Adaptive Order Flow Forecasting with Multiplicative Error Models

Wolfgang Karl Härdle
Andrija Mihoci
Christopher Hian-Ann Ting

Ladislaus von Bortkiewicz Chair of Statistics
C.A.S.E. – Center for Applied Statistics and Economics
Humboldt–Universität zu Berlin
Lee Kong Chian School of Business
Singapore Management University
http://lvb.wiwi.hu-berlin.de
http://www.case.hu-berlin.de
http://business.smu.edu.sg

Trading Volume

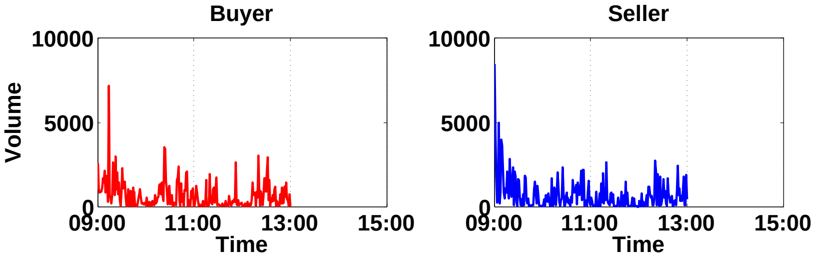


Figure 1: One-minute aggregated size of buyer-initiated and seller-initiated trades for the 'mini Nikkei 225 index futures' traded at the Osaka Securities Exchange on 20130305 ▸ Volume (Description)

Trading Volume

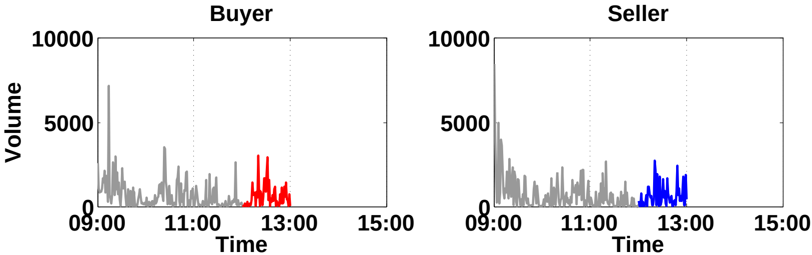


Figure 2: One-minute aggregated size of buyer-initiated and seller-initiated trades for the 'mini Nikkei 225 index futures' traded at the Osaka Securities Exchange on 20130305 with an estimation window length of 60

Trading Volume

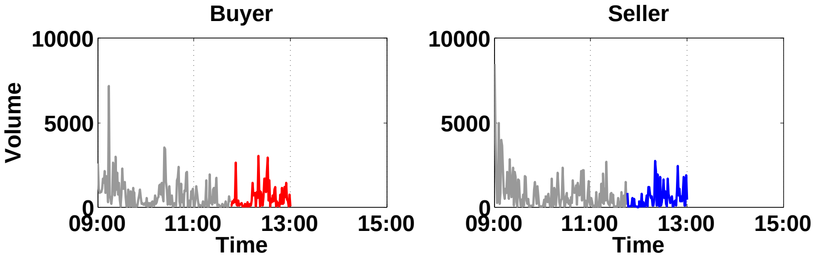


Figure 3: One-minute aggregated size of buyer-initiated and seller-initiated trades for the 'mini Nikkei 225 index futures' traded at the Osaka Securities Exchange on 20130305 with an estimation window length of 75

Trading Volume

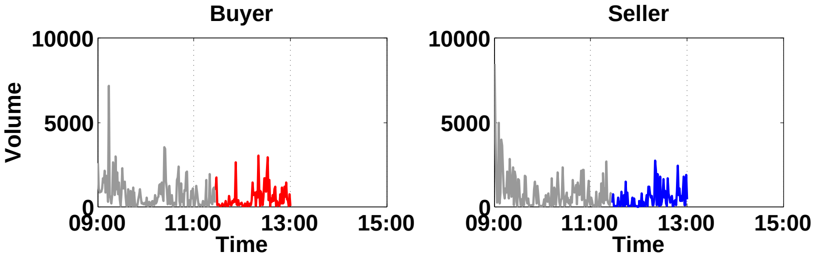


Figure 4: One-minute aggregated size of buyer-initiated and seller-initiated trades for the 'mini Nikkei 225 index futures' traded at the Osaka Securities Exchange on 20130305 with an estimation window length of 94

Trading Volume

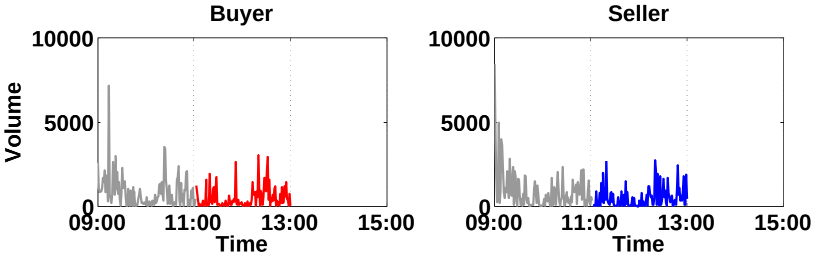


Figure 5: One-minute aggregated size of buyer-initiated and seller-initiated trades for the 'mini Nikkei 225 index futures' traded at the Osaka Securities Exchange on 20130305 with an estimation window length of 118

Order Flow

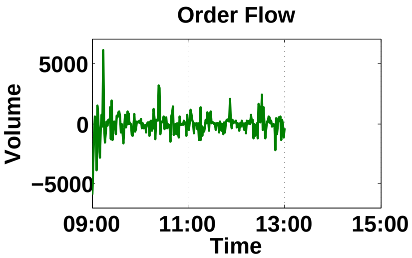


Figure 6: Order flow (difference between the buyer-initiated and the seller-initiated volume) for the 'mini Nikkei 225 index futures' traded at the Osaka Securities Exchange on 20130305

Example

A brokerage decides to trade contracts over a forecasted period (e.g., 5min.) a trading day. A position (buy/sell) is entered ahead the forecasted period and closed afterwards.

Trading Strategies

- Strategy (i) 'Buy if predicted order flow is positive and sell if negative'
- Strategy (ii): 'Buy if predicted order flow is positive'
- Strategy (iii): 'Sell if predicted order flow is negative'

Objectives

- Short-Term Order Flow Forecasting
  - Local adaptive Multiplicative Error Model (MEM)
  - Balance between modelling bias and parameter variability
  - Estimation windows with potentially varying lengths
- Intra-Day Trading
  - Buyer- and seller-initiated trades and order flow dynamics
  - Calibration and evaluation of trading strategies

Statistics and Quantitative Finance

Statistics

- Modelling bias vs. parameter variability
- Flexible framework and predictive accuracy

Quantitative Finance Practice

- Trading strategies, threshold selection
- Performace evaluation, equity curves

Outline

1. Motivation ✓
2. Local Adaptive Multiplicative Error Model (MEM)
3. Order Flow Dynamics
4. Order Flow Forecasting
5. Intra-Day Trading
6. Conclusions

Multiplicative Error Model (MEM)

- Engle (2002), MEM(p,q), $\mathcal{F}_i$ - information set up to i

$$y_i = \mu_i \varepsilon_i, \qquad \mathsf{E}\left[\varepsilon_i \,|\mathcal{F}_{i-1}\right] = 1$$

$$\mu_i = \omega + \sum_{j=1}^{p} \alpha_j y_{i-j} + \sum_{j=1}^{q} \beta_j \mu_{i-j}, \qquad \omega > 0, \alpha_j, \beta_j \geq 0$$

- Hautsch (2012) - comprehensive MEM literature overview
  - y_i - squared (de-meaned) log return: GARCH(p,q)
  - y_i - volume, bid-ask spread, duration: ACD(p,q)

Engle, Robert F. on BBI:

Autoregressive Conditional Duration (ACD)

1. Exponential-ACD, Engle and Russel (1998) ▸ EACD

   $\varepsilon_i \sim Exp(1)$, $\boldsymbol{\theta}_E = (\omega, \boldsymbol{\alpha}, \boldsymbol{\beta})^\top$, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_p)$, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_q)$

2. Weibull-ACD, Engle and Russel (1998) ▸ WACD

   $\varepsilon_i \sim \mathcal{G}(s, 1)$, $\boldsymbol{\theta}_W = (\omega, \boldsymbol{\alpha}, \boldsymbol{\beta}, s)^\top$

Weibull, E. H. Waloddi on BBI:

Parameter Estimation

- Consistent parameter estimation
- Data calibration with time-varying intervals
- Quasi maximum likelihood estimates (QMLEs) of $\boldsymbol{\theta}_E$ and $\boldsymbol{\theta}_W$

$$\widetilde{\boldsymbol{\theta}}_I = \arg\max_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} L_I\left(y; \boldsymbol{\theta}\right) \tag{1}$$

- $I = [i_0 - n, i_0]$ - interval of $(n+1)$ observations at i_0
- $L_I(\cdot)$ - log likelihood, EACD and WACD

Estimation Quality

- Mercurio and Spokoiny (2004), Spokoiny (2009)
- Quality of estimating *true* parameter vector $\boldsymbol{\theta}^*$ by QMLE $\widetilde{\boldsymbol{\theta}}_I$ in terms of Kullback-Leibler divergence; $\mathcal{R}_r(\boldsymbol{\theta}^*)$ - risk bound

$$\mathsf{E}_{\boldsymbol{\theta}^*}\left|L_I(\widetilde{\boldsymbol{\theta}}_I) - L_I(\boldsymbol{\theta}^*)\right|^r \leq \mathcal{R}_r(\boldsymbol{\theta}^*)$$

▸ Gaussian Regression

- Likelihood based confidence sets

- 'Modest' risk, $r = 0.5$ (shorter intervals of homogeneity)
- 'Conservative' risk, $r = 1$ (longer intervals of homogeneity)

Kullback, Solomon and Leibler, Richard A. on BBI:

Local Parametric Approach (LPA)

- LPA, Spokoiny (1998, 2009)
  - Time series parameters can be locally approximated
  - Finding the (longest) *interval of homogeneity*
  - Balance between modelling bias and parameter variability

- Time series literature
  - GARCH$(1,1)$ models - Čížek et al. (2009)
  - Realized volatility - Chen et al. (2010)
  - Multiplicative Error Models - Härdle et al. (2014)

Interval Selection

- $(K+1)$ nested intervals with length $n_k = |I_k|$

$$\begin{array}{ccccccc} I_0 & \subset & I_1 & \subset \cdots \subset & I_k & \subset \cdots \subset & I_K \\ \widetilde{\boldsymbol{\theta}}_0 & & \widetilde{\boldsymbol{\theta}}_1 & & \widetilde{\boldsymbol{\theta}}_k & & \widetilde{\boldsymbol{\theta}}_K \end{array}$$

Example: Trading volumes aggregated over 1-min periods

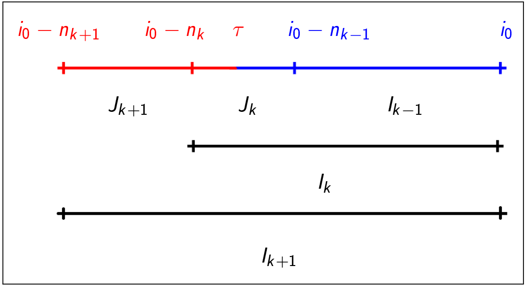
Fix i_0, $I_k = [i_0 - n_k, i_0]$, $n_k = \left[n_0 c^k \right]$, $c > 1$

$\{n_k\}_{k=0}^{14} = \{15 \text{ min.}, 19 \text{ min.}, \ldots, 1 \text{ day}\}$, $c = 1.25$

Local Change Point Detection ▸ Example

- Fix i_0, sequential test $(k = 1, \ldots, K)$

H_0 : parameter homogeneity within I_k vs. H_1 : $\exists$ change point within J_k



$$T_k = \sup_{\tau \in J_k} \left\{ L_{A_{k,\tau}}\left(\widetilde{\theta}_{A_{k,\tau}}\right) + L_{B_{k,\tau}}\left(\widetilde{\theta}_{B_{k,\tau}}\right) - L_{I_{k+1}}\left(\widetilde{\theta}_{I_{k+1}}\right) \right\},$$

with $J_k = I_k \setminus I_{k-1}$, $A_{k,\tau} = [i_0 - n_{k+1}, \tau]$ and $B_{k,\tau} = (\tau, i_0]$

Critical Values, $\mathfrak{z}_k$

▸ Critical Values

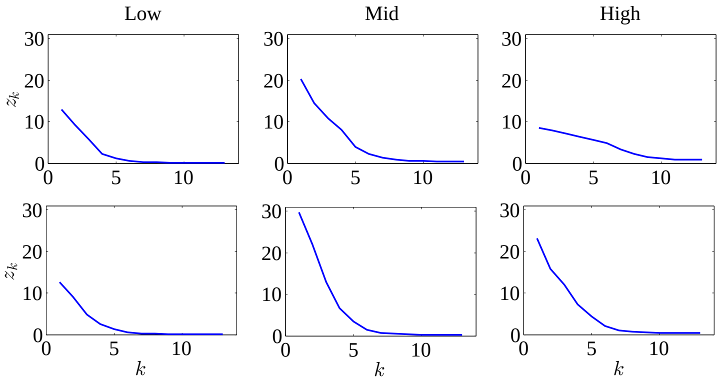


Figure 7: Simulated critical values of an EACD(1, 1) model and chosen parameter constellations according to Table 2. Volume series upper panel, order flow lower panel.

Adaptive Estimation

- Compare T_k at every step k with $\mathfrak{z}_k$
- Data window index of the *interval of homogeneity* - $\widehat{k}$
- Adaptive estimate

$$\widehat{\boldsymbol{\theta}} = \widetilde{\boldsymbol{\theta}}_{\widehat{k}}, \quad \widehat{k} = \max_{k \leq K} \{k : T_\ell \leq \mathfrak{z}_\ell, \ell \leq k\}$$

- Note: rejecting the null at $k = 1$, $\widehat{\theta}$ equals QMLE at I_0
  If the algorithm goes until K, $\widehat{\theta}$ equals QMLE at I_K

Data

- Osaka Securities Exchange
  - Series: mini Nikkei 225 index futures
  - NO12U 20120628-20120913, NO12Z 20120914-20121213, NO13H 20121214-20130307, NO13M 20130308-20130524
- Span: 20120628 - 20130524
  - Data on 20130304 removed as trading stopped between 11:06-14:10
  - 218 trading days (78480 minutes), 09:01-15:00

Notation

- One-minute cumulated volumes at day d and minute i
  - Buyer-initiated $\breve{y}^b_{i,d}$, seller-initiated $\breve{y}^s_{i,d}$
  - Order flow $\breve{y}^b_{i,d} - \breve{y}^s_{i,d}$, relative order flow $\breve{y}^b_{i,d}/(\breve{y}^b_{i,d} + \breve{y}^s_{i,d})$
- Seasonally adjusted volume Periodicity component $s_{i,d-1}$
  - Buyer-initiated volume, $y^b_{i,d} = \breve{y}^b_{i,d}/s^b_{i,d-1}$
  - Seller-initiated volume, $y^s_{i,d} = \breve{y}^s_{i,d}/s^s_{i,d-1}$
  - $d = 31, \ldots, 218$ (20120813 – 20130524), $i = 1, \ldots, 360$

Intraday Periodicity

Figure 8: Estimated intraday periodicity factors for the buyer-initiated and the seller-initiated trades for the 'mini Nikkei 225 index futures' traded at the Osaka Securities Exchange from 20120813-20130524

Adaptive Estimation

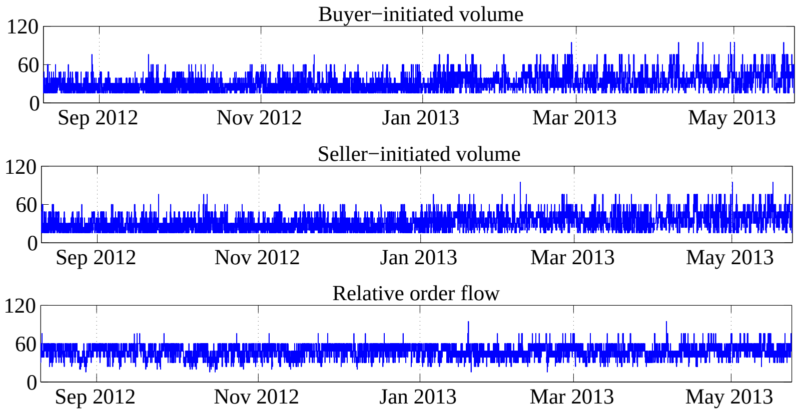


Figure 9: Estimated length of intervals of homogeneity (in minutes)

Order Flow Forecasting

Setup

- Forecasting period: 20120815 - 20130523 (186 days)
- Forecasts at each minute
  - Computed recursively
  - Multiplied by the seasonality component associated with the previous 30 days
- Predicted order flow: difference between predicted buyer and seller-initiated volume; predicted relative order flow

Intra-Day Trading

Transactions

- Profit or loss from 'trading' one futures contract in ¥
- New transaction - end of the current minute at current transaction price
- Offset transaction - after 5 min.

Results

- Calibration phase: 20120814-20121231
- Evaluation phase: 20130104-20130523

Trading Strategies

Strategies

(i) 'Buy if positive and sell if negative'

(ii) 'Buy if positive'

(iii) 'Sell if negative'

Strategies

- Compare: order flow with 0, relative order flow with 0.5
- Threshold - e.g., 95th percentile of the observed series

Trading Profit - Calibration Phase

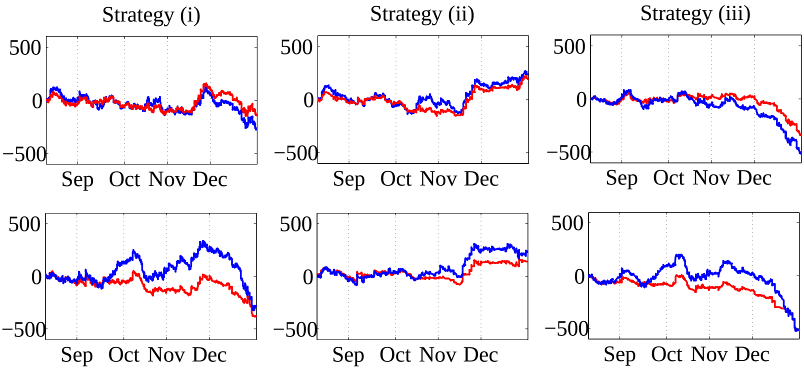


Figure 10: Cumulative profit in thousands ¥ per one contract between 20120814-20121231. Order flow (upper panel), relative order flow (lower panel), non-zero threshold in red.

Trading Profit - Evaluation Phase

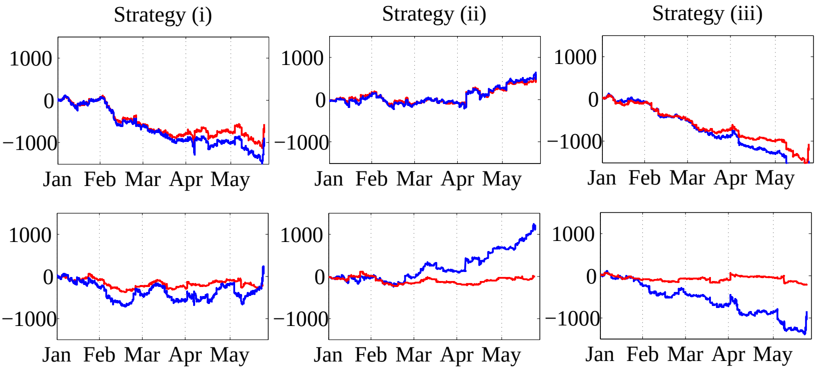


Figure 11: Cumulative profit in thousands ¥ per one contract between 20130104-20130523. Order flow (upper panel), relative order flow (lower panel), non-zero threshold in red.

Conclusions

Short-Term Order Flow Forecasting

- Order flow predicted successfully using the local adaptive MEMs
- Adaptive estimation requires up to 2 hours of data

Intra-Day Trading

- Best strategy: (ii) 'Buy if positive'
- Trading profit achieved, threshold selection

Adaptive Order Flow Forecasting with Multiplicative Error Models

Wolfgang Karl Härdle
Andrija Mihoci
Christopher Hian-Ann Ting

Ladislaus von Bortkiewicz Chair of Statistics
C.A.S.E. – Center for Applied Statistics
and Economics
Humboldt–Universität zu Berlin
Lee Kong Chian School of Business
Singapore Management University
http://lvb.wiwi.hu-berlin.de
http://www.case.hu-berlin.de
http://business.smu.edu.sg

References

Chen, Y. and Härdle, W. and Pigorsch, U.
Localized Realized Volatility
Journal of the American Statistical Association **105**(492): 1376–1393, 2010

Čížek, P., Härdle, W. and Spokoiny, V.
Adaptive Pointwise Estimation in Time-Inhomogeneous Conditional Heteroscedasticity Models
Econometrics Journal **12**: 248–271, 2009

Engle, R. F.
New Frontiers for ARCH Models
Journal of Applied Econometrics **17**: 425–446, 2002

References

Engle, R. F. and Rangel, J. G.
The Spline-GARCH Model for Low-Frequency Volatility and Its Global Macroeconomic Causes
Review of Financial Studies **21**: 1187–1222, 2008

Engle, R. F. and Russell, J. R.
Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data
Econometrica **66**(5): 1127–1162, 1998

Gallant, A. R.
On the bias of flexible functional forms and an essentially unbiased form
Journal of Econometrics **15**: 211–245, 1981

References

Härdle, W. K., Hautsch, N. and Mihoci, A.
Local Adaptive Multiplicative Error Models for High-Frequency Forecasts
Journal of Applied Econometrics, 2014

Hautsch, N.
Econometrics of Financial High-Frequency Data
Springer, Berlin, 2012

Mercurio, D. and Spokoiny, V.
Statistical inference for time-inhomogeneous volatility models
The Annals of Statistics **32**(2): 577–602, 2004

References

Spokoiny, V.
Estimation of a function with discontinuities via local polynomial fit with an adaptive window choice
The Annals of Statistics **26**(4): 1356–1378, 1998

Spokoiny, V.
Multiscale Local Change Point Detection with Applications to Value-at-Risk
The Annals of Statistics **37**(3): 1405–1436, 2009

Volume Description

▸ Trading Volume

- Buyer-initiated volume at minute i
  - Consider market orders only
  - Number of contracts that have been bought during minute i (at the ask price)

- Seller-initiated volume at minute i
  - Consider market orders only
  - Number of contracts that have been sold during minute i (at the bid price)

Exponential-ACD (EACD)

▸ ACD ▸ Parameter Estimation

- Engle and Russel (1998), $\varepsilon_i \sim Exp(1)$

$$L_I(y;\boldsymbol{\theta_E}) = \sum_{i=max(p,q)+1}^{n} \left(-\log\mu_i - \frac{y_i}{\mu_i}\right) \mathbf{I}\{i \in I\} \qquad (2)$$

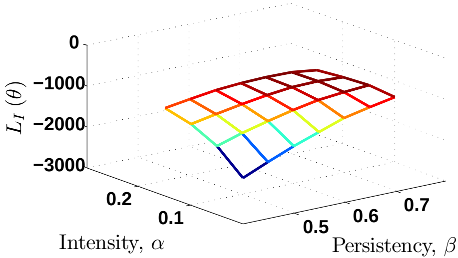


Figure 12: Log likelihood - EACD(1,1), $\boldsymbol{\theta_E^*} = (0.10, 0.20, 0.65)^\top$

Weibull-ACD (WACD)

▸ ACD ▸ Parameter Estimation

- Engle and Russel (1998), $\varepsilon_i \sim \mathcal{G}(s, 1)$

$$L_I(y; \boldsymbol{\theta}_W) = \sum_{i \in I} \left[\log \frac{s}{y_i} + s \log \frac{\Gamma(1 + 1/s)\, y_i}{\mu_i} - \left\{ \frac{\Gamma(1 + 1/s)\, y_i}{\mu_i} \right\}^s \right] \mathbf{I}\{i \in I\} \tag{3}$$

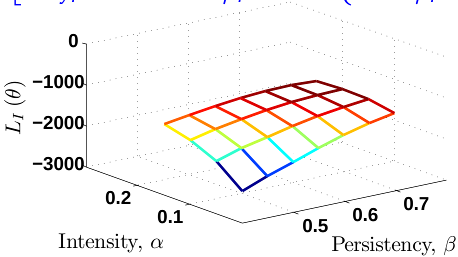


Figure 13: Log likelihood - WACD(1,1), $\boldsymbol{\theta}_W^* = (0.10, 0.20, 0.65, 0.85)^\top$

Gaussian Regression

▸ Estimation Quality

$Y_i = f(X_i) + \varepsilon_i, i = 1, \ldots, n$, weights $W = \{w_i\}_{i=1}^{n}$

$L(W, \theta) = \sum_{i=1}^{n} \ell\{Y_i, f_\theta(X_i)\} w_i$, log-density $\ell(\cdot)$, $\widetilde{\theta} = \arg\max_{\theta \in \Theta} L(W, \theta)$

1. Local constant, $f(X_i) \approx \theta^*$, $\varepsilon_i \sim \mathrm{N}(0, \sigma^2)$

$$\mathrm{E}_{\theta^*} \left| L(W, \widetilde{\theta}) - L(W, \theta^*) \right|^r \leq \mathrm{E}\,|\xi|^{2r}, \quad \xi \sim \mathrm{N}(0, 1)$$

2. Local linear, $f(X_i) \approx \theta^{*\top} \Psi_i$, $\varepsilon_i \sim \mathrm{N}(0, \sigma^2)$, basis functions $\Psi = \{\psi_1(X_1), \ldots, \psi_p(X_p)\}$, multivariate ξ

$$\mathrm{E}_{\theta^*} \left| L(W, \widetilde{\theta}) - L(W, \theta^*) \right|^r \leq \mathrm{E}\,|\xi|^{2r}, \quad \xi \sim \mathrm{N}(0, \mathcal{I}_p)$$

Local Change Point Detection ▸ LCP

Example: Trading volumes aggregated over 1-min periods

- Scheme with $(K+1) = 14$ intervals and fix i_0
- Assume $I_0 = 60$min. is homogeneous

- H_0 : parameter homogeneity within $I_1 = 75$min.
  - Define $J_1 = I_1 \setminus I_0$ - observations from y_{i_0-75} up to y_{i_0-60}
  - For each $\tau \in J_1$ fit log likelihoods over $A_{1,\tau}$, $B_{1,\tau}$ and I_2
  - Find the largest likelihood ratio - T_{I_1,J_1}

Critical Values, $\mathfrak{z}_k$

▸ Critical Values

- Simulate $\mathfrak{z}_k$ - homogeneity of the interval sequence $I_0, \ldots, I_k$
- 'Propagation' condition (under H_0)

$$\mathsf{E}_{\theta^*} \left| L_{I_k}(\widetilde{\theta}_k) - L_{I_k}(\widehat{\theta}_k) \right|^r \leq \rho_k \mathcal{R}_r(\theta^*), \quad k = 1, \ldots, K \qquad (4)$$

$\rho_k = \rho k / K$ for given significance level ρ, ▸ $\widehat{\theta}_k$ - adaptive estimate

- Check $\mathfrak{z}_k$ for (nine) different θ^* ▸ Parameter Dynamics - Quartiles
  - EACD and WACD, $K \in \{8, 13\}$, $r \in \{0.5, 1\}$, $\rho \in \{0.25, 0.50\}$
  - Findings: $\mathfrak{z}_k$ are virtually invariable w.r.t. θ^* given a scenario
    Largest differences at first two or three steps

Parameter Dynamics

Estimation window	EACD(1,1)			WACD(1,1)		
	Q25	Q50	Q75	Q25	Q50	Q75
1 week	0.85	0.89	0.93	0.82	0.88	0.92
2 days	0.77	0.86	0.92	0.74	0.84	0.91
1 day	0.68	0.82	0.90	0.63	0.79	0.89
3 hours	0.54	0.75	0.88	0.50	0.72	0.87
2 hours	0.45	0.70	0.86	0.42	0.67	0.85
1 hour	0.33	0.58	0.80	0.31	0.57	0.80

Table 1: Quartiles of estimated persistence levels $\left(\widetilde{\alpha} + \widetilde{\beta}\right)$ for all five stocks at each minute from 20080222-20081231 (215 trading days). Calibration period: 20080102 - 20080221, Härdle et al. (2014)

Parameter Dynamics

▸ Critical Values

Model	Buyer-initiated			Seller-initiated			Order Flow		
	Low	Mid	High	Low	Mid	High	Low	Mid	High
$\widetilde{\omega}$	0.10	0.22	0.46	0.10	0.21	0.38	0.17	0.23	0.32
$\widetilde{\alpha}$	0.11	0.16	0.19	0.12	0.15	0.17	0.12	0.18	0.22
$\widetilde{\beta}$	0.45	0.63	0.73	0.52	0.66	0.74	0.24	0.36	0.45
$\widetilde{\alpha}+\widetilde{\beta}$	0.56	0.79	0.92	0.64	0.81	0.91	0.36	0.54	0.67

Table 2: Quartiles of estimated MEM parameters based on an estimation window covering 360 observations from 14 August 2012 to 24 May 2013. We label the first quartile as 'low', the second quartile as 'mid' and the third quartile as 'high'.

Intra-day Periodicity

▸ Notation

- Flexible Fourier Series (FFS) approximation, Gallant (1981)
  Intraday periodicity components $\left(s_{1,d-30}, \dots, s_{360,d-1}\right)^\top$

- Estimation: 30-day rolling window with $s_{i,d-1} = \dots = s_{i,d-30}$, Engle and Rangel (2008)

$$s_{i,d-1} = \delta \bar{\imath}_i + \sum_{m=1}^{M} \{\delta_{c,m} \cos(\bar{\imath}_i \cdot 2\pi m) + \delta_{s,m} \sin(\bar{\imath}_i \cdot 2\pi m)\}$$

$\bar{\imath} = (\bar{\imath}_1, \dots, \bar{\imath}_{360})^\top = (1/360, \dots, 360/360)^\top$ - intraday time trend