Econometric Modeling of Inter-Order Durations

K. Bień-Barkowska

Institute of Econometrics, Warsaw School of Economics, Madalińskiego 6/8, 02-513 Warsaw, Poland

We investigate the dynamics of inter-order durations, i.e. times elapsing between consecutive orders submitted to the Reuters Dealing 3000 Spot Matching System, an automated brokerage platform for interbank EUR/PLN spot trading. Strong autocorrelation of the inter-order waiting times combined with the significant cross-correlations among individual order types (i.e. market buy, market sell, limit buy, limit sell) has been captured with the Multistate Asymmetric Box-Cox Autoregressive Conditional Duration (MABCACD) model. Our empirical study provides new insights about the microstructure of the interbank FX spot markets.

DOI: 10.12693/APhysPolA.127.A-7
PACS: 89.65 Gh

1. Introduction

Empirical studies of the market microstructure have undergone a rapid upsurge over the last two decades, mainly due to the access to intraday financial databases that provide academics with the extremely detailed information on exact times and characteristics (i.e. marks) of individual orders and trades. Usually, the tick-by-tick data refers to stock markets and is saved on-the-fly by the automated brokerage systems at major trading venues, which provides a standpoint for in-depth analyses of price or liquidity formation on equity markets [1–5].

Depending on an economic question at hand, the intraday tick-by-tick data can be thinned with respect to selected ‘micro-scale’ events, such as order arrivals, order cancellations, price changes or buy or sell trades of a pre-defined volume. In the literature on high-frequency finance, such sequences of events are often described as dynamic intensity-based models of marked point processes in continuous time [6–9]. Second strand of the literature covers models of financial durations (i.e. times between selected events) in discrete time, such as Autoregressive Conditional Duration (ACD) models, originally proposed by R.F. Engle and J.R. Russell in Ref. [10]. In shortcut, the ACD model is an econometric time series model of the time intervals between arbitrarily defined events. It seeks to describe the dynamic properties of the time variable, and thus it is able to depict observable accelerations in trading intensity, volatility or temporary swings in the pace of liquidity provision. The survey of several extensions of the ACD models can be found in Ref. [3, 11, 12]. Empirical investigations of financial durations can be used to augment the intraday measures of market risk. First instantaneous volatility measures were derived by R.F. Engle and J.R. Russell in Ref. [10]. R.F. Engle as well as G. Dionne et al. propose the so-called Ultra-High-Frequency Generalised Autoregressive Conditional Heteroskedasticity (GARCH) models which describe volatility per unit of time [13, 14]. J. Grammig and M. Wellner extend this approach by allowing for interdependence between durations and volatility in Ref. [15]. L. Bauwens and P. Giot propose the Asymmetric ACD models to construct a simple trading strategy [16]. R.F. Engle and J. Lange develop the VNET liquidity measure which reflects volume of buy or sell trades over a price duration (i.e. time span till price changes by a given increment) [17]. N. Hautsch uses time spans between net volume changes to infer about the intra-day process of liquidity formation [18]. J. Large investigates the resiliency of the limit order book to different market shocks with the application of order submission intensities [19].

In our study we rely on the duration-based representation of the order submission process in the Reuters Dealing 3000 Spot Matching System, the major order-driven platform for the interbank trading of the EUR/PLN. The aim of this study is threefold. First, we extend the Multistate Asymmetric Autoregressive Conditional Duration (MAACD) model of [20] using the Box-Cox transformation of the time variable, so that it is able to capture possible nonlinear relations between the current and lagged durations. The model can capture the strong autocorrelation of inter-order durations and the causal effects of the trading actions undertaken by individual currency dealers. Second, we present how to derive the time-varying conditional probabilities of observing a particular order type given past orders and durations. Finally, we show how to derive residuals of the model and prove that the MABCACD model is quite successful in describing the dynamics of the trading process in the interbank EUR/PLN spot market.

2. Empirical data

The Reuters Dealing 3000 Spot Matching System is the automated trading platform for the interbank FX spot market. In this automatic brokerage system, all incoming buy and sell orders are being continuously matched once their prices agree. Currency dealers who want to execute a buy/sell transaction can either submit a market order which is always immediately executed at the...
most competitive bid/ask quotes or post a limit order that waits for an execution in future at a more favorable price\(^1\). The EUR/PLN exchange rate is quoted as a quantity of Zlotys per one Euro and the transaction currency is Euro.

Our data is comprised of exact times (rounded to the nearest centisecond) and characteristics (quoted prices, sizes and the buy/sell indicators) of all market and limit orders submitted to the system during January and February 2008\(^1\). In the empirical study, we focus on the submissions of most frequent orders which can be classified to four categories \((k = 1, 2, 3, 4)\): market sell orders or marketable limit sell orders \((MS, k = 1)\); inside-the-quote limit sell order that improve the best ask quote \((IQS, k = 2)\); market buy orders or a marketable limit buy orders \((MB, k = 3)\); inside-the-quote limit buy orders that improve the best bid quote \((IQB, k = 4)\). All in all there are 15,839 MS orders, 21,998 IQS orders, 16,205 MB orders and 20,703 IQB orders in the sample. Inter-order durations are calculated as \(x_i = t_i - t_{i-1}\) (where \(t_i\) denotes a time of the \(i\)-th order for \(i = 1, 2, ..., n\))\(^3\). The process of submitting orders, and hence the process of trading itself, features two striking regularities: (1) orders cluster in time and (2) given order types tend to be followed by some order types and not by the others.

A clustering scheme for order arrivals is visible at first sight in Fig. 1. This feature might be attributed to the phenomenon of information clustering that is well-recognised in finance. Times of news releases are not independent in time but gravitate to each other. The pieces of private information that are dispersed among market participants are being continuously impounded in quoted prices, which results in systematic fluctuation of volatility. This stylised fact of high-frequency data gave birth to the seminal GARCH models that rely on news clustering, and hence volatility clustering schemes (see Ref. [22] for the original ‘plain-vanilla’ ARCH specification and Ref. [23] for its generalisation). The GARCH models, which are nowadays perceived as the cornerstone of financial econometrics, focus on autocorrelations in returns squared (but data must be converted to equidistant time series). However, their ‘siblings’, the ACD models, depict the dynamic properties of financial durations, and hence clustering of orders, trades or price changes. Economic background for the duration-based modelling strategy dates back to the theoretical models of market microstructure, where time between trades is assumed to be of informational value for market participants. D. Easley and M. O’Hara develop a model where temporal accelerations in trading intensity might signal the presence of informed traders, who buy an asset if the news is good, sell an asset if the news is bad and if there is no news, they do not trade [24]. Such buying/selling strategies might encourage other technical traders, which leads to clustering of trades in periods of information arrival. Therefore, short inter-order waiting times might indicate news arrival (see also related models [25–27]). Quite different explanation for duration clustering can be derived from Ref. [28], where liquidity traders minimize their transaction costs and enter the market if there are many other market participants and the bid-ask spread is narrow. Moreover, informed traders often tend to split their large trades into a sequence of smaller buy or sell transactions in order to camouflage the information they possess [29]. This means that order-splitting strategies can be explained by so called stealth-trading behaviour. As suggested in Ref. [30], most of price changes can be attributed to trades of medium sizes, which means that informed traders split their large orders into a sequence of medium-sized ones in order to limit the information leakage. Stealth trading behaviour has also been empirically documented in Ref. [31–33].

\(^1\)Market orders are usually treated as more ‘aggressive’ than limit orders since they consume liquidity at the best ask quote (buy market orders) or at the best bid quote (sell limit orders) by being immediately executed against best (most competitive) limit orders in the limit order book (LOB). Limit orders usually provide liquidity, because, if they cannot be immediately matched, they are inserted on the appropriate level of the LOB and are to be executed only if the price moves in their direction.

\(^1\)EUR/PLN trading is particularly active on working days between 8:00 and 18:00 CET. The quantity of orders submitted beyond these time frames is negligible and we excluded them from the sample.

\(^3\)The inter-order durations were adjusted for the time-of-day effect as suggested in Ref. [21].

---

![Fig. 1. Illustration of the order submission process — different symbols correspond to selected types of orders.](image-url)
factors that constitute private information can be found in Ref. [36]. FX dealers might form expectations about the directions of the FX rate with the self-assessment of market fundamentals or the technical analysis (chartists approach) [37]. Recently, it has been shown in Ref. [38], that chartists might even outperform fundamentalists in FX markets at short horizons.

The viewpoint of the theoretical models of the market microstructure analyses in Ref. [42–44]. Under the competing risks framework of the MAACD model, the (conditional) bivariate density for the inter-order durations (\(x_i\)) and the indicators of the order types (\(y_i = k\) where \(k = 1, 2, 3, 4\)) is given as:

\[
f(x_i, y_i | \mathcal{F}_{i-1}) = \prod_{k=1}^{4} h_{x_k}(x_i | \mathcal{F}_{i-1}) S_{x_k}(x_i | \mathcal{F}_{i-1}),
\]

where \(\mathcal{F}_{i-1}\) denotes an information set up to a time point \(t - 1\) that contains past realizations of \(x_i\) and \(y_i\), \(x_{i,k}\) (for \(k = 1, 2, 3, 4\)) corresponds to an inter-order duration at the end of which state \(k\) would be observed. Because at time \(t_i\) only one type of order can be observed, \(x_i = \min(x_{i,1}, x_{i,2}, x_{i,3}, x_{i,4})\). \(h_{x_k}\) and \(S_{x_k}\) denote a hazard and a survival function for \(x_k\), respectively. \(I^k\) is a dummy variable (\(I^k = 1\) if a state \(y_i = k\) is observed at time \(t_i\) and \(I^k = 0\) if a state \(y_i \neq k\) is observed at time \(t_i\).

Each of \(x_{i,k}\) can be decomposed as:

\[
x_{i,k} = \Phi_{i,k} x_{i,k},
\]

where \(\Phi_{i,k} = \Psi_{i,k} \Gamma(1+\gamma^{-1})^{-1}, \Psi_{i,k} = E(x_{i,k} | \mathcal{F}_{i-1})\) and \(x_{i,k}\) is an i.i.d. Weibull-distributed random variable with a shape parameter \(\gamma\), satisfying \(E(x_{i,k}) = \Gamma(1+\gamma^{-1})\). We propose the following specification for the non-linear transformations of conditional mean functions:

\[
\lambda_{i,k} - 1 = \sum_{l=1}^{4} \alpha_{l,k} I_{l-1}^k + \omega_{l,k} x_{i,k} - 1, \quad \lambda_{i,k} \geq 0,
\]

\[
+ \beta_{l,k} \psi_{l,k}^\lambda_{i,k} - 1, \quad \psi_{l,k} \geq 0,
\]

where \(\lambda_{1,k}, \lambda_{2,k}, \omega_{l,k}, \alpha_{l,k}, \beta_{l,k}\) (for \(l = 1, 2, 3, 4\)) denote the model parameters to be estimated. The autoregressive specification of Eq. (3) allows to capture strong autocorrelation of inter-order durations. Intercepts and parameters corresponding to lagged durations change with the previously observed orders (according to the binary indicators \(I_{l-1}^k\)), which allows for capturing significant cross-correlations among orders as indicated in Fig. 2.

In our data, we cannot identify orders placed according to the algorithmic trading (AT) strategies, however the presence of high-frequency traders in FX markets has been lately investigated in Ref. [41]. It has been shown that AT improves market efficiency i.e. there is a reduction of arbitrage opportunities and faster price discovery.

3. Multistate Asymmetric Box-Cox ACD model (MABCACD)

We suggest a non-linear extension of the Multistate Asymmetric ACD (MAACD) model proposed in Ref. [20]. The MAACD model itself can be treated as a generalisation of the Asymmetric ACD Model proposed by L. Bauwens and P. Giot [16] and applied in empirical market microstructure analyses in Ref. [42–44].

\[
E_{\text{Econometric Modeling of Inter-Order Durations}}
\]
The Box-Cox parameters $\lambda_{1,k}$ and $\lambda_{2,k}$ allow for concave, convex or linear relationship between 1-period lagged duration and the current duration expectation conditional upon $F_{t-1}$. This Box-Cox version of the MAACD model nests the logarithmic version of the model for $\lambda_{1,k} \to 0$ and linear specification for $\lambda_{1,k} \lambda_{2,k} = 1$. Standard univariate Box-Cox ACD models were used in Ref. [11].

In this study, we suggest to impose the same shape parameter $\lambda$ for the Weibull-distributed error terms $\varepsilon_{i,k}$ (for $k = 1, 2, 3, 4$)\(^*\). As we will show, such a restriction allows for closed-form expressions for conditional duration expectations $E(x_i|F_{t-1})$ and conditional transition probabilities. Given the Weibull-distributed error terms $\varepsilon_{i,k}$, the joint conditional density of $x_i$ and $y_i$ can be rewritten as:

$$f(x_i, y_i|F_{t-1}) = \prod_{k=1}^{4} \left[ \frac{\gamma}{\Phi_{1,k}} \left( \frac{x_i}{\Phi_{1,k}} \right)^{-\gamma-1} \right] t_i^\gamma e^{-\left(\frac{x_i}{\Phi_{1,k}}\right)^\gamma}. \quad (4)$$

Given the bivariate conditional density of $x_i$ and $y_i$, the MAACD model can be estimated with the Maximum Likelihood (ML) method. The log-likelihood function can be derived by taking natural logarithms of both sides of Eq. (4) and summing over all observations.

The conditional transition probabilities can be derived as:

$$P(y_i = k|F_{t-1}) = \int_0^\infty \left[ \frac{\gamma}{\Phi_{1,k}} \left( \frac{x_i}{\Phi_{1,k}} \right)^{-\gamma-1} \right] \frac{\Phi_{1,k}}{\Phi_{1,k}} e^{\frac{x_i}{\Phi_{1,k}}} dx_i = \frac{\Phi_{1,k}}{\sum_{k=1}^{4} \Phi_{1,k}} \Phi_{1,k}^{-\gamma}. \quad (5)$$

Furthermore,

$$f(x_i|F_{t-1}) = \sum_{k=1}^{4} \left[ \frac{\gamma}{\Phi_{1,k}} \left( \frac{x_i}{\Phi_{1,k}} \right)^{-\gamma-1} \right] \prod_{k=1}^{4} \frac{\Phi_{1,k}}{\Phi_{1,k}} e^{-\left(\frac{x_i}{\Phi_{1,k}}\right)^\gamma} = \gamma x_i^{-\gamma} \left( \sum_{k=1}^{4} \Phi_{1,k}^{-\gamma} \right) \left( \sum_{k=1}^{4} \Phi_{1,k}^{-\gamma} \right)^{-1}. \quad (6)$$

Accordingly, the conditional distribution of $x_i$ is Weibull\(^*\) with a shape parameter $\gamma$ and a scale parameter $\left( \sum_{k=1}^{4} \Phi_{1,k}^{-\gamma} \right)^{-1/\gamma}$. The goodness-of-fit of the MAACD model can be assessed by checking for the lack of autocorrelation in residuals:

$$e_i = \frac{x_i}{\left( \sum_{k=1}^{4} \Phi_{1,k}^{-\gamma} \right)^{-1/\gamma}}. \quad (7)$$

4. Estimation results and diagnostic checks

Estimation results of the MABCACD model are reported in Table. The log-likelihood function has been pre-programmed and maximised with the BFGS algorithm in the Maximum Likelihood (Maxlik) application of the Gauss vs. 13. software. All the parameter estimates are statistically different from zero. The null hypothesis of $\gamma = 1$ can also be rejected, thus the exponential distribution is too restrictive for $\varepsilon_{i,k}$. Therefore, each of four point processes corresponding to four order types (cf. Eq. (2)) is not memoryless and the corresponding baseline hazard is monotonically decreasing. In contrast to the logarithmic specification of the MAACD model [20], parameter estimates of the Box-Cox transformations are significantly different from zero, which confirms the nonlinear relationship between 1-period lagged durations and the current durations expectations for all types of orders.

\(^*\)The validity of this restriction can be tested with empirical data.

\(^*\)This property is also shown in Ref. [16].

### Table: ML parameter estimates of the Multistate Asymmetric Box-Cox ACD model for order submissions in the interbank EUR/PLN spot market.

<table>
<thead>
<tr>
<th>MS, $k = 1$</th>
<th>MB, $k = 2$</th>
<th>IQS, $k = 3$</th>
<th>IQB, $k = 4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\omega_{1,k}$</td>
<td>0.443</td>
<td>1.789</td>
<td>1.597</td>
</tr>
<tr>
<td>$\omega_{2,k}$</td>
<td>1.535</td>
<td>0.593</td>
<td>0.723</td>
</tr>
<tr>
<td>$\omega_{3,k}$</td>
<td>1.091</td>
<td>0.786</td>
<td>0.712</td>
</tr>
<tr>
<td>$\omega_{4,k}$</td>
<td>0.788</td>
<td>1.105</td>
<td>1.265</td>
</tr>
<tr>
<td>$\alpha_{1,k}$</td>
<td>0.701</td>
<td>0.695</td>
<td>0.732</td>
</tr>
<tr>
<td>$\alpha_{2,k}$</td>
<td>0.278</td>
<td>0.418</td>
<td>0.699</td>
</tr>
<tr>
<td>$\alpha_{3,k}$</td>
<td>0.321</td>
<td>0.328</td>
<td>0.464</td>
</tr>
<tr>
<td>$\alpha_{4,k}$</td>
<td>0.311</td>
<td>0.336</td>
<td>0.442</td>
</tr>
<tr>
<td>$\beta_{1,k}$</td>
<td>0.289</td>
<td>0.302</td>
<td>0.586</td>
</tr>
<tr>
<td>$\lambda_{1,k}$</td>
<td>0.148</td>
<td>0.183</td>
<td>0.302</td>
</tr>
<tr>
<td>$\lambda_{2,k}$</td>
<td>0.138</td>
<td>0.206</td>
<td>0.554</td>
</tr>
</tbody>
</table>

In order to judge about the superiority of the MAB-CACD model, we also compared its Bayesian information criterion (BIC) with the criteria corresponding to linear and logarithmic version of the MAACD. The BIC for the Box-Cox model was the smallest one, thus the non-linear extension of the model allows for the best fit among these three specifications.

In order to check for the adequacy of the dynamic specification, in Fig. 3 we compare the ACF of the model residuals (cf. Eq. (7)), with the ACF for the inter-order duration series. We can see that the huge amount of autocorrelation has been successfully explained. There is still some significant autocorrelation left, but taking into account the scale of order clustering, the results are satisfactory.

In Fig. 4 we plotted the time-varying conditional probabilities (cf. Eq. (5)). Their visual inspection allows to confirm striking lead- and lag dependencies among order types. We see that each of order types clusters in times when its conditional probability is systematically higher. Moreover, the MS orders are often followed by the IQB orders as indicated by the significant cross-correlation at the lag 1 (cf. Fig. 2). Symmetrically,
Fig. 3. Comparison between the ACF of the Multistate Asymmetric Box-Cox ACD residuals (in black) with the ACF for the inter-order durations (in grey). Horizontal lines depict 99% confidence interval.

MB orders cluster together with the IQS orders. We can also see that conditional probabilities of the MS orders are positively correlated with the conditional probabilities of the IQB orders, but negatively correlated with the probabilities of observing an MB order or an IQS order. As the theoretical probabilities of registering a given order type agree with their observed frequency, the MABCACD model can serve as a tool for depicting dealers’ trading strategies.

Fig. 4. Comparison between the time-varying transition probabilities for different order types from the MABCACD model and observed order submissions for 100 consecutive data points.

5. Conclusions

In this study we have investigated two dynamic properties of the order submission process: (1) strong autocorrelation of inter-order durations and (2) significant lead-lag relationships among four major order types. We proposed the Multistate Asymmetric Box-Cox ACD model that aims at capturing non-linear relations in the specification of conditional duration expectations. This joint model for the indicators of the order type and corresponding waiting times is able to depict both order clustering and causal patterns embedded in dealers’ actions. The restriction that we posed on the shape parameter of the duration distribution leads to simple formulas for the time-varying conditional probabilities of order types and for conditional duration expectations. We showed how these can be used to construct model residuals.

Further extensions to this study may take into account different distribution families for the duration variable and more general specifications of their conditional mean functions, for example taking the long range dependence into account, as it is done for the univariate fractionally integrated ACD models [45, 46]. In the current study, we took advantage from the vast data set covering two months of inter-dealer trading in January–February 2008, which corresponds to constant appreciation of the Polish zloty. However, it might be reasonable to verify the stability of the obtained relations for the periods of reverse market trends, as well as to allow for a possibly non-linear impact of additional exogenous explanatory variables (bid-ask spread, bid/ask depth, volatility).

Acknowledgments

We thank the Thomson Reuters for providing the data for this study. This research has been carried out within the project “The Microstructure of the Interbank FX Spot Market” financed by the National Science Centre in Poland upon decision No. DEC-2013/09/B/HS4/01319.

References