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

Statistical Challenges

- Understanding high-frequency dynamics
  - Time-varying parameters [Parameter Dynamics]
  - Regime shifts
- Modelling using procrustrean assumptions
  - Time-invariant parameters
  - Transition form, number of regimes, transition variable type

Objectives

(i) Localising Multiplicative Error Models (MEM)

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

(ii) Short-term forecasting

- Case study: trading volume
- Evaluation against standard approach - fixed estimation length on an ad hoc basis

Example

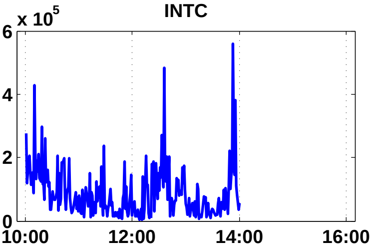


Figure 1: One-minute cumulated trading volume for Intel Corporation (INTC) on 20080902

Example

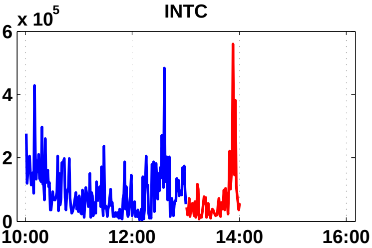


Figure 2: One-minute cumulated trading volume for Intel Corporation (INTC) on 20080902 with an estimation window length of 60

Example

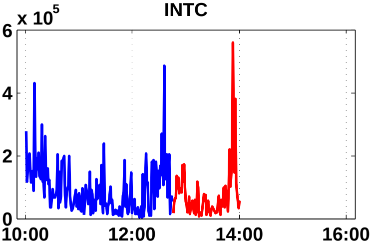


Figure 3: One-minute cumulated trading volume for Intel Corporation (INTC) on 20080902 with an estimation window length of 75

Example

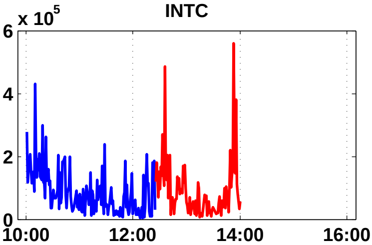


Figure 4: One-minute cumulated trading volume for Intel Corporation (INTC) on 20080902 with an estimation window length of 95

Example

Example: Short-term forecasting

Forecasting strategies up to the next one hour:

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

(ii) LPA technique with adaptively selected interval of homogeneity

Research Questions

- How strong is the variation of MEM parameters over time?
- What are typical interval lengths of parameter homogeneity?
- How good are LPA short-term forecasts relative to procedures with ad hoc fixed estimation windows?

Outline

1. Motivation ✓
2. Multiplicative Error Models (MEM)
3. Local Parametric Approach (LPA)
4. Forecasting Trading Volumes
5. Conclusions

Multiplicative Error Models (MEM)

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

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

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

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

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

  - $I = [i_0 - n, i_0]$ - interval of $(n+1)$ observations at i_0
  - $L_I(\cdot)$ - log likelihood, see (7) for EACD and (8) for WACD

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 window, Engle and Rangel (2008)

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

$\bar{\imath} \in (0, 1]$ - number of minutes from opening until i relative to 360

Intraday Periodicity

Figure 5: Estimated intraday periodicity components for AAPL, order $M = 6$ selected by BIC

Parameter Dynamics

▸ Statistical Challenges

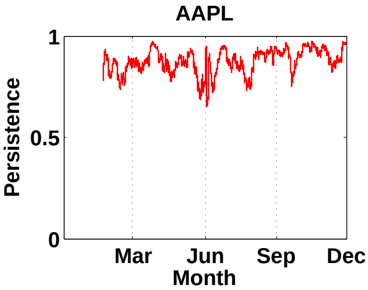


Figure 6: Estimated weekly $(n = 1800)$ and daily $(n = 360)$ persistence $\widetilde{\alpha}_i + \widetilde{\beta}_i$ for seasonally adjusted trading volume using an EACD$(1, 1)$ at each minute in 2008

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 22 February to 31 December 2008 (215 trading days).

Parameter Dynamics

Model	Low Persistence			Moderate Persistence			High Persistence		
	Q25	Q50	Q75	Q25	Q50	Q75	Q25	Q50	Q75
EACD, $\widetilde{\alpha}$	0.28	0.22	0.18	0.30	0.23	0.19	0.31	0.24	0.20
EACD, $\widetilde{\beta}$	0.56	0.62	0.67	0.59	0.66	0.71	0.62	0.68	0.73
WACD, $\widetilde{\alpha}$	0.28	0.21	0.17	0.30	0.23	0.18	0.32	0.24	0.19
WACD, $\widetilde{\beta}$	0.54	0.60	0.65	0.58	0.65	0.70	0.60	0.68	0.74

Table 2: Quartiles of 774,000 estimated ratios $\widetilde{\beta}/\left(\widetilde{\alpha}+\widetilde{\beta}\right)$ (estimation windows covering 1800 observations) for all five stocks at each minute from 22 February to 31 December 2008 (215 trading days) conditional on the persistence level (low, moderate or high)

Parameter Dynamics - Summary

- MEM parameters, their variability and distribution properties change over time
- Longer local estimation windows increase estimation precision and the misspecification risk
- Tradeoff between estimation (in)efficiency and local flexibility

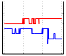
Estimation Quality

- Mercurio and Spokoiny (2004), Spokoiny (2009)
- 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)

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

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

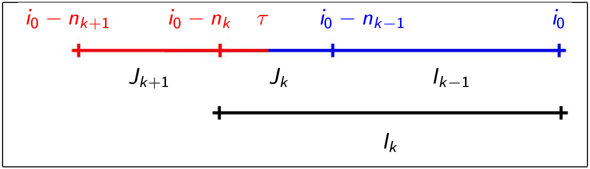
Fix i_0, $I_k = [i_0 - n_k, i_0]$, $n_k = \left[n_0 c^k\right]$, $c > 1$

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

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 $\mathfrak{z}_k$ - homogeneity of the interval sequence $I_0, \ldots, I_k$
- 'Propagation' conditions

$$\mathrm{E}_{\theta^*} \left| L_{I_k}(\widetilde{\theta}_{I_k}) - L_{I_k}(\widehat{\theta}_{I_k}) \right|^r \leq \rho_k \mathcal{R}_r(\theta^*), \quad k = 1, \ldots, K \quad (2)$$

$\rho_k = \rho k/K$ for given significance level ρ

- Check $\mathfrak{z}_k$ for (nine) different θ^*
  - 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

Critical Values, $\mathfrak{z}_k$

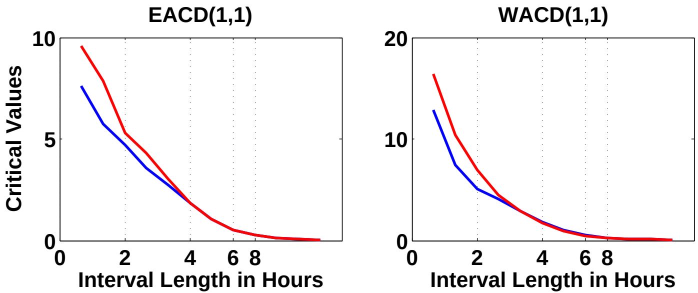


Figure 7: Critical values for low $(\widetilde{\alpha}+\widetilde{\beta}=0.84)$ and high $(\widetilde{\alpha}+\widetilde{\beta}=0.93)$ weekly persistence and 'modest' risk $(r=0.5)$ with $\rho=0.25$

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

Adaptive Estimation - Results

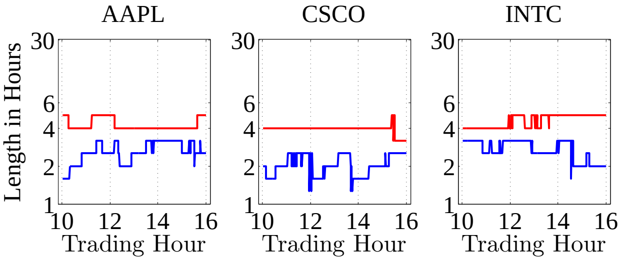


Figure 8: Estimated length $n_{\hat{k}}$ of *intervals of homogeneity* given the conservative $(r = 1)$ and modest $(r = 0.5)$ risk case on 20080222 using the EACD$(1,1)$ model with $\rho = 0.25$

Adaptive Estimation - Results

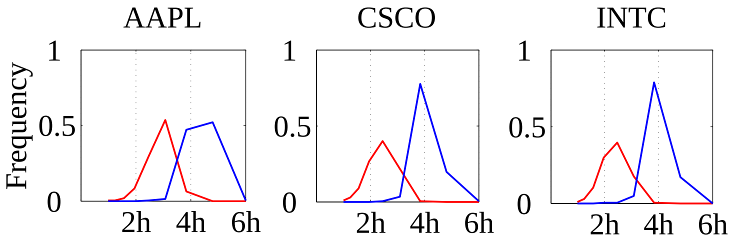


Figure 9: Distribution of estimated interval length $n_{\hat{k}}$ (in hours) given the conservative ($r = 1$) and modest ($r = 0.5$) risk case from 20080222 to 20081231 using the EACD(1, 1) model with $\rho = 0.25$

Forecasting Trading Volumes

Setup

- 5 stocks, forecasting period: 20080222 - 20081222 (210 days)
- Forecasts at each minute (horizon $h = 1, \dots, 60$ min.)
- EACD(1, 1) and WACD(1, 1) model, $r \in \{0.5, 1\}$, $\rho \in \{0.25, 0.5\}$

Strategies

- LPA technique - prediction $\widehat{y}_{i+h}$, error $\widehat{\varepsilon}_{i+h} = \breve{y}_{i+h} - \widehat{y}_{i+h}$
- 'Standard' method: 360 (1 day) or 1800 observations (1 week) - prediction $\widetilde{y}_{i+h}$, error $\widetilde{\varepsilon}_{i+h} = \breve{y}_{i+h} - \widetilde{y}_{i+h}$

Forecasting Superiority

- Diebold and Mariano (1995) tests, loss differential

$$d_h = \{d_{i+h}\}_{i=1}^{n} = \widetilde{\varepsilon}^2_{i+h} - \widehat{\varepsilon}^2_{i+h} \tag{3}$$

- Ratio of root mean squared errors

$$\sqrt{n^{-1}\sum_{i=1}^{n}\widehat{\varepsilon}^2_{i+h}} \Big/ \sqrt{n^{-1}\sum_{i=1}^{n}\widetilde{\varepsilon}^2_{i+h}} \tag{4}$$

Forecasting Superiority

- Qualitative test (qualitatively lower prediction errors)

$$T_{ST,h} = \left\{ \sum_{i=1}^{n} \mathrm{I}\left(d_{i+h} > 0\right) - 0.5n \right\} / \sqrt{0.25n} \xrightarrow{\mathcal{L}} \mathrm{N}(0,1) \tag{5}$$

- Quantitative test, $H_0 : \mathrm{E}\left[d_h\right] = 0$

$$T_{DM,h} = \bar{d}_h / \sqrt{2\pi \widehat{f}_{d_h}\left(0\right)/n} \xrightarrow{\mathcal{L}} \mathrm{N}(0,1) \tag{6}$$

$\bar{d}_h = n^{-1} \sum_{i=1}^{n} d_{i+h}$, $\widehat{f}_{d_h}\left(0\right)$ - spectral density estimate at frequency zero

Forecasting Superiority

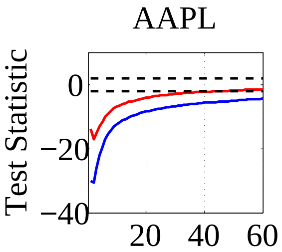


Figure 10: Test statistic $T_{DM,h}$ across all 60 forecasting horizons from 20100222 to 20101222 (210 trading days): LPA against a fixed-window scheme using 360 and 1800 observations using the EACD(1,1) model, $r = 0.5$ and $\rho = 0.25$

Forecasting Superiority

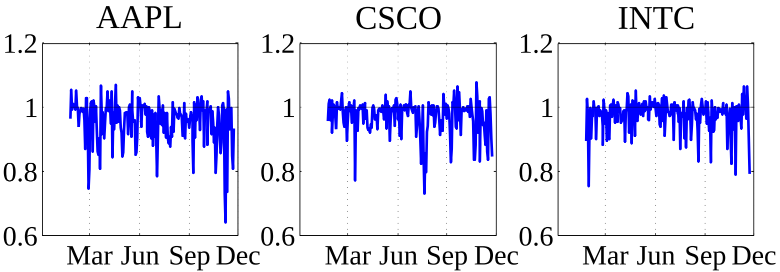


Figure 11: Ratio between the RMSPEs of the LPA and of a fixed-window approach (6 trading hours) over the sample from 20100222 to 20101222 (210 trading days) using an EACD(1, 1) model, $r = 0.5$ and $\rho = 0.25$

Forecasting Superiority

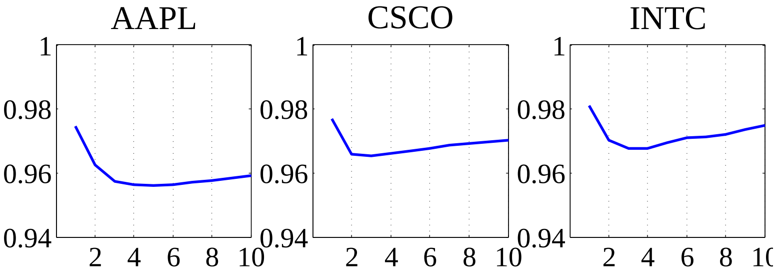


Figure 12: Ratio between the RMSPEs of the LPA and of a fixed-window approach (6 trading hours) across horizon from 20100222 to 20101222 (210 trading days) using an EACD(1,1) model, $r = 0.5$ and $\rho = 0.25$

Forecasting Superiority - Summary

1. *Overall performance* - LPA qualitatively and quantitatively outperforms 'standard' methods
2. *Forecasting horizon* - overperformance best at short horizons (approx. 3-4 minutes)
3. *Sample* - excellent during market distress
4. *Model specifications and tuning parameters* - robust results

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

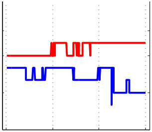
Forecasting Trading Volumes

- 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

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

Diebold, F. and Mariano, R. S.
Comparing Predictive Accuracy
Journal of Business and Economic Statistics **13**(3): 253–263, 1995

References

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

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

References

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

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

Exponential-ACD (EACD)

Parametric Modelling

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

$$L_I\,(y;\boldsymbol{\theta}_E) = \sum_{i=max(p,q)+1}^{n} \left(-\log\mu_i - \frac{y_i}{\mu_i}\right) \mathrm{I}\,\{i \in I\} \tag{7}$$

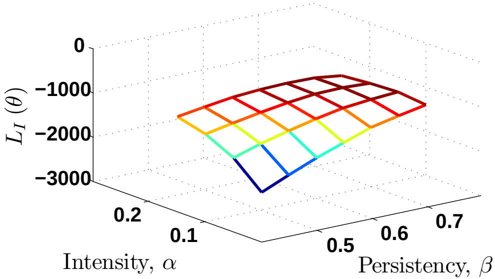


Figure 13: Log likelihood - EACD(1,1), $\boldsymbol{\theta}_E^* = (0.10, 0.20, 0.65)^\top$

Weibull-ACD (WACD) ▸ Parametric Modelling

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

$$L_I(y; \theta_W) = \sum_{i \in I} \left[\log \frac{s}{y_i} + s \log \frac{\Gamma(1 + 1/s)\, y_i}{\mu_i} - \left\{ \frac{\Gamma(1 + 1/s)\, y_i}{\mu_i} \right\}^s \right] \mathbf{I}\{i \in I\} \tag{8}$$

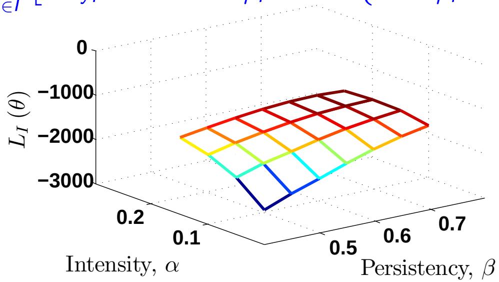


Figure 14: Log likelihood - WACD(1,1), $\theta_W^* = (0.10, 0.20, 0.65, 0.85)^\top$

Local Change Point Detection ▸ LCP

Example: Trading volumes aggregated over 1-min periods

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

- H_0 : parameter homogeneity within $I_1 = 75\text{min.}$
  - 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}