Modeling Financial (Inter-Order) Durations in the Reuters 3000 Spot Matching System

Katarzyna Bień – Barkowska

Institute of Econometrics, Warsaw School of Economics
& Financial System Department, National Bank of Poland

FENS
Lublin, 15^{th} *May 2014*

Aim of the Study:

Examination of the dealers' behavior on the order-driven market for the EUR/PLN currency pair (order submission rules)

Contribution to the Existing Literature:

1. Generalization of the Asymmetric ACD model of Bauwens and Giot (2003) (to the case of more than two competing risks).
2. Bivariate modeling of order timing and the strength of order aggressiveness (inspiration: Lo and Sapp (2008) study).
3. Empirical verification of selected theoretical market microstructure hypotheses for the FX spot market.

Plan of the Presentation:

1. Trading in the Reuters Dealing 3000 Spot Matching System.
2. Multistate version of the asymmetric ACD model of Bauwens & Giot (2003) (more than two competing risks).
3. Posible generalisations of the Multistate Asymmetric ACD model of Bien-Barkowska (2014) (1. Spillover effects, 2. Box-Cox transformations).
4. Simulation results.

Market and Data:

The FX market of the polish zloty is the most liquid among all the currency markets of Central European emerging economies. According to the survey of the FX market activity conducted by the Bank for International Settlements, the average daily turnover in interbank spot transactions accounted to 4,851 million USD in April 2007.

Reuters Dealing 3000 Spot Matching System accounts for 40 % of the whole turnover (the offshore and the local market).

Datasets from the Reuters Dealing 3000 Spot Matching System:

- Order Data: Exact date and time of submission as well as execution/cancellation, a firm quote, a size and an indicator for a market side of a quote

Exemplary order book

Level of order book	ASK/OFFER – Orders to sell euro		BID – Orders to buy euro	
1	3.9220	2	3.9208	2
2	3.9228	1	3.9185	1
3	3.9240	1	3.9175	1
4	3.9244	1	3.9166	2
5	3.9245	1	3.9135	1
6	3.9255	1	3.8930	8
7	3.9266	3		
8	3.9285	1		
9	3.9300	1		
10	3.9310	2		
11	3.9415	1		
12	3.9425	1		
13	3.9435	1		
14	3.9455	1		
15	3.9465	1		
16	3.9500	2		
17	3.9500	1		

Diurnality pattern

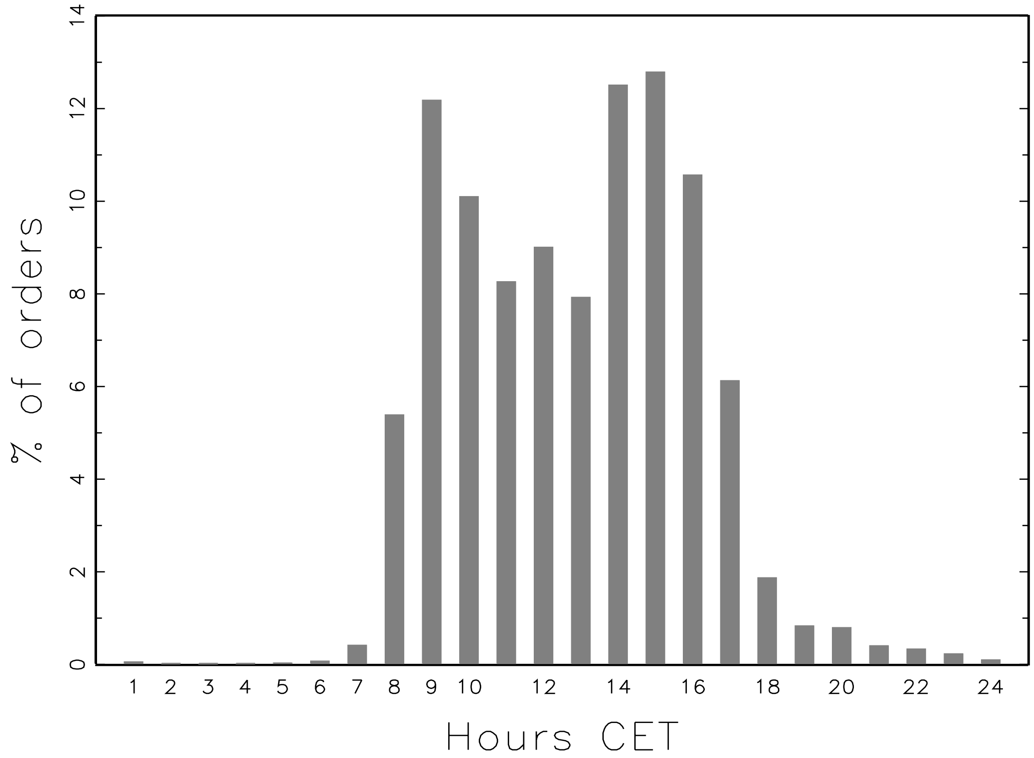


Figure 1: Intraday order submission frequency.

Four different events have been defined for orders registered in January-February 2009 (74,745 obs):

- $k = 1$ (MS) – Submission of a market sell order or a marketable limit sell order; in this case the price of the incoming sell order is lower or equal to the most competitive bid price prevailing in the order book (15,839 obs).
- $k = 2$ (MB) – Submission of a market buy order or a marketable limit buy order; in this case the price of the incoming buy order is higher or equal to the most competitive ask price prevailing in the order book (16,205 obs).
- $k = 3$ (IQS)– Submission of an inside-the-quote limit sell order; in this case the price of the incoming sell order is lower than the best ask price but higher than the best bid price (21,998 obs).
- $k = 4$ (IQB)– Submission of an inside-the-quote limit buy order; in this case the price of the incoming buy order is higher than the best bid price but lower than the best ask price (20,703 obs).

Asymmetric ACD models (AACD)

Asymmetric Autoregressive Conditional Duration (ACD) model of (Bauwens, Giot 2003) is a flexible model for the conditional density of financial durations that can elapse once one of two possible end states (i.e. events pointed on the micro-scale by appropriate "thinning" the data) occurs.

The AACD model describes the marked point process $\{x_i, y_i\}$, where $x_i = t_i - t_{i-1}$ is a duration between the moments in which certain events occur: t_i and t_{i-1}, and y_i is a qualitative variable indicating a type of an event: $y_i \in \{a, b\}$.

The duration x_i is treated as an outcome variable of a function $x_i = \min(x_{i,a}, x_{i,b})$, where $x_{i,a}$ and $x_{i,b}$ are durations that would end up in state a and b, respectively.

In a very close analogy to Bauwens, Giot, 2003, we consider the model for the marked point process $\{x_i, y_i\}$, where $x_i = t_i - t_{i-1}$ is a duration between the moments in which subsequent orders arrive to the system and y_i is an indicator variable for a particular type of an event $y_i = k$ (where $k = 1, 2, 3, 4$).

Accordingly, x_i can be treated as an outcome variable of a function $x_i = \min(x_{i,1}, x_{i,2}, x_{i,3}, x_{i,4})$, where each of variables $x_{i,k}$ (for $k = 1, 2, 3, 4$) corresponds to an order duration that would end in the state k.

As in the standard framework of a competing risks model, we consider a joint conditional bivariate density for x_i and y_i[a]:

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

where I^k is a dummy variable ($I_i^k = 1$ if a state $y_i = k$ is observed at time t_i and $I_i^k = 0$ if a state $y_i \neq k$ is observed at time t_i). $\mathcal{F}_{i-1}$ denotes an information set up to a time point $t-1$ that contains past realizations of x_i and y_i, h_{x_k} and S_{x_k} denote a hazard and a survival function for x_k, respectively.

[a] The model assumes independence (conditionally on $\mathcal{F}_{i-1}$) between durations $x_{i,1}, x_{i,2}, .x_{i,3}, x_{i,4}$.

For example, if a state MS is observed at t_i, the conditional bivariate density of the pair $\{x_i, y_i\}$ is given by:

$$\begin{aligned} f(x_i, y_i = 1|\mathcal{F}_{i-1}) &= h_{x_1}(x_i|\mathcal{F}_{i-1})^{I_i^1} S_{x_1}(x_i|\mathcal{F}_{i-1}) \prod_{k=2}^{4} S_{x_k}(x_i|\mathcal{F}_{i-1}) \\ &= f_{x_1}(x_i|\mathcal{F}_{i-1}) \prod_{k=2}^{4} S_{x_k}(x_i|\mathcal{F}_{i-1}) \end{aligned}$$

Therefore, if a duration x_i ends with an MS order $(y_i = 1)$, x_i contributes to the density function via: (1) the conditional density of $x_{i,1}$ evaluated at x_i, i.e. $f_{x_1}(x_i|\mathcal{F}_{i-1})$ and (2) the joint conditional probability that all other unobserved durations $x_{i,k}$ (with end states $k = 2, 3, 4$) are longer than the realized duration x_i: $\prod_{k=2}^{4} S_{x_k}(x_i|\mathcal{F}_{i-1})$.

The conditional expectations are modelled in a dynamic fashion, such as previous states and previously observed durations could exert an influence on their length. In the standard framework of the ACD model of Engle and Russell (1998) $x_{i,k}$ is given as:

$$x_{i,k} = \Phi_{i,k}\varepsilon_{i,k} \tag{2}$$

where $\Phi_{i,k} = \Psi_{i,k} \cdot \mu_{\varepsilon_i}^{-1}$, $\Psi_{i,k} = E(x_{i,k}|\mathcal{F}_{i-1})$ and μ_{ε_i} is the mean of the Weibull distribution.

Conditional (with respect to $\mathcal{F}_{i-1}$) duration expectations are modelled with the Box-Cox Generalization of the ACD model. The transformed duration expectations: $\psi_{i,k} = (\Psi_{i,k}^{\lambda_{1,k}} - 1)/\lambda_{1,k}$ are:

$$\psi_{i,k} = \sum_{l=1}^{4} \omega_{l,k} I_{i-1}^{l} + \alpha_k (x_{i-1}^{\lambda_{2,k}} - 1)/\lambda_{2,k} + \sum_{j=1}^{4} \beta_j \psi_{i-1,j} \qquad (3)$$

where $l = 1,2,3,4$, $j = 1,2,3,4$ and I_i^l is a dummy indicator ($I_i^l = 1$, if a state $y_i = l$ at the end of duration x_i and $I_i^l = 0$, if $y_i \neq l$).

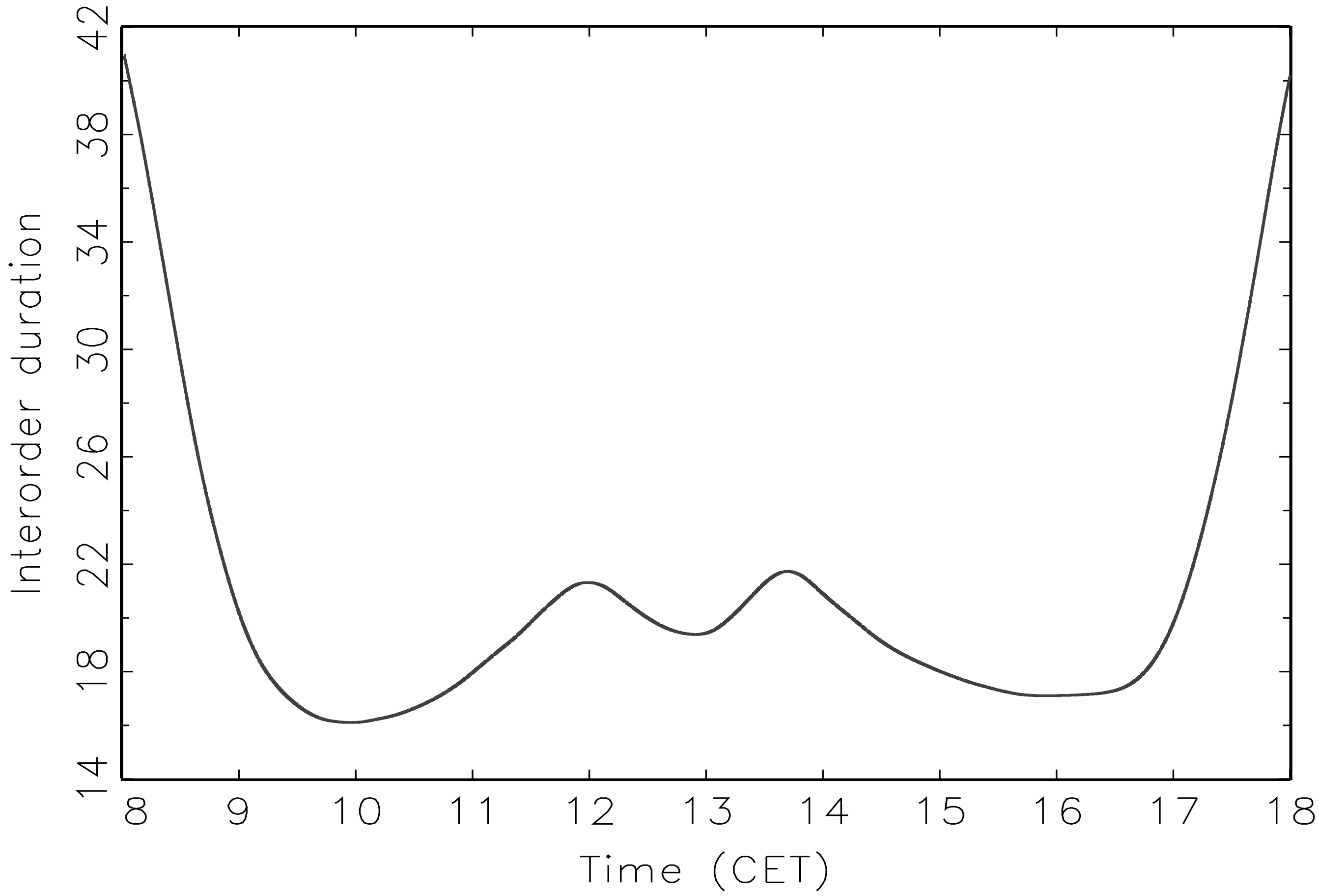
Assuming the Weibull distribution for the error terms $\varepsilon_{i,k}$, the joint conditional density function for the pair $\{x_i, y_i\}$ can be derived as:

$$f(x_i, y_i|\mathcal{F}_{i-1}) = \prod_{k=1}^{4} \left[\frac{\gamma_k}{\Phi_{i,k}} \left[\frac{x_i}{\Phi_{i,k}} \right]^{\gamma_k - 1} \right]^{I_i^k} \cdot e^{-\left[\frac{x_i}{\Phi_{i,k}} \right]^{\gamma_k}} \quad (4)$$

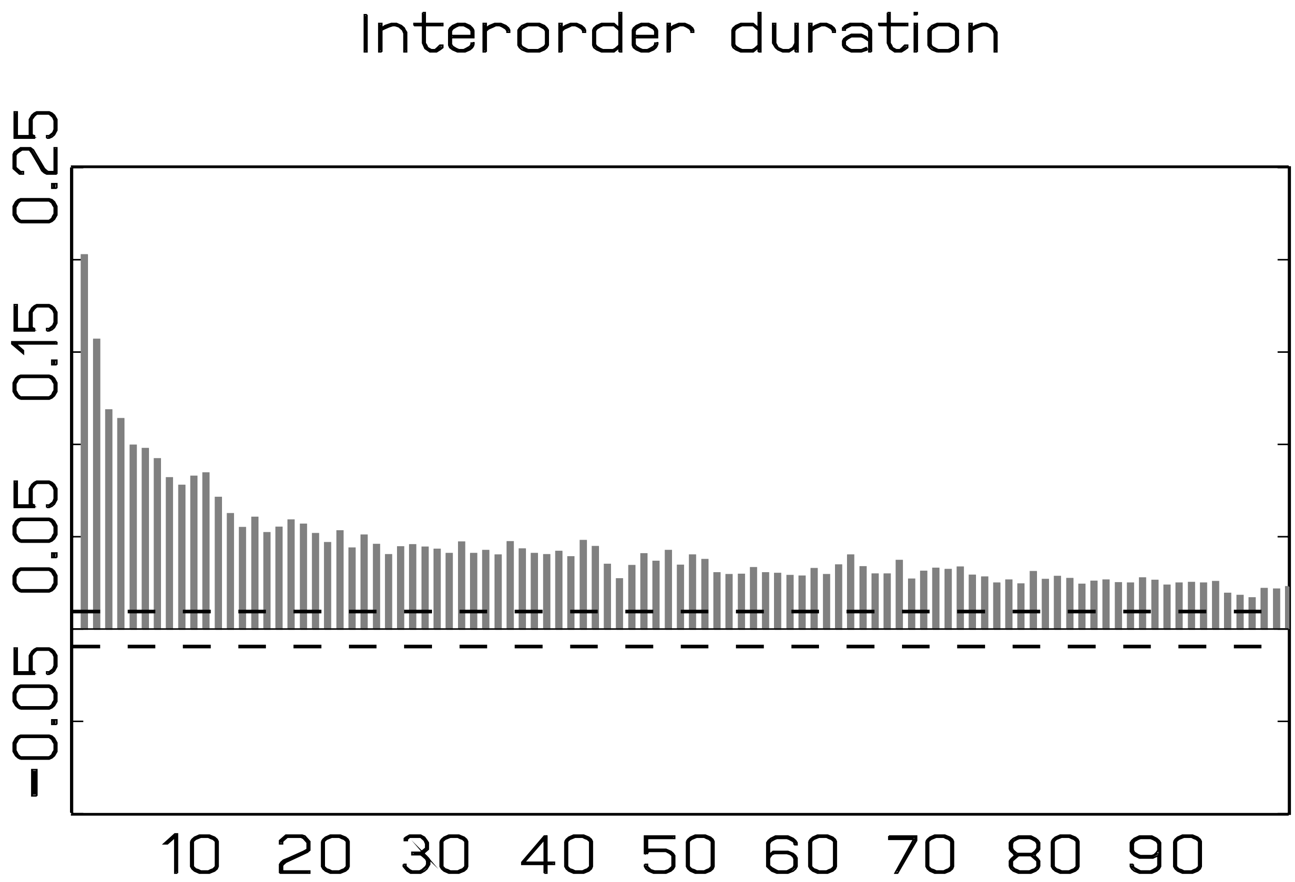
Accordingly, the joint log-likelihood function can be given as the sum of four log-likelihoods:

$$\begin{aligned} \ln L(\Theta|x_i, y_i, \mathcal{F}_{i-1}) &= \sum_{i=1}^{N} \sum_{k=1}^{4} \ln L_k(\Theta_k|x_i, y_i, \mathcal{F}_{i-1}) \\ &= \sum_{i=1}^{N} \sum_{k=1}^{4} I_i^k \left[\ln \left[\frac{\gamma_k}{\Phi_{i,k}} \left[\frac{x_i}{\Phi_{i,k}} \right]^{\gamma_k - 1} \right] \right] - \sum_{i=1}^{N} \sum_{k=1}^{4} \left[\frac{x_i}{\Phi_{i,k}} \right]^{\gamma_k} \end{aligned}$$

Intraday seasonality of inter-order durations:



Autocorrelation function of inter-order durations:



Estimation results:

	MS, k=1		MB, k=2		IQS, k=3		IQB, k=4	
parameters	est.	p	est.	p	est.	p	est.	p
$\omega_{1,k}$ MS	0.9140	0.0000	1.8360	0.0000	0.1595	0.3721	-0.4948	0.0005
$\omega_{2,k}$ MB	1.8123	0.0000	0.9266	0.0000	-0.4939	0.0019	0.1979	0.1998
$\omega_{3,k}$ IQS	1.3473	0.0000	0.5746	0.0052	0.0134	0.9314	0.0214	0.8767
$\omega_{4,k}$ IQB	0.8025	0.0000	1.2617	0.0000	0.0087	0.9532	0.0772	0.5678
$\beta_{1,k}$ MS	0.2110	0.0000	0.1054	0.5241	0.1853	0.0925	0.4574	0.0000
$\beta_{2,k}$ MB	-0.0804	0.4179	0.1748	0.0048	0.3493	0.0000	0.1273	0.0294
$\beta_{3,k}$ IQS	-0.1685	0.2891	0.9148	0.0000	0.3766	0.0000	0.1962	0.0455
$\beta_{4,k}$ IQB	0.6845	0.0000	-0.4474	0.0400	0.0591	0.6929	0.2014	0.0393
$\alpha_{1,k}$	2.1272	0.0000	1.8491	0.0000	-1.1859	0.0000	1.9293	0.0000
$\alpha_{2,k}$	-0.7191	0.0001	-0.8430	0.0000	1.9923	0.0000	-1.1202	0.0000
$\lambda_{1,k}$	0.1256	0.0000	0.2087	0.0000	0.3184	0.0000	0.2582	0.0000
$\lambda_{2,k}$	0.1693	0.0000	0.2138	0.0000	0.2274	0.0000	0.2128	0.0000
γ_k	0.7309	0.0000	0.7315	0.0000	0.7392	0.0000	0.7306	0.0000

Estimation results

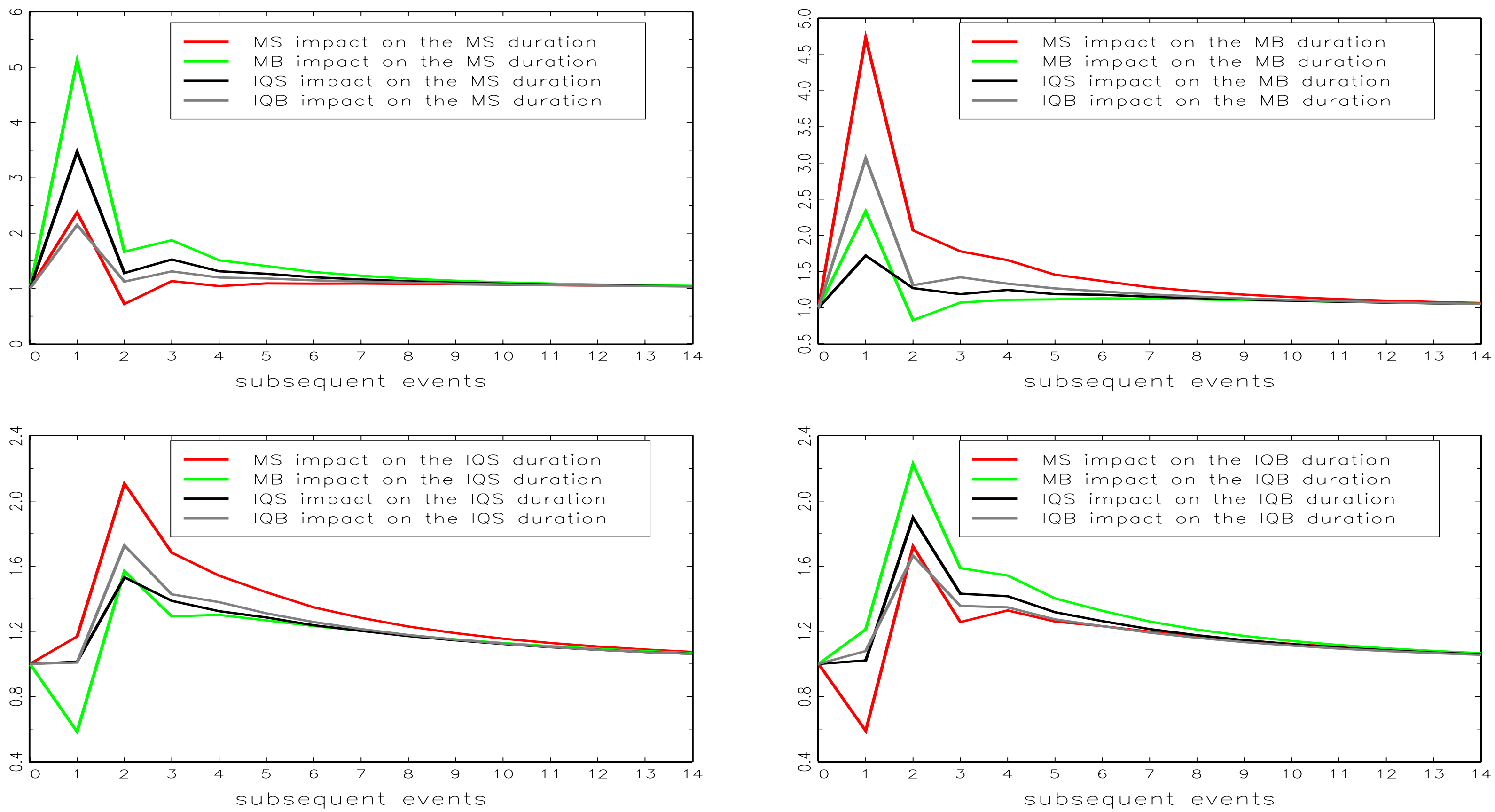


Figure 2: Impulse response functions: Expected duration in result of a given event. Upper left: Expected MS duration, upper right: expected MB duration, lower left: expected IQS duration, lower right: expected IQB duration.

ESTIMATION RESULTS

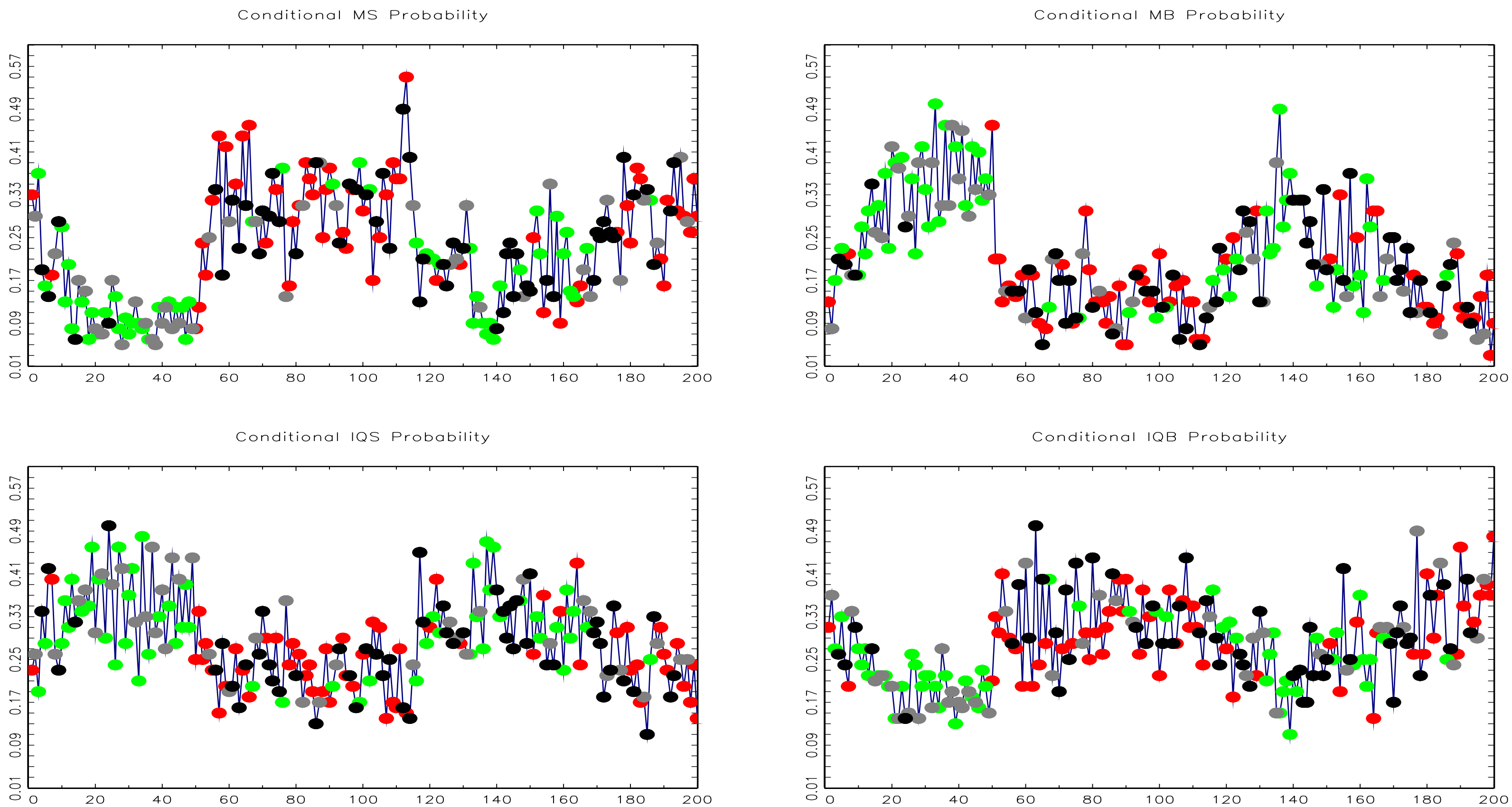
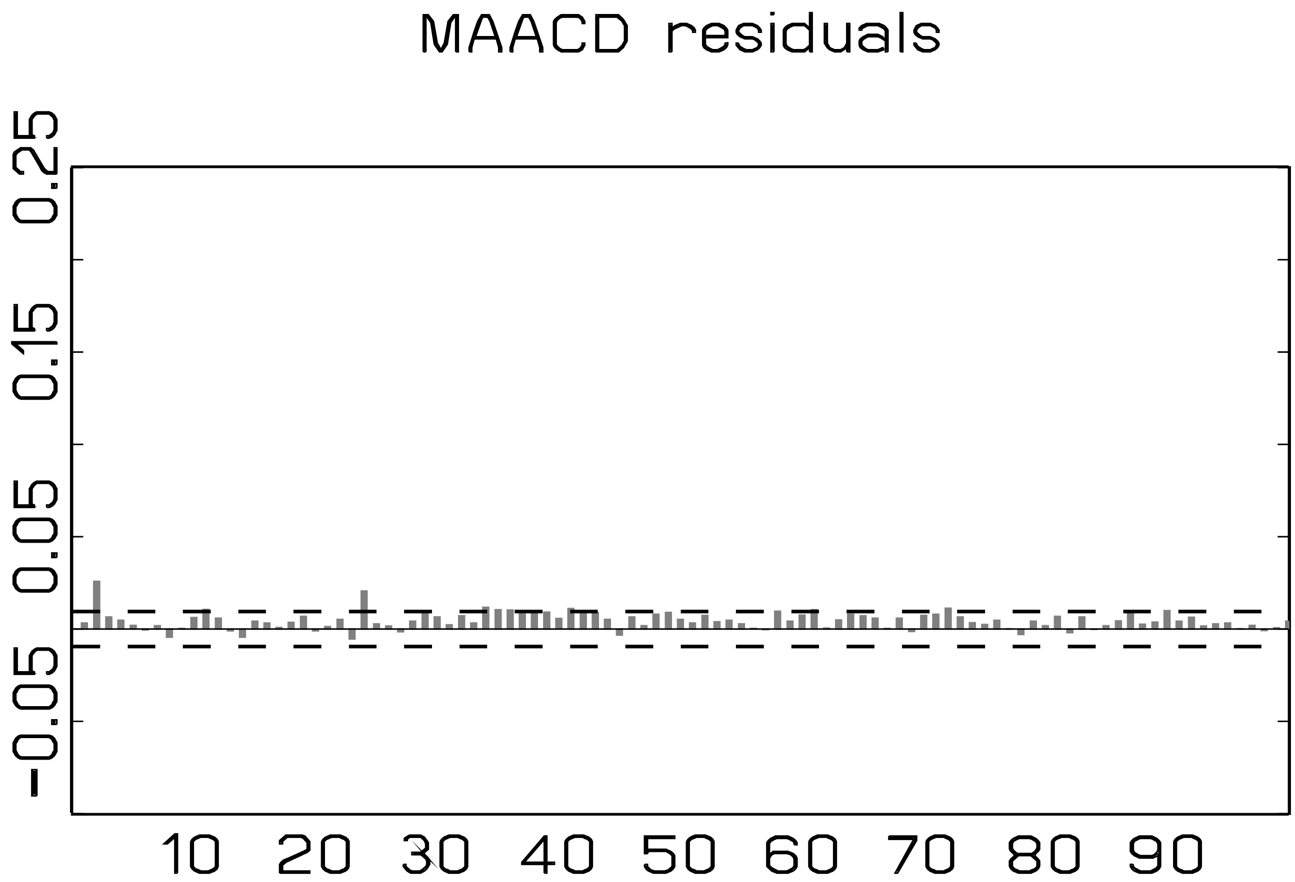


Figure 3: Time-varying conditional probabilities of order arrivals. Red dots – market sell orders, green dots – market buy orders, grey dots – in-the-quote sell orders, black dots – in-the-quote buy orders.

Autocorrelation function of MAACD residuals:



The simulation algorithm can be outlined as following:

1. For $i = 1$, set initial values for $\hat{\Psi}_{i,k}$ as mean estimates of (observable) series $x_{i,k}$ (for $k = 1, 2, 3, 4$).
2. For $k = 1, 2, 3, 4$, draw n values of $\varepsilon_{j,k}$ $(j = 1, 2, ..., n)$ from independent Weibull distributions. Each distribution is characterized by the corresponding shape parameter $\hat{\gamma}_k$.
3. Compute $x_{i,k} = \hat{\Phi}_{i,k}\varepsilon_{i,k}$, where $\hat{\Phi}_{i,k} = \hat{\Psi}_{i,k} \cdot \mu_{\varepsilon_k}^{-1} = \hat{\Psi}_{i,k}(\Gamma(1 + \hat{\gamma_k}^{-1}))^{-1}$.
4. Set $y_i = l$ and $x_i = x_{i,l}$ if a duration $x_{i,l}$ is the shortest (for $l = 1, 2, 3, 4$), i.e. $x_{i,l} = \min\{x_{i,1}, ..., x_{i,l}, ..., x_{i,10}\}$.
5. Compute $\hat{\Psi}_{i+1,k}$ with obtained y_i and x_i.
6. Iterate from (3) where i=i+1.

Simulation results:

Transition probabilities – simulation results
(column 1 contains type of a directly preceding event):

	MS	MB	IQS	IQB
MS (simulated)	0.2393	0.1365	0.2599	**0.3641**
MS (real data)	0.2646	0.1301	0.2411	**0.3641**
MB (simulated)	0.1164	0.2437	**0.3765**	0.2632
MB (real data)	0.1292	0.2537	**0.3784**	0.2385
IQS (simulated)	0.1692	0.2862	**0.2883**	0.2571
IQS (real data)	0.1765	**0.2875**	0.2826	0.2532
IQB (simulated)	**0.2743**	0.1769	0.2892	0.2594
IQB (real data)	**0.2747**	0.1829	0.2707	0.2715

Thank you very much for your attention!

katarzyna.bien@sgh.waw.pl