Local Adaptive Multiplicative Error Models for High-Frequency Forecasts

Wolfgang Karl Härdle
Nikolaus Hautsch
Andrija Mihoci

Ladislaus von Bortkiewicz Chair of Statistics
Chair of Econometrics
C.A.S.E. – Center for Applied Statistics and Economics
Humboldt–Universität zu Berlin
http://lvb.wiwi.hu-berlin.de
http://www2.hu-berlin.de/oekonometrie/
http://case.hu-berlin.de

Trading Volume

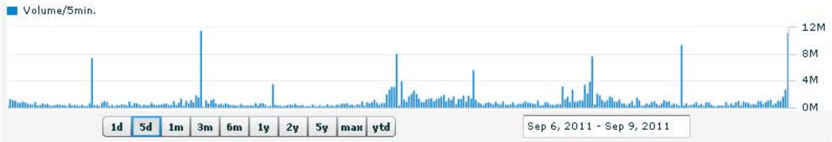


Figure 1: Five-minute cumulated trading volume for Intel Corporation (INTC) at NASDAQ

Statistical Challenges

- High-frequency dynamics subject to regime shifts
  - Time-varying parameters ▸ Parameter Dynamics
  - Structural breaks
- Modelling using procrustrean assumptions
  - Fixed local estimation window and overparametrisation
  - Transition form, number of regimes, transition variable type

Objectives

▸ Economic Benefits

(i) Localising Multiplicative Error Models (MEM)

- Local parametric approach (LPA)
- Balance between modelling bias and parameter variability
- Estimation windows with potentially varying lengths

(ii) Forecasting trading volumes

- Rolling window out-of-sample forecasting exercise
- Evaluation against standard approach: fixed estimation length on an ad hoc basis

Example

Example: Short-term forecasting

An investor decides to forecast one-minute cumulated trading volume for INTC up to the next one hour

Forecasting strategies

(i) 'Standard' method - fixed estimation window (one week)

(ii) LPA technique with adaptively selected interval of homogeneity

Outline

1. Motivation ✓
2. Multiplicative Error Models (MEM)
3. Local Parametric Approach
4. Empirical Results
5. Conclusions

Multiplicative Error Models (MEM)

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

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

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

- 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:

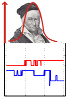
Parametric Modelling

1. Exponential-ACD, Engle and Russel (1998) ‣ EACD
   $\varepsilon_i \sim Exp\,(1)$, $\boldsymbol{\theta}_E = (\omega, \boldsymbol{\alpha}, \boldsymbol{\beta})^\top$, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_p)^\top$, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_q)^\top$
2. Weibull-ACD, Engle and Russel (1998) ‣ WACD
   $\varepsilon_i \sim \mathcal{G}\,(s, 1)$, $\boldsymbol{\theta}_W = (\omega, \boldsymbol{\alpha}, \boldsymbol{\beta}, s)^\top$

⊡ Quasi maximum likelihood estimates (QMLEs) of $\boldsymbol{\theta}_E$ and $\boldsymbol{\theta}_W$

$$\widetilde{\boldsymbol{\theta}}_I = \arg\max_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} L_I\,(y; \boldsymbol{\theta}) \qquad (1)$$

- Data interval (fix i_0, length n) $I = [i_0 - n, i_0]$
- Quasi log likelihood $L_I\,(\cdot)$, see (2) and (3)

Weibull, E. H. Waloddi on BBI: 

Parameter Dynamics

Statistical Challenges

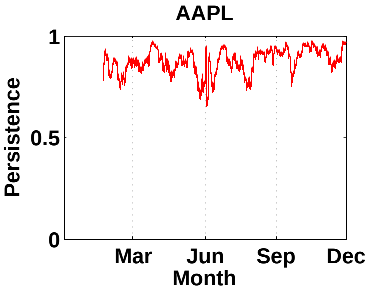


Figure 2: Estimated weekly $(n = 1800)$ and daily $(n = 360)$ persistence $\widetilde{\alpha}_i + \widetilde{\beta}_i$ for seasonally adjusted trading volume using an EACD$(1,1)$ at each minute in 2008

Estimation Quality

- Quality of estimating $\boldsymbol{\theta}^*$ by QMLE $\widetilde{\boldsymbol{\theta}}_I$ in terms of Kullback-Leibler divergence; $\mathcal{R}_r(\boldsymbol{\theta}^*)$ - risk bound

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

'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)

- Volatility modelling - Mercurio and Spokoiny (2004)
- GARCH$(1, 1)$ models - Čížek et al. (2009)
- Realized volatility - Chen et al. (2010)

- LPA: time series parameters can be locally approximated
- Balance between modelling bias and parameter variability

Statistical Framework

- LPA: Spokoiny (1998, 2009)

- *Interval of homogeneity* - balance between bias and variability
- Small modelling bias (SMB) - basis for the 'oracle' interval

- Adaptive estimate $\widehat{\theta}$ - QMLE at the *interval of homogeneity*
- Risk of $\widehat{\theta}$ is within a close range of the minimal oracle risk

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

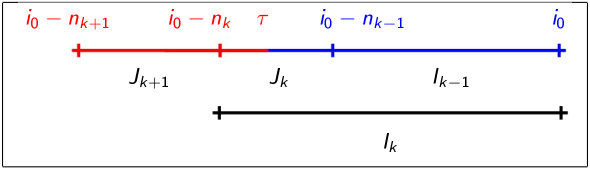
Fix i_0, $I_k = [i_0 - n_k, i_0]$, $n_k = \left[n_0 c^k\right]$, $c > 1$

$\{n_k\}_{k=0}^{13} = \{60 \text{ min.}, 75 \text{ min.}, \ldots, 1 \text{ week}\}$, $c = 1.25$

Local Change Point Detection ▸ Example

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

H_0 : parameter homogeneity within I_k vs. H_1 : change point within I_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, $\mathfrak{z}_k$

- Simulate $\mathfrak{z}_k$ under the null of the homogeneity of the interval sequence $I_0, \ldots, I_k$
- Check $\mathfrak{z}_k$ for different θ^*
  - Nine different parameters θ^* for each model (EACD and WACD) and risk level ($r = 0.5$ and $r = 1$)
  - Findings: $\mathfrak{z}_k$ are virtually invariable w.r.t. θ^* given a scenario Largest differences at first two or three steps

Critical Values, $\mathfrak{z}_k$

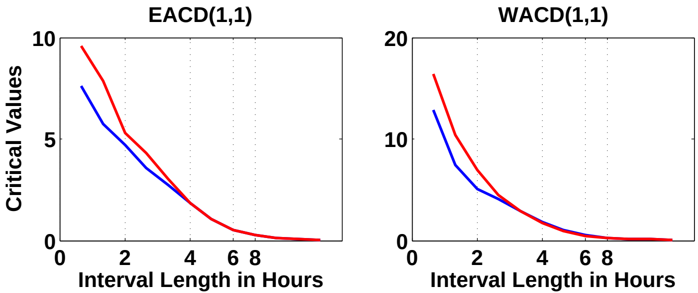


Figure 3: Critical values for low $(\widetilde{\alpha}+\widetilde{\beta}=0.84)$ and high $(\widetilde{\alpha}+\widetilde{\beta}=0.93)$ weekly persistence. Modest risk $(r=0.5)$ and median ratio $\widetilde{\beta}/\left(\widetilde{\alpha}+\widetilde{\beta}\right)$.

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

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

- 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

- NASDAQ Stock Market in 2008, 250 trading days
  - 5 stocks: AAPL, CSCO, INTC, MSFT and ORCL
  - $\breve{y}_i$ - one-minute cumulated trading volume from 10:00-16:00
  - y_i - seasonally adjusted trading volume
- Periodicity effect - FFS approximation, Gallant (1981)
  - 30-days rolling windows, Engle and Rangel (2008)
  - Order $M = 1, \ldots, 6$ selected by BIC, $\bar{\imath} \in [0, 1]$

$$y_i = \breve{y}_i / [\delta \cdot \bar{\imath} + \sum_{m=1}^{M} \{\delta_{c,m} \cos(\bar{\imath} \cdot 2\pi m) + \delta_{s,m} \sin(\bar{\imath} \cdot 2\pi m)\}]$$

Intraday Periodicity

Figure 4: Estimated intraday periodicity components for AAPL, $M = 6$

Adaptive Estimation

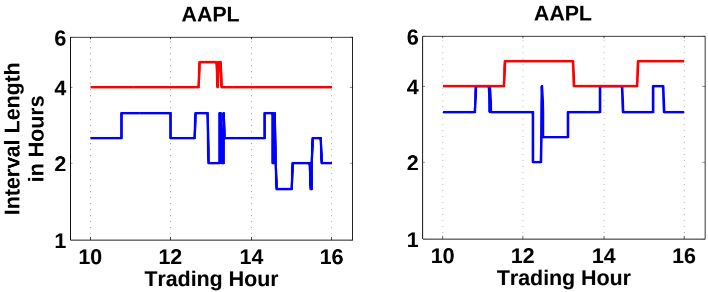


Figure 5: Estimated interval length $n_{\hat{k}}$ for the conservative $(r = 1)$ and modest $(r = 0.5)$ risk case on 20080902 (left, lowest volume) and 20081030 (right, highest volume) using the EACD model

Adaptive Estimation

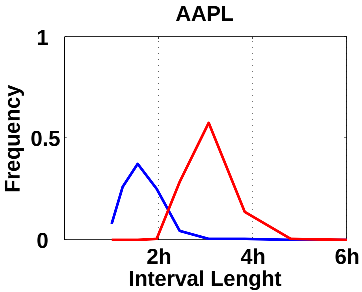


Figure 6: Estimated interval length $n_{\hat{k}}$ for the conservative $(r = 1)$ and modest $(r = 0.5)$ risk case in 2008 using the EACD (left) and WACD (right) models

Forecasting

Strategies

- LPA technique with adaptively selected interval of homogeneity
- 'Standard' method: 360 (1 day) or 1800 observations (1 week)

Setup

- Forecasting period: 20080222 - 20081222 (210 trading days)
- Forecasts at each minute (horizon $h = 1, \ldots, 60$ min.)

(a) Sign test, (b) Diebold-Mariano test, (c) Forecasting regressions

Forecasting Accuracy

1. *Overall performance* - LPA qualitatively outperforms 'standard' methods, quantitatively equal to 1-day 'standard' method
2. *Forecasting horizon* - performance best at approx. 3-4 minutes
3. *Course of a typical trading day* - best results after 14:00
4. *Model specifications and tuning parameters* - WACD slightly better, insensitivity

Forecasting Accuracy

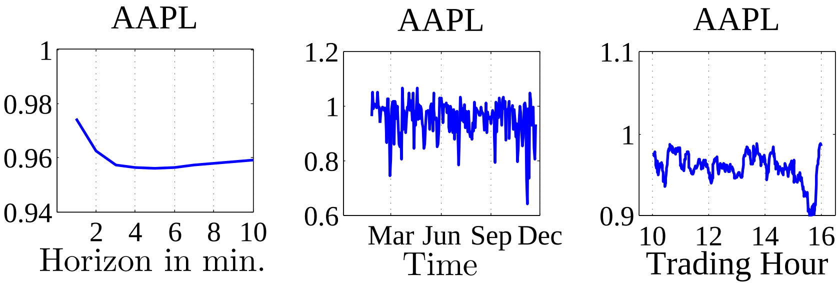


Figure 7: Ratio between the RMSPE of the LPA method and the RMSPE of the standard approach using an EACD(1, 1) model for Apple Inc. (AAPL) from 20100222 to 20101222 (210 trading days)

Forecasting Accuracy

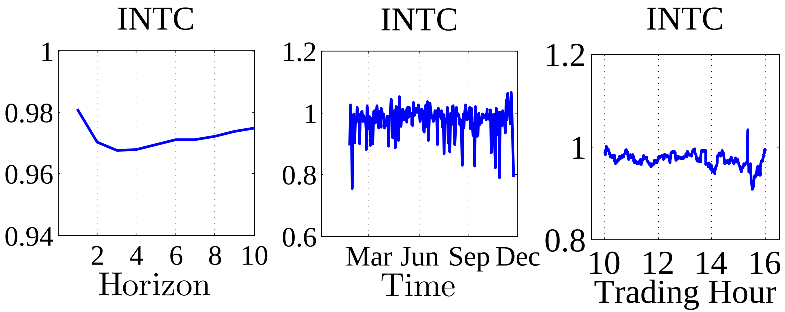


Figure 8: Ratio between the RMSPE of the LPA method and the RMSPE of the standard approach using an EACD$(1,1)$ model for Intel Corporation (INTC) from 20100222 to 20101222 (210 trading days)

Conclusions

Localising MEM

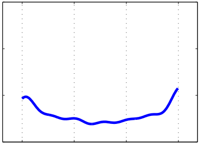
- Time-varying parameters and estimation quality
- LPA - 5 stocks in 2008 (79200 minutes): AAPL, CSCO, INTC, MSFT and ORCL
- Precise adaptive estimation ($r = 1$) requires 4-5 hours of data, modest risk approach ($r = 0.5$) requires 2-3 hours

Forecasting Trading Volume

- LPA outperforms the 'standard' method
- Overall performance, horizon, trading day, tuning parameters

Local Adaptive Multiplicative Error Models for High-Frequency Forecasts

Wolfgang Karl Härdle
Nikolaus Hautsch
Andrija Mihoci



Ladislaus von Bortkiewicz Chair of Statistics
Chair of Econometrics
C.A.S.E. – Center for Applied Statistics and Economics
Humboldt–Universität zu Berlin
http://lvb.wiwi.hu-berlin.de
http://www2.hu-berlin.de/oekonometrie/
http://case.hu-berlin.de

References

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

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

References

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

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

Economic Benefits ▸ Objectives

Flexible statistical framework

- MEM parameter dynamics: strength and frequency
- Typical interval lengths of parameter homogeneity
- Tuning parameters (sensitivity)
- Short-term forecasting

Exponential-ACD (EACD) ▸ Parametric Modelling

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

$$L_I\,(\boldsymbol{\theta}) = \sum_{t \in I} \left(-\log \mu_t - \frac{y_t}{\mu_t} \right) \tag{2}$$

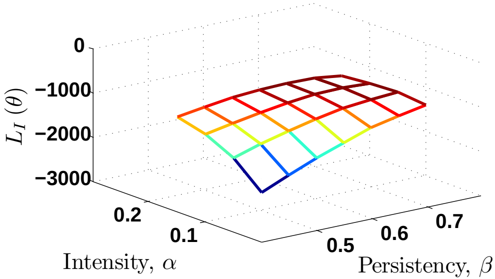


Figure 9: Quasi log likelihood - EACD(1,1), $\boldsymbol{\theta}^* = (0.10, 0.20, 0.65)^\top$

Weibull-ACD (WACD)

Parametric Modelling

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

$$L_I(\theta) = \sum_{t \in I} \left[\log \frac{s}{y_t} + s \log \frac{\Gamma(1 + 1/s)\, y_t}{\mu_t} - \left\{ \frac{\Gamma(1 + 1/s)\, y_t}{\mu_t} \right\}^s \right] \qquad (3)$$

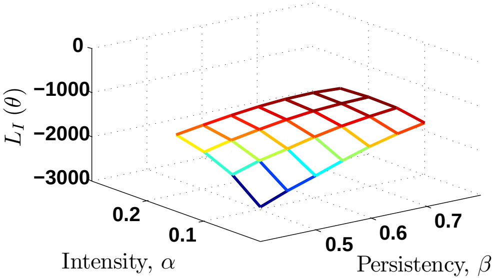


Figure 10: Quasi log likelihood - WACD(1,1), $\theta^* = (0.10, 0.20, 0.65, 0.85)^\top$

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}