

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 Procrustean 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

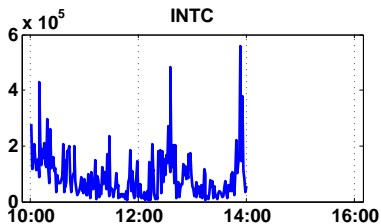


Figure 1: One-minute cumulated trading volume for Intel Corporation (INTC) on 20080902



Example

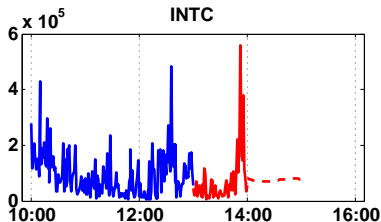


Figure 2: One-minute cumulated trading volume for Intel Corporation (INTC) on 20080902 with an **EACD(1,1)** estimation window length of **60** with volume forecasts to the next one hour (dashed)



Example

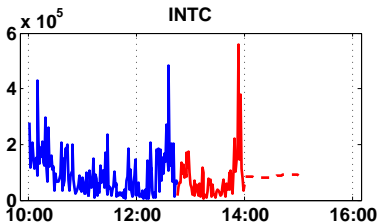


Figure 3: One-minute cumulated trading volume for Intel Corporation (INTC) on 20080902 with an **EACD(1,1)** estimation window length of **75** with volume forecasts to the next one hour (dashed)



Example

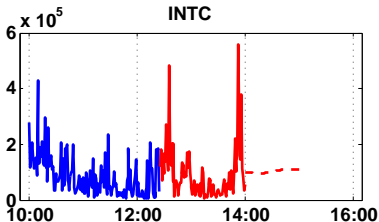


Figure 4: One-minute cumulated trading volume for Intel Corporation (INTC) on 20080902 with an **EACD(1,1)** estimation window length of **95** with volume forecasts up to the next one hour (dashed)



Research Questions

▶ Example

- 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?



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

$$y_i = \mu_i \varepsilon_i, \quad E[\varepsilon_i | \mathcal{F}_{i-1}] = 1$$
$$\mu_i = \omega + \sum_{j=1}^p \alpha_j y_{i-j} + \sum_{j=1}^q \beta_j \mu_{i-j}, \quad \omega > 0, \alpha_j, \beta_j \geq 0$$

- Hautsch (2012) - comprehensive MEM literature overview
 - y_i - squared (de-means) 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

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

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 θ_E and θ_W

$$\tilde{\theta}_I = \arg \max_{\theta \in \Theta} L_I(y; \theta) \quad (1)$$

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



Data

- NASDAQ Stock Market in 2008, 250 trading days
 - ▶ 5 stocks: AAPL, CSCO, INTC, MSFT and ORCL
 - ▶ \check{y}_i - one-minute cumulated trading volume from 10:00-16:00
 - ▶ y_i - seasonally adjusted trading volume

- Periodicity effect - FFS approximation, Gallant (1981)
 - ▶ Engle and Rangel (2008) - 30-days rolling window

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

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



Intraday Periodicity

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



Parameter Dynamics

▶ Statistical Challenges

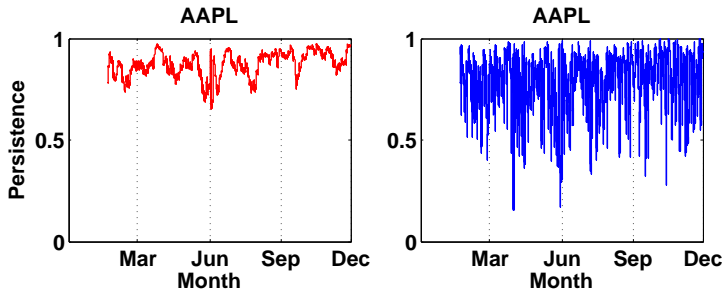


Figure 6: Estimated **weekly** ($n = 1800$) and **daily** ($n = 360$) persistence $\tilde{\alpha}_i + \tilde{\beta}_i$ for seasonally adjusted trading volume using an EACD(1, 1) at each minute in 2008



Estimation Quality

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

$$E_{\theta^*} \left| L_I(\tilde{\theta}_I) - L_I(\theta^*) \right|^r \leq \mathcal{R}_r(\theta^*)$$

- ▶ '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 $n_k = |I_k|$

$$\begin{matrix} I_0 & \subset & I_1 & \subset & \cdots & \subset & I_k & \subset & \cdots & \subset & I_K \\ \tilde{\theta}_0 & & \tilde{\theta}_1 & & & & \tilde{\theta}_k & & & & \tilde{\theta}_K \end{matrix}$$

Example: Trading volumes aggregated over 1-min periods

$$\text{Fix } i_0, I_k = [i_0 - n_k, i_0], n_k = \lceil n_0 c^k \rceil, c > 1$$

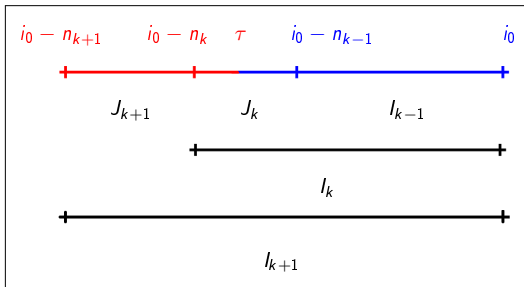
$$\{n_k\}_{k=0}^{13} = \{60 \text{ min.}, 75 \text{ min.}, \dots, 1 \text{ week}\}, c = 1.25$$



Local Change Point Detection ▶ Example

□ Fix i_0 , sequential test ($k = 1, \dots, K$)

H_0 : parameter homogeneity within I_k vs. H_1 : change point within I_k



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

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

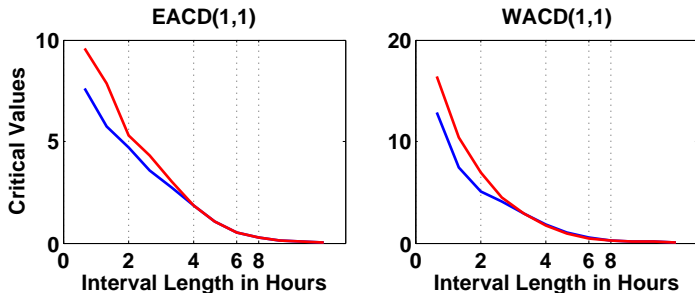


Figure 7: Critical values for low ($\tilde{\alpha} + \tilde{\beta} = 0.84$) and high ($\tilde{\alpha} + \tilde{\beta} = 0.93$) weekly persistence and 'modest' risk ($r = 0.5$) with $\rho = 0.25$



Adaptive Estimation

- Compare T_k at every step k with β_k
- Data window index of the *interval of homogeneity* - \hat{k}
- Adaptive estimate

$$\hat{\theta} = \tilde{\theta}_{\hat{k}}, \quad \hat{k} = \max_{k \leq K} \{k : T_l \leq \beta_l, l \leq k\}$$

- Note: rejecting the null at $k = 1$, $\hat{\theta}$ equals QMLE at l_0
If the algorithm goes until K , $\hat{\theta}$ equals QMLE at l_K



Adaptive Estimation - Results

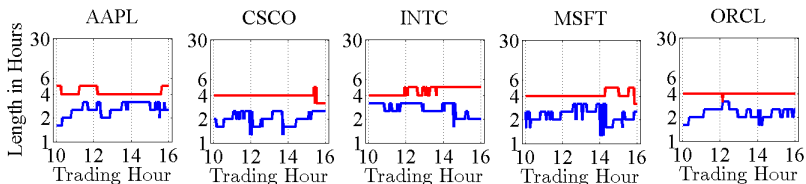


Figure 8: Estimated length $n_{\hat{k}}$ of *intervals of homogeneity* given the **modest** ($r = 0.5$) and the **conservative** ($r = 1$) risk case on 20080222 using the EACD(1, 1) model with $\rho = 0.25$



Adaptive Estimation - Results

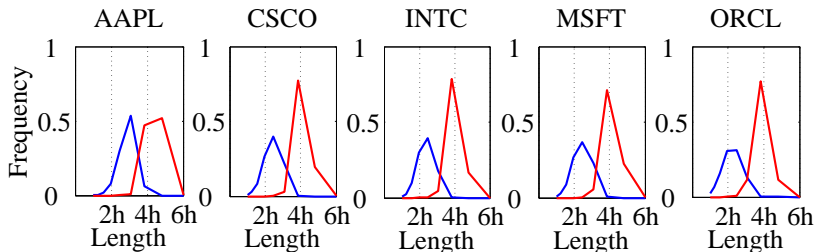


Figure 9: Relative frequency of the estimated interval length $n_{\hat{k}}$ (discrete variable, in hours) given the **modest** ($r = 0.5$) and **conservative** ($r = 1$) risk case from 20080222 to 20081231 using the EACD(1,1) model with $\rho = 0.25$



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), $r \in \{0.5, 1\}$, $\rho \in \{0.25, 0.5\}$

Strategies

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



Forecasting Superiority

- Diebold and Mariano (1995) tests, loss differential

$$d_h = \{d_{i+h}\}_{i=1}^n = \tilde{\varepsilon}_{i+h}^2 - \hat{\varepsilon}_{i+h}^2 \quad (2)$$

- Ratio of root mean squared errors

$$\sqrt{n^{-1} \sum_{i=1}^n \hat{\varepsilon}_{i+h}^2} / \sqrt{n^{-1} \sum_{i=1}^n \tilde{\varepsilon}_{i+h}^2} \quad (3)$$



Forecasting Superiority

- Qualitative test ▶ EACD ▶ WACD

$$T_{ST,h} = \left\{ \sum_{i=1}^n \mathbf{I}(d_{i+h} > 0) - 0.5n \right\} / \sqrt{0.25n} \xrightarrow{\mathcal{L}} N(0, 1) \quad (4)$$

- Quantitative test, $H_0 : E[d_h] = 0$

$$T_{DM,h} = \bar{d}_h / \sqrt{2\pi \hat{f}_{d_h}(0) / n} \xrightarrow{\mathcal{L}} N(0, 1) \quad (5)$$

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



Forecasting Superiority

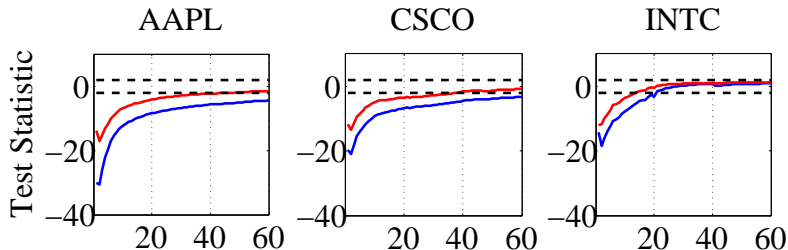


Figure 10: Test statistic $T_{DM,h}$ across all 60 forecasting horizons from 20080222 to 20081222 (210 trading days): LPA against a fixed-window scheme using 360 and 1800 observations using the EACD(1,1) model, $r = 0.5$ and $\rho = 0.25$



Forecasting Superiority

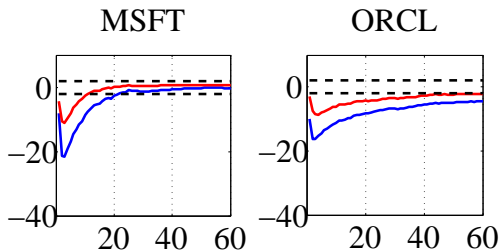


Figure 11: Test statistic $T_{DM,h}$ across all 60 forecasting horizons from 20080222 to 20081222 (210 trading days): LPA against a fixed-window scheme using 360 and 1800 observations using the EACD(1,1) model, $r = 0.5$ and $\rho = 0.25$



Forecasting Superiority

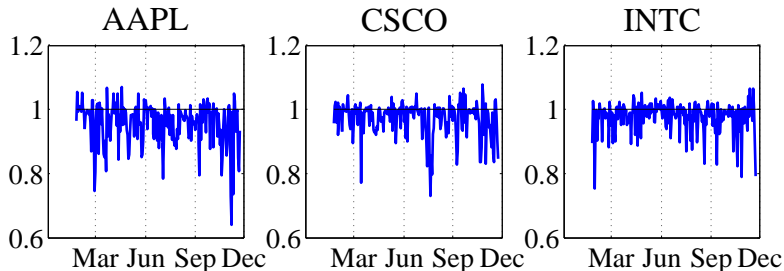


Figure 12: Ratio between the RMSPEs of the LPA and of a fixed-window approach (6 trading hours) over the sample from 20080222 to 20081222 (210 trading days) using an EACD(1,1) model, $r = 0.5$ and $\rho = 0.25$



Forecasting Superiority

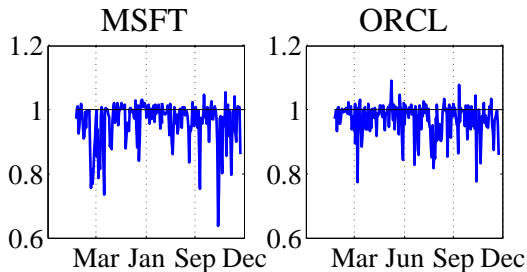


Figure 13: Ratio between the RMSPEs of the LPA and of a fixed-window approach (6 trading hours) over the sample from 20080222 to 20081222 (210 trading days) using an EACD(1,1) model, $r = 0.5$ and $\rho = 0.25$



Forecasting Superiority

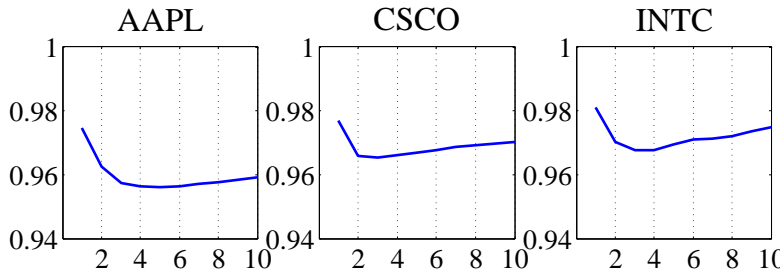


Figure 14: Ratio between the RMSPEs of the LPA and of a fixed-window approach (6 trading hours) across horizon from 20080222 to 20081222 (210 trading days) using an EACD(1,1) model, $r = 0.5$ and $\rho = 0.25$



Forecasting Superiority

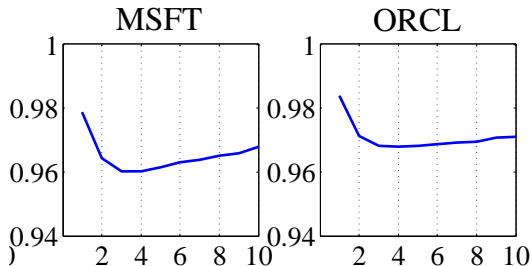


Figure 15: Ratio between the RMSPEs of the LPA and of a fixed-window approach (6 trading hours) across horizon from 20080222 to 20081222 (210 trading days) using an EACD(1,1) model, $r = 0.5$ and $\rho = 0.25$



Conclusions

(i) Localising MEM

- Time-varying parameters and estimation quality
- NASDAQ blue chips: AAPL, CSCO, INTC, MSFT and ORCL
- 'Conservative' adaptive estimation ($r = 1$) requires 4-5 hours of data, modest risk approach ($r = 0.5$) requires 2-3 hours

(ii) 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

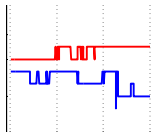
C.A.S.E. – Center for Applied Statistics
and Economics

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



Example ▶ Research Questions

Volume Weighted Average Price (VWAP) trading strategy

- Volume and market VWAP forecasting strategies
 - (i) 'Standard' method - fixed estimation window (one day/week)
 - (ii) LPA technique with selected interval of homogeneity
- VWAP trading strategy
 - ▶ Minimize liquidity impact costs (trading large orders at an inferior price compared to smaller orders) - order splitting
 - ▶ Benchmark the order price against (best predicted) market VWAP - trade at better than the market VWAP



Exponential-ACD (EACD)

▸ ACD

▸ Parameter Estimation

□ Engle and Russel (1998), $\varepsilon_i \sim \text{Exp}(1)$

$$L_I(y; \theta_E) = \sum_{i=\max(p,q)+1}^n \left(-\log \mu_i - \frac{y_i}{\mu_i} \right) \mathbf{1}\{i \in I\} \quad (6)$$

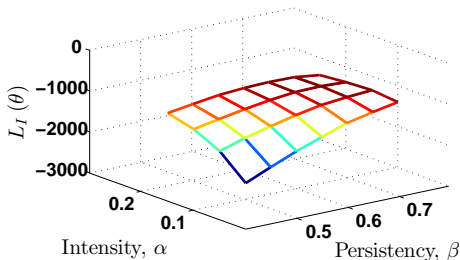


Figure 16: Log likelihood - EACD(1,1), $\theta_E^* = (0.10, 0.20, 0.65)^T$



Weibull-ACD (WACD)

▸ ACD

▸ Parameter Estimation

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

$$L_I(y; \theta_W) = \sum_{i \in I} \left[\log \frac{s}{y_i} + s \log \frac{\Gamma(1 + 1/s) y_i}{\mu_i} - \left\{ \frac{\Gamma(1 + 1/s) y_i}{\mu_i} \right\}^s \right] \mathbb{I}\{i \in I\} \quad (7)$$

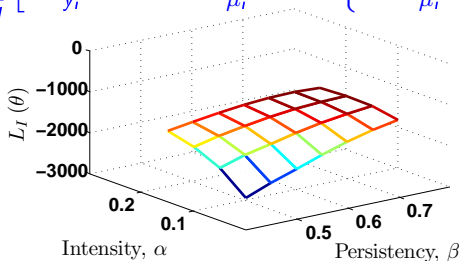


Figure 17: Log likelihood - WACD(1,1), $\theta_W^* = (0.10, 0.20, 0.65, 0.85)^\top$



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 $(\tilde{\alpha} + \tilde{\beta})$ for all five stocks at each minute from 20080222-20081231 (215 trading days). Calibration period: 20080102 - 20080221



Parameter Dynamics ▶ Critical Values

Model	Low Persistence			Moderate Persistence			High Persistence		
	Q25	Q50	Q75	Q25	Q50	Q75	Q25	Q50	Q75
EACD, $\tilde{\alpha}$	0.28	0.22	0.18	0.30	0.23	0.19	0.31	0.24	0.20
EACD, $\tilde{\beta}$	0.56	0.62	0.67	0.59	0.66	0.71	0.62	0.68	0.73
WACD, $\tilde{\alpha}$	0.28	0.21	0.17	0.30	0.23	0.18	0.32	0.24	0.19
WACD, $\tilde{\beta}$	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 $\tilde{\beta}/(\tilde{\alpha} + \tilde{\beta})$ (estimation windows covering 1800 observations) for all five stocks at each minute from 20080222-20081231 (215 trading days) conditional on the persistence level: low { EACD (0.85), WACD (0.82) }, moderate { EACD (0.89), WACD (0.88) } or high { EACD (0.93), WACD (0.92) }. Calibration period: 20080102-20080221



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



Critical Values, \mathfrak{z}_k

▶ Critical Values

- Simulate \mathfrak{z}_k - homogeneity of the interval sequence l_0, \dots, l_k
- 'Propagation' condition (under H_0)

$$E_{\theta^*} \left| L_{l_k}(\tilde{\theta}_{l_k}) - L_{l_k}(\hat{\theta}_{l_k}) \right|^r \leq \rho_k \mathcal{R}_r(\theta^*), \quad k = 1, \dots, K \quad (8)$$

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

- 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



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}



Forecasting Superiority

▶ Forecasting Superiority

	EACD(1, 1)				
	AAPL	CSCO	INTC	MSFT	ORCL
1 week					
$r = 0.5, \rho = 0.25$	-38.9	-28.6	-24.1	-33.8	-31.4
$r = 0.5, \rho = 0.50$	-38.7	-28.7	-24.2	-33.8	-31.4
$r = 1.0, \rho = 0.25$	-40.5	-31.4	-23.3	-39.1	-32.8
$r = 1.0, \rho = 0.50$	-40.4	-31.3	-23.3	-39.0	-32.9
1 day					
$r = 0.5, \rho = 0.25$	-10.8	-6.0	-13.1	-5.7	-15.1
$r = 0.5, \rho = 0.50$	-10.6	-6.0	-12.8	-5.5	-15.0
$r = 1.0, \rho = 0.25$	-6.9	-8.6	-8.7	-4.4	-12.9
$r = 1.0, \rho = 0.50$	-7.1	-8.6	-8.8	-4.4	-13.0

Table 3: Largest (in absolute terms) test statistic $T_{ST,h}$ across all 60 forecasting horizons from 20080222-20081222 (210 trading days). LPA against a fixed-window scheme using 360 and 1800 observations



Forecasting Superiority ▶ Forecasting Superiority

	WACD(1, 1)				
	AAPL	CSCO	INTC	MSFT	ORCL
1 week					
$r = 0.5, \rho = 0.25$	-22.6	-25.7	-20.2	-26.7	-26.6
$r = 0.5, \rho = 0.50$	-22.7	-25.5	-20.3	-26.7	-26.6
$r = 1.0, \rho = 0.25$	-27.9	-30.8	-21.5	-31.3	-29.8
$r = 1.0, \rho = 0.50$	-28.1	-30.8	-21.5	-31.5	-29.7
1 day					
$r = 0.5, \rho = 0.25$	-6.4	-3.5	-6.1	-4.9	-12.6
$r = 0.5, \rho = 0.50$	-6.3	-3.2	-6.2	-4.8	-12.7
$r = 1.0, \rho = 0.25$	-4.1	-5.1	-6.5	-4.2	-11.5
$r = 1.0, \rho = 0.50$	-3.9	-5.2	-6.5	-4.1	-11.4

Table 4: Largest (in absolute terms) test statistic $T_{ST,h}$ across all 60 forecasting horizons from 20080222-20081222 (210 trading days). LPA against a fixed-window scheme using 360 and 1800 observations

