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
 - 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

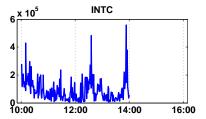


Figure 1: One-minute cumulated trading volume for Intel Corporation (INTC) on 20080902

Example

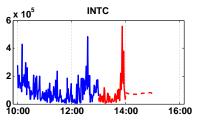


Figure 2: One-minute cumulated trading volume for Intel Corporation (INTC) on 20080902 with an EACD(1,1) estimation window length of 60 with volume forecasts up to the next one hour (dashed)

Example

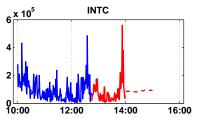


Figure 3: One-minute cumulated trading volume for Intel Corporation (INTC) on 20080902 with an EACD(1,1) estimation window length of 75 with volume forecasts up to the next one hour (dashed)

Example

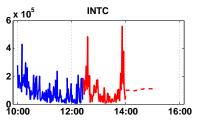


Figure 4: One-minute cumulated trading volume for Intel Corporation (INTC) on 20080902 with an EACD(1,1) estimation window length of 95 with volume forecasts up to the next one hour (dashed)

<u>~</u>~~~

Research Questions PExample

- How strong is the variation of MEM parameters over time?
- What are typical interval lengths of parameter homogeneity?



Outline

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



Multiplicative Error Models (MEM)

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

$$y_i = \mu_i \varepsilon_i,$$
 $\mathbb{E}\left[\varepsilon_i \mid \mathcal{F}_{i-1}\right] = 1$
 $\mu_i = \omega + \sum_{j=1}^p \alpha_j y_{i-j} + \sum_{j=1}^q \beta_j \mu_{i-j},$ $\omega > 0, \alpha_j, \beta_j \ge 0$

- □ Hautsch (2012) comprehensive MEM literature overview
 - y_i squared (de-meaned) log return: GARCH(p, q)
 - \triangleright 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 $\varepsilon_i \sim \mathsf{Exp}(1), \ \theta_{\mathsf{E}} = (\omega, \alpha, \beta)^{\top}, \ \alpha = (\alpha_1, \dots, \alpha_p), \ \beta = (\beta_1, \dots, \beta_q)$
- 2. Weibull-ACD, Engle and Russel (1998) → WACD $\varepsilon_i \sim \mathcal{G}(s,1), \ \theta_W = (\omega, \alpha, \beta, s)^{\top}$

Weibull, E. H. Waloddi on BBI:



Parameter Estimation

- □ Consistent parameter estimation
- oxdot Quasi maximum likelihood estimates (QMLEs) of $oldsymbol{ heta}_E$ and $oldsymbol{ heta}_W$

$$\widetilde{\boldsymbol{\theta}}_{I} = \arg\max_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} L_{I}(\boldsymbol{y}; \boldsymbol{\theta})$$
 (1)

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

Data

- NASDAQ Stock Market in 2008, 250 trading days
 - ▶ 5 stocks: AAPL, CSCO, INTC, MSFT and ORCL
 - $ightharpoonup reve{y}_i$ 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

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

 $\overline{\imath} \in (0,1]$ - 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



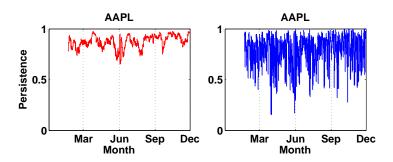


Figure 6: Estimated weekly (n=1800) and daily (n=360) persistence $\widetilde{\alpha}_i + \widetilde{\beta}_i$ for seasonally adjusted trading volume using an EACD(1,1) at each minute in 2008

Local Adaptive MEM



Estimation Quality

- $oxed{oxed}$ Quality of estimating *true* parameter vector $oldsymbol{ heta}^*$ by QMLE $oldsymbol{ heta}_I$ in terms of Kullback-Leibler divergence; $\mathcal{R}_r(oldsymbol{ heta}^*)$ risk bound

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

- ightharpoonup 'Modest' risk, r = 0.5 (shorter intervals of homogeneity)
- ightharpoonup '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
- - Volatility modelling Mercurio and Spokoiny (2004)
 - ► GARCH(1,1) models Čížek et al. (2009)
 - Realized volatility Chen et al. (2010)



Interval Selection

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

Example: Trading volumes aggregated over 1-min periods

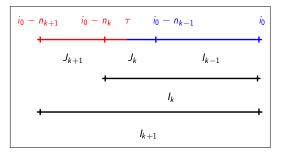
Fix
$$i_0$$
, $I_k = [i_0 - n_k, i_0]$, $n_k = [n_0 c^k]$, $c > 1$

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

Local Change Point Detection **PExample**

Fix i_0 , sequential test (k = 1, ..., K)

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



$$\begin{split} T_k &= \sup_{\tau \in J_k} \left\{ L_{A_{k,\tau}} \left(\widetilde{\theta}_{A_{k,\tau}} \right) + L_{B_{k,\tau}} \left(\widetilde{\theta}_{B_{k,\tau}} \right) - L_{I_{k+1}} \left(\widetilde{\theta}_{I_{k+1}} \right) \right\}, \\ \text{with } J_k &= I_k \setminus I_{k-1}, \ A_{k,\tau} = [i_0 - n_{k+1}, \tau] \text{ and } B_{k,\tau} = (\tau, i_0] \end{split}$$

Local Adaptive MEM



Critical Values, 3k Critical Values

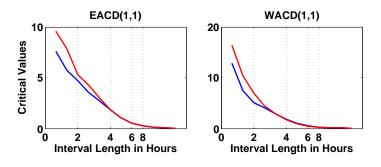


Figure 7: Critical values for low $(\widetilde{\alpha} + \widetilde{\beta} = 0.84)$ and high $(\widetilde{\alpha} + \widetilde{\beta} = 0.93)$ weekly persistence and 'modest' risk (r = 0.5) with $\rho = 0.25$

Adaptive Estimation

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

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

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

Adaptive Estimation - Results

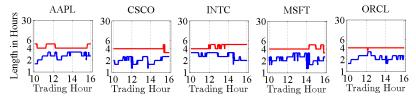


Figure 8: Estimated length $n_{\hat{k}}$ of intervals of homogeneity given the conservative (r=1) and modest (r=0.5) risk case on 20080222 using the EACD(1,1) model with $\rho=0.25$

Adaptive Estimation - Results

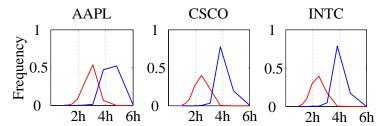


Figure 9: Relative frequency of the estimated interval length $n_{\hat{k}}$ (discrete variable, in hours) given the modest (r=0.5) and conservative (r=1) risk case from 20080222 to 20081231 using the EACD(1,1) model with $\rho=0.25$

Adaptive Estimation - Results

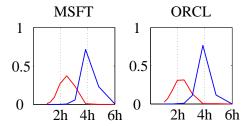


Figure 10: Relative frequency of the estimated interval length $n_{\hat{k}}$ (discrete variable, in hours) given the modest (r=0.5) and conservative (r=1) risk case from 20080222 to 20081231 using the EACD(1,1) model with $\rho=0.25$

Forecasting Trading Volumes

Setup

- $oxed{\Box}$ Forecasts at each minute, horizon $h=1,\ldots,60$ min.
- □ EACD(1, 1) and WACD(1, 1), $r ∈ \{0.5, 1\}$, $ρ ∈ \{0.25, 0.5\}$

Strategies

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

Diebold and Mariano (1995) tests, loss differential

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

Ratio of root mean squared errors

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

Qualitative test ▶ EACD ▶ WACD

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

 $oxed{\Box}$ Quantitative test, $H_0: E[d_h] = 0$

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

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

Local Adaptive MEM

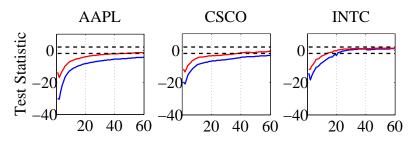


Figure 11: Test statistic $T_{DM,h}$ across all 60 forecasting horizons from 20100222 to 20101222 (210 trading days): LPA against a fixed-window scheme using 360 and 1800 observations using the EACD(1,1) model, r=0.5 and $\rho=0.25$

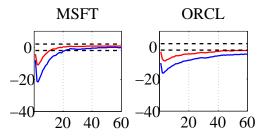


Figure 12: Test statistic $T_{DM,h}$ across all 60 forecasting horizons from 20100222 to 20101222 (210 trading days): LPA against a fixed-window scheme using 360 and 1800 observations using the EACD(1,1) model, r=0.5 and $\rho=0.25$

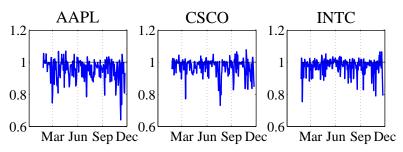


Figure 13: Ratio between the RMSPEs of the LPA and of a fixed-window approach (6 trading hours) over the sample from 20100222 to 20101222 (210 trading days) using an EACD(1,1) model, r=0.5 and $\rho=0.25$

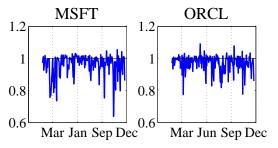


Figure 14: Ratio between the RMSPEs of the LPA and of a fixed-window approach (6 trading hours) over the sample from 20100222 to 20101222 (210 trading days) using an EACD(1,1) model, r=0.5 and $\rho=0.25$

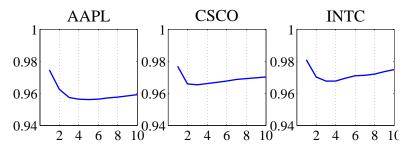


Figure 15: Ratio between the RMSPEs of the LPA and of a fixed-window approach (6 trading hours) across horizon from 20100222 to 20101222 (210 trading days) using an EACD(1,1) model, r=0.5 and $\rho=0.25$

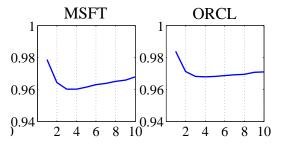


Figure 16: Ratio between the RMSPEs of the LPA and of a fixed-window approach (6 trading hours) across horizon from 20100222 to 20101222 (210 trading days) using an EACD(1,1) model, r=0.5 and $\rho=0.25$

Conclusions — 5-1

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

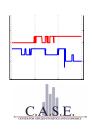
- LPA outperforms the 'standard' method
- Overall performance, horizon, trading day, tuning parameters



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

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

Local Adaptive MEM



References — 6-2

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 — 6-3

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

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

oxdot Engle and Russel (1998), $arepsilon_i \sim {\it Exp}\,(1)$

$$L_{I}(y; \boldsymbol{\theta}_{E}) = \sum_{i=\max(p,q)+1}^{n} \left(-\log \mu_{i} - \frac{y_{i}}{\mu_{i}}\right) \mathbb{I}\left\{i \in I\right\}$$
 (6)

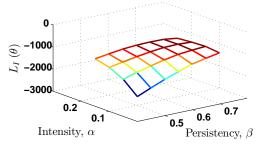


Figure 17: Log likelihood - EACD(1,1), $\theta_{E}^{*}=(0.10,0.20,0.65)$



Weibull-ACD (WACD)

→ ACD → Parameter Estimation

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

Figure 18: Log likelihood - WACD(1,1), $\theta_W^* = (0.10, 0.20, 0.65, 0.85)^{\top}$ Local Adaptive MEM

Parameter Dynamics

Estimation	EACD(1,1)			WACD(1, 1)		
window	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). Calibration period: 2 January 2008 - 21 February 2008

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{eta}	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{eta}	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). Calibration period: 2 January 2008 - 21 February 2008

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, 3k Critical Values

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

$$\mathsf{E}_{\theta^*} \left| L_{l_k}(\widetilde{\theta}_{l_k}) - L_{l_k}(\widehat{\theta}_{l_k}) \right|^r \le \rho_k \mathcal{R}_r(\theta^*), \quad k = 1, \dots, K \quad (8)$$

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

- \odot 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 ••••

Example: Trading volumes aggregated over 1-min periods

- \odot Scheme with (K+1)=14 intervals and fix i_0
- oxdot Assume $I_0=60$ min. is homogeneous
- - ▶ Define $J_1 = I_1 \setminus I_0$ observations from y_{i_0-75} up to y_{i_0-60}
 - lacksquare For each $au\in J_1$ fit log likelihoods over $A_{1, au}$, $B_{1, au}$ 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

Local Adaptive MEM



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

Local Adaptive MEM