.010	**0.069**	**0.046**
$\alpha_{\text{BMO-inter}}$	**-0.010**	**-0.016**	**-0.011**	**0.051**	**0.082**	**0.067**
$\alpha_{\text{BMO-small}}$	**0.009**	0.008	0.004	0.024	**0.020**	0.017
$\alpha_{\text{BLO-inside}}$	0.003	**0.011**	0.001	**0.049**	**0.016**	**0.026**
$\alpha_{\text{BLO-at}}$	**0.005**	**0.009**	**-0.004**	0.009	0.011	**0.022**
$\alpha_{\text{BLO-outside}}$	-0.002	-0.004	0.003	-0.001	0.002	**0.022**
$\alpha_{\text{SMO-agg}}$	**0.071**	0.004	**0.016**	**0.094**	-0.005	**0.060**
$\alpha_{\text{SMO-inter}}$	**0.022**	**0.035**	0.008	0.014	-0.003	-0.007
$\alpha_{\text{SMO-small}}$	0.009	-0.002	**0.032**	-0.005	0.022	0.018
$\alpha_{\text{SLO-inside}}$	**0.013**	**0.036**	**0.012**	**0.132**	0.015	**0.025**
$\alpha_{\text{SLO-at}}$	**0.009**	**0.007**	**0.007**	**0.053**	**0.103**	**0.052**
$\alpha_{\text{SLO-outside}}$	0.002	**0.006**	**0.006**	**0.040**	**0.037**	**0.123**
β	**0.682**	**0.632**	**0.677**	**0.559**	**0.676**	**0.512**
γ_{BFACT1}	**0.008**	-0.003	**0.035**	4.08E-4	0.001	-0.002
$\gamma_{\text{DIFFSLOPE}}$	0.006	**0.037**	2.05E-4	**0.022**	0.006	**-0.019**
γ_{BFACT3}	0.004	-0.023	**0.072**	-0.013	0.010	**-0.025**
γ_{SFACT1}	**0.009**	**0.014**	-1.93E-4	**0.007**	**0.011**	**-0.014**
γ_{SFACT3}	-0.026	0.001	**0.049**	0.005	-0.016	0.007
γ_{SPREAD}	**-1.349**	**-2.467**	**-1.902**	-0.456	-0.274	-0.378
γ_{BIDVOL}	**-3.37E-5**	**-1.865E-5**	**3.86E-5**	**5.82E-6**	5.24E-6	1.68E-6
γ_{ASKVOL}	**1.39E-5**	**1.07E-5**	**1.31E-5**	**1.67E-5**	**-9.53E-6**	-1.96E-6
γ_{VOLAT}	**-0.493**	**-0.334**	-0.142	**-0.120**	**0.431**	0.487
ϕ	**1.388**	**1.195**	**1.402**	**0.652**	**0.773**	**0.775**
$W(\psi's = 0)$	(0.00)	(0.00)	(0.01)	(0.01)	(0.71)	(0.00)
Var(ϵ_t)	1.07	1.04	1.09	1.01	1.03	1.02
Log likelihood	-5516.6	-6952.0	-5497.3	-16220.2	-13507.1	-13613.9

The table reports Maximum Likelihood estimates of a MDACP system. The dependent variables are the one-minute counts of category 1 orders (BMO-agg/SMO-agg, market orders that walk up or down the book), category 2 orders (BMO-inter/SMO-inter, market orders which consume all volume available at the best quote), category 3 orders (BMO-small/SMO-small, orders are market orders that consume part of the depth at the best quote), category 4 orders (BLO-inside/SLO-inside, limit orders submitted inside the best quotes), category 5 orders (BLO-at/SLO-at, limit orders submitted at the best quote), category 6 orders (BLO-outside/SMO-outside, limit orders outside the best quotes) and cancelations (BCANC/SCANC). The table reports the results of the ask side equations. The mean equations are specified as:

$$\mu^*_{t,i} = \mu_{t,i} \exp\left(X_{t-1}\gamma_i + \textstyle\sum_{p=1,2}(\psi_{c,p} \cos \frac{2\pi p\, Re[t,N]}{N} + \psi_{s,p} \sin \frac{2\pi p\, Re[t,N]}{N})\right),$$

where $\mu_{t,i} = \omega_i + \sum_{j=1}^{12} \alpha_{i,j} N_{t-1,j} + \beta_i \mu_{t-1,i}$, for $t = 1, \ldots, 10731$. $Re[t, N]$ is the remainder of the integer division of t by N, the number of periods in a trading session. X_{t-1} collects the vector of predetermined variables, the inside spread (SPREAD), the volume at the best bid (BIDVOL), the volume at the best ask (ASKVOL) and volatility measured by the standard deviation of the last 5 minutes midquote returns (VOLAT). BFACT1 (SFACT1) denotes the first factor (liquidity factor) extracted by PCA at the bid (ask) side, DIFFSLOPE is the difference of the absolute values of the second factors (informational factor). BFACT3 and SFACT3 denote the third factor for the bid (ask) side. ϕ is the dispersion parameter of the Double Poisson. Parameters significant at the 5% level are printed boldface. For ϕ the null hypothesis is that the parameter is one, for β and the α parameters the null is that the true parameter is zero.The $W(\psi's = 0)$ row reports the p-values for a Wald-test of the hypothesis that the seasonality parameters $\psi_{s,1}$, $\psi_{s,2}$, $\psi_{c,1}$ and $\psi_{c,2}$ are jointly zero. $Var(\varepsilon_t)$ is the variance of the Pearson residual which should be close to one for a correctly specified model.

Table 7: **Estimation results for a bivariate MDACP system of buy and sell market orders with cancelation counts as predetermined variables**

	BMO	SMO
ω	**0.034**	**0.057**
α_{BMO}	**0.086**	0.002
α_{SMO}	**0.020**	0.003
β	**0.848**	**0.698**
$\gamma_{\text{BCANC-0-2}}$	**-0.020**	-0.007
$\gamma_{\text{BCANC-3-5}}$	-0.013	-0.007
$\gamma_{\text{BCANC-5+}}$	-0.009	0.010
$\gamma_{\text{SCANC-0-2}}$	**-0.016**	**-0.031**
$\gamma_{\text{SCANC-3-5}}$	0.014	-0.006
$\gamma_{\text{SCANC-5+}}$	**0.012**	0.004
ϕ	**0.714**	**0.753**
$W(\psi's = 0)$	(0.00)	(0.00)
Var(ϵ_t)	1.02	0.99
Log likelihood	-14656.7	-12042.91

The table reports Maximum Likelihood estimates of a bivariate MDACP system. The dependent variables are one minute counts of buy (BMO) and sell (SMO) market orders. The mean equations are specified as:

$$\mu^*_{t,i} = \mu_{t,i} \exp\left(X_{t-1}\gamma_i + \sum_{p=1,2} \psi_{c,p} \cos \frac{2\pi p\, Re[t,N]}{N} + \psi_{s,p} \sin \frac{2\pi p\, Re[t,N]}{N}\right)$$

where $\mu_{t,i} = \omega_i + \sum_{j=1}^{2} \alpha_{i,j} N_{t-1,j} + \beta_i \mu_{t-1,i}$, for $t = 1, \ldots, 10731$. $Re[t,N]$ is the remainder of the integer division of t by N, the number of periods in a trading session. X_{t-1} collects the vector of predetermined variables which consist of cancelations categorized according to their position away from the best quotes. BCANC-0-2 and SCANC-0-2 denote cancelations of standing limit orders at, or one or two quotes away from the best quotes. The categories BCANC-3-5, SCANC-3-5, BCANC-5+ and SCANC-5+ are defined accordingly. ϕ is the dispersion parameter of the Double Poisson. Parameters significant at the 5% level are printed boldface. For ϕ the null hypothesis is that the parameter is one, for β and the α parameters the null is that the true parameter is zero. The $W(\psi's = 0)$ row reports the p-values for a Wald-test of the hypothesis that the seasonality parameters $\psi_{s,1}$, $\psi_{s,2}$, $\psi_{c,1}$ and $\psi_{c,2}$ are jointly zero. $Var(\varepsilon_t)$ is the variance of the Pearson residual which should be close to one for a correctly specified model.

Table 8: **Estimation results for a MDACP system of buy and sell limit order categories with cancelation counts as predetermined variables**

Parameters	BLO-inside	BLO-at	BLO-outside	SLO-inside	SLO-at	SLO-outside
ω	**0.133**	**0.041**	**0.057**	**0.137**	**0.048**	**0.109**
$\alpha_{\text{BLO-inside}}$	**0.143**	**0.035**	-0.007	**0.065**	**0.045**	**0.044**
$\alpha_{\text{BLO-at}}$	**0.069**	**0.109**	0.005	0.003	-0.002	**0.015**
$\alpha_{\text{BLO-outside}}$	**0.040**	**0.077**	**0.091**	-0.009	**-0.014**	0.012
$\alpha_{\text{SLO-inside}}$	**0.058**	**0.041**	**0.043**	**0.155**	**0.024**	0.004
$\alpha_{\text{SLO-at}}$	**-0.023**	-4.91E-4	-0.002	**0.064**	**0.115**	**0.019**
$\alpha_{\text{SLO-outside}}$	**-0.036**	-0.010	0.010	**0.046**	**0.052**	**0.087**
β	**0.680**	**0.713**	**0.685**	**0.581**	**0.707**	**0.533**
$\gamma_{\text{BCANC-0-2}}$	**-0.037**	**-0.036**	**0.053**	0.012	0.005	0.015
$\gamma_{\text{BCANC-3-5}}$	**-0.021**	**-0.037**	**0.093**	0.009	**0.027**	**0.024**
$\gamma_{\text{BCANC-5+}}$	**-0.027**	0.006	**0.025**	0.011	**0.017**	-0.007
$\gamma_{\text{SCANC-0-2}}$	-1.49E-5	-0.006	-0.005	-0.014	**-0.039**	**0.089**
$\gamma_{\text{SCANC-3-5}}$	**0.061**	0.011	0.016	-0.028	**-0.044**	**0.103**
$\gamma_{\text{SCANC-5+}}$	**0.029**	**0.049**	**0.017**	0.007	**0.016**	0.013
ϕ	**0.649**	**0.786**	**0.765**	**0.649**	**0.771**	**0.777**
$W(\psi's = 0)$	(0.00)	(0.00)	(0.01)	(0.01)	(0.94)	(0.00)
Var(ϵ_t)	0.98	1.00	1.02	1.01	1.03	1.03
Log likelihood	-17525.1	-14027.4	-13866.3	-16242.7	-13519.5	-13601.6

The table reports the Maximum Likelihood estimates of a MDACP system. The dependent variables are the one-minute counts of category 4 orders (BLO-at and SLO-at, limit orders submitted inside the best quotes), category 5 orders (BLO-inside and SLO-inside, limit orders submitted at the best quote) and category 6 orders (BLO-outside and SLO-outside, limit orders submitted outside the best quotes). The mean equations are specified as:

$$\mu^*_{t,i} = \mu_{t,i} \exp\left(X_{t-1}\gamma_i + \sum_{p=1,2} \psi_{c,p} \cos \frac{2\pi p\, Re[t,N]}{N} + \psi_{s,p} \sin \frac{2\pi p\, Re[t,N]}{N}\right)$$

where $\mu_{t,i} = \omega_i + \sum_{j=1}^{6} \alpha_{i,j} N_{t-1,j} + \beta_i \mu_{t-1,i}$, for $t = 1, \ldots, 10731$. $Re[t,N]$ is the remainder of the integer division of t by N, the number of periods in a trading session. X_{t-1} collects the vector of predetermined variables which consist of cancelations categorized according to their position away from the best quotes. BCANC-0-2 and SCANC-0-2 denote cancelations of standing limit orders at one or two quotes away from the best quotes. The categories BCANC-3-5, SCANC-3-5, BCANC-5+ and SCANC-5+ are defined accordingly. ϕ is the dispersion parameter of the Double Poisson. Parameters significant at the 5% level are printed boldface. For ϕ the null hypothesis is that the parameter is one, for β and the α parameters the null is that the true parameter is zero. The $W(\psi's = 0)$ row reports the p-values for a Wald-test of the hypothesis that the seasonality parameters $\psi_{s,1}$, $\psi_{s,2}$, $\psi_{c,1}$ and $\psi_{c,2}$ are jointly zero. $Var(\varepsilon_t)$ is the variance of the Pearson residual which should be close to one for a correctly specified model.

Table 9: **Estimation results for a MDACP system of buy and sell limit order categories (submitted outside the best quotes) with cancelation counts as predetermined variables**

	BLO-out-1-2	BLO-out-3-5	BLO-out-5+	SLO-out-1-2	SLO-out-3-5	SLO-out-5+
ω	**0.019**	**0.014**	**0.025**	**0.038**	**0.021**	**0.032**
$\alpha_{\text{BLO-outside-1-2}}$	**0.058**	0.005	**0.011**	-0.004	**0.012**	**0.013**
$\alpha_{\text{BLO-outside-3-5}}$	**0.022**	**0.074**	**0.025**	0.011	0.001	**-0.012**
$\alpha_{\text{BLO-outside-5+}}$	0.001	**0.015**	**0.084**	**0.012**	-0.001	-0.002
$\alpha_{\text{SLO-outside-1-2}}$	0.006	**0.011**	0.005	**0.079**	**0.011**	**-0.008**
$\alpha_{\text{SLO-outside-3-5}}$	0.006	0.001	0.006	**0.021**	**0.046**	**0.013**
$\alpha_{\text{SLO-outside-5+}}$	0.003	**0.011**	**-0.013**	**-0.014**	-0.001	**0.087**
β	**0.684**	**0.716**	**0.768**	**0.558**	**0.692**	**0.663**
$\gamma_{\text{BCANC-0-2}}$	**0.046**	**0.023**	**-0.009**	**0.025**	**0.012**	**0.008**
$\gamma_{\text{BCANC-3-5}}$	**0.029**	**0.060**	**-0.009**	**0.012**	**0.014**	**0.009**
$\gamma_{\text{BCANC-5+}}$	**0.007**	-0.005	**0.027**	0.002	-0.001	**0.0097**
$\gamma_{\text{SCANC-0-2}}$	**0.011**	0.005	-0.004	**0.084**	**0.016**	**-0.007**
$\gamma_{\text{SCANC-3-5}}$	**0.014**	0.006	**0.008**	**0.021**	**0.063**	**0.007**
$\gamma_{\text{SCANC-5+}}$	**0.020**	3.36E-4	**0.012**	0.003	-0.003	**0.023**
ϕ	**1.030**	**1.149**	**1.347**	**1.0223**	**1.197**	**1.287**
$W(\psi' s = 0)$	(0.00)	(0.02)	(0.01)	(0.01)	(0.00)	(0.00)
Var(ϵ_t)	1.01	1.04	1.06	1.02	1.02	1.04
Log likelihood	-8615.7	-7683	-6701.9	-8560.2	-7364.3	-6238.8

The table reports the Maximum Likelihood estimates of a MDACP model. The dependent variables are the counts of limit orders submitted outside the best quotes. BLO-outside-1-2 and SLO-outside-1-2 count the number of buy and sell orders submitted one or two quotes away from the best quotes. The categories BLO-outside-3-5, SLO-outside-3-5, BLO-outside-5+ and SLO-outside-5+ are defined accordingly (note that we use *out* instead of *outside* in the upper part of the table, due to space limitations). The mean equations are specified as:

$$\mu^*_{t,i} = \mu_{t,i} \exp\left(X_{t-1}\gamma_i + \sum_{p=1,2} \psi_{c,p} \cos \frac{2\pi p\, Re[t,N]}{N} + \psi_{s,p} \sin \frac{2\pi p\, Re[t,N]}{N}\right)$$

where $\mu_{t,i} = \omega_i + \sum_{j=1}^{6} \alpha_{i,j} N_{t-1,j} + \beta_i \mu_{t-1,i}$, for $t = 1, \ldots, 10731$. $Re[t,N]$ is the remainder of the integer division of t by N, the number of periods in a trading session. X_{t-1} collects the vector of predetermined variables which consist of cancelations categorized according to their position away from the best quotes. BCANC-0-2 and SCANC-0-2 denote cancelations of standing limit orders at, or one or two quotes away from the best quotes. The categories BCANC-3-5, SCANC-3-5, BCANC-5+ and SCANC-5+ are defined accordingly. ϕ is the dispersion parameter of the Double Poisson. Parameters significant at the 5% level are printed boldface. For ϕ the null hypothesis is that the parameter is one, for β and the α parameters the null is that the true parameter is zero. The $W(\psi' s = 0)$ row reports the p-values for a Wald-test of the hypothesis that the seasonality parameters $\psi_{s,1}$, $\psi_{s,2}$, $\psi_{c,1}$ and $\psi_{c,2}$ are jointly zero. $Var(\varepsilon_t)$ is the variance of the Pearson residual which should be close to one for a correctly specified model.

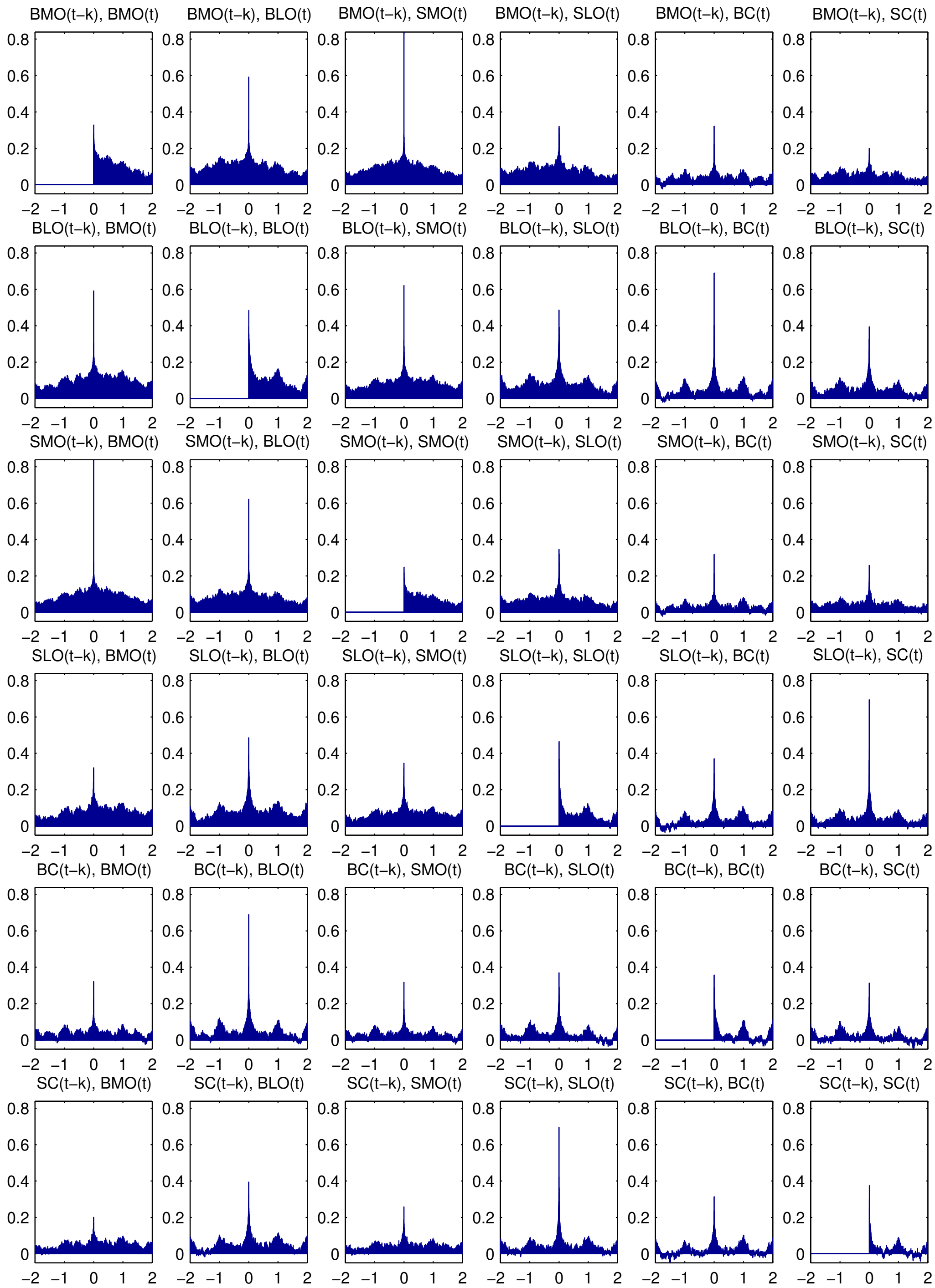

Figure 1: **Cross-correlation of market events**

The figure depicts two days auto- and cross-correlograms of the aggregated market event counts for Daimler-Chrysler. BMO denotes buy market orders, SMO sell market orders, BLO buy limit orders, SLO sell limit orders, BC buy cancelations, and SC sell cancelations.

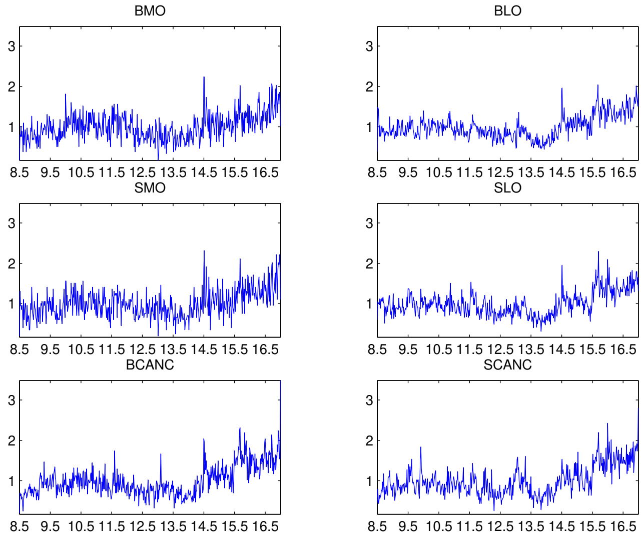

Figure 2: **Seasonality in market event count series**

The figure depicts the daily seasonality of the aggregated market event counts for Daimler-Chrysler. Xetra trading hours at the FSE extended from 8.30 a.m to 5.00 p.m. CET. BMO denote buy market orders, SMO market orders, BLO buy limit orders, SLO sell limit orders, BCANC buy cancelations and SCANC sell cancelations.

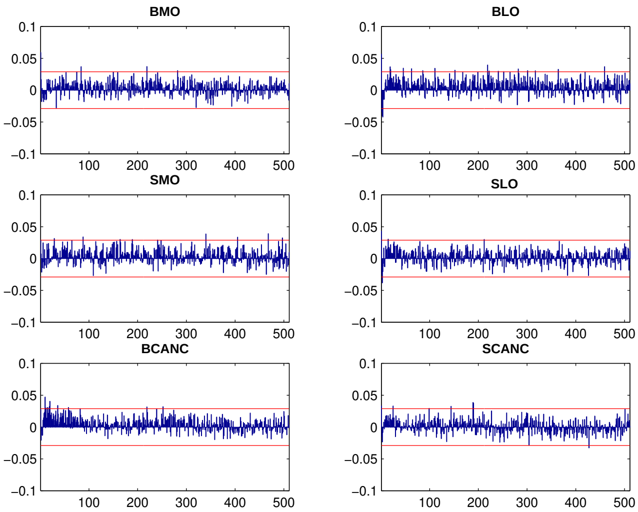

Figure 3: **Autocorrelogram of the Pearson residuals - Aggregated MDACP system**

The figures depict one-day (510 one-minute intervals) autocorrelograms of the Pearson residuals of the aggregated MDACP system (estimation results in table 4). BMO denote buy market orders, SMO sell market orders, BLO buy limit orders, SLO sell limit orders, BCANC buy cancelations, and SCANC sell cancelations.