Localized Realized Volatility Modeling

Ying Chen

Wolfgang Karl Härdle

Uta Pigorsch

National University of Singapore

Humboldt-Universität zu Berlin

University of Mannheim

An important realized volatility fact

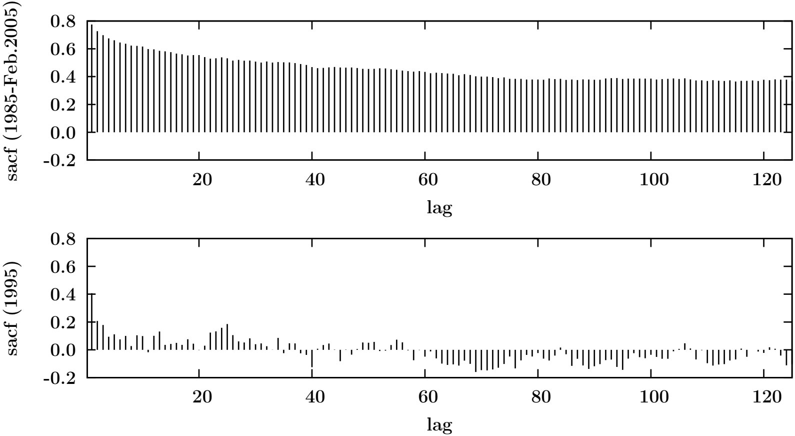


Figure 1: Sample autocorrelations of log RV for different sample periods.

A dual view

- **The long memory point of view:**
  Volatility is generated by long memory processes, i.e. fractionally integrated, $I(d)$, processes.
- **The short memory point of view:**
  Volatility may equally well be generated by a short memory process with structural changes.
  Example: GARCH model with changing parameters.

Realized volatility

- Volatility forecasts are important for an adequate risk management and derivative pricing.
- Realized volatility is based on high-frequency information.
- It is a more precise volatility estimator than daily squared or absolute returns.
- Exhibits better forecast properties, Andersen, Bollerslev, Diebold and Labys (2001).

Localized realized volatility

- For a point τ in time, find a past time interval for which a local volatility model is a good approximator.
- The time interval is determined by adaptive statistical methods.
- Represents a local analysis, i.e. changes are detected close to the forecasting time point.

Outline

1. Motivation ✓
2. Realized volatility
3. Localized realized volatility
4. Long memory models
5. Empirical analysis
6. Conclusion

Realized volatility

Daily realized volatility

$$\widetilde{RV}_t = \sum_{j=1}^{M} r_{t,j}^2,$$

with $r_{t,j} = p_{t,n_j} - p_{t,n_{j-1}}$, $j = 1, \ldots, M$, and p_{t,n_j} the log price observed at time point n_j of trading day t.
It converges to the quadratic variation for $M \to \infty$ (Andersen and Bollerslev, 1998; Barndorff-Nielsen and Shephard, 2002b).

Realized volatility (smoothed)

Tukey-Hanning kernel

$$RV_t = \widetilde{RV}_{t,1} + \sum_{h=1}^{H^*} k\left(\frac{h-1}{H^*}\right)(\gamma_{t,h} + \gamma_{t,-h})$$

$k(x) = \sin^2\left\{\frac{\pi}{2}(1-x)^2\right\}$,
$\gamma_{t,h} = \sum_{j=1}^{M} r_{t,j} r_{t,j-h}$ (one-minute returns),
$H^* = 5.74 \frac{\widetilde{RV}_{t,1}/2M}{\widetilde{RV}_{t,15}} \sqrt{M}$ with $RV_{t,i}$ the realized variance estimator based on i minute returns.

(Barndorff-Nielsen, Hansen, Lunde and Shephard, 2008)

Data

S&P500 index futures from January 2, 1985 to February 4, 2005.

Series	Mean	Std.Dev.	Skewness	Kurtosis	LB(21)[1]
RV_t	1.07	8.16	59.08	3861	1375
$\log(RV_t)$	-0.51	0.87	0.43	4.99	46809

[1] The critical value of this Ljung-Box test is 32.671.

Table 1: Descriptive statistics.

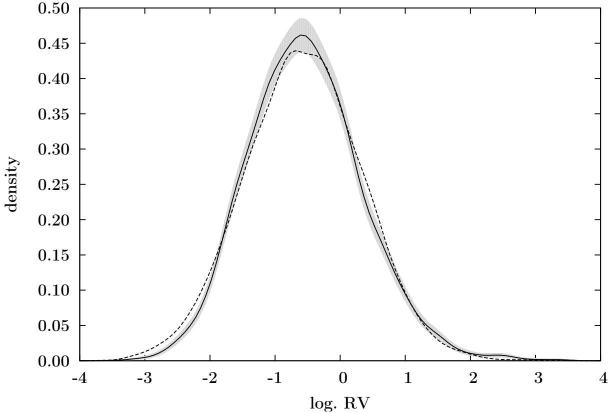


Figure 2: Kernel density estimates (solid line: log RV, shaded area: pointwise 95% confidence intervals, dashed line: normal distribution).

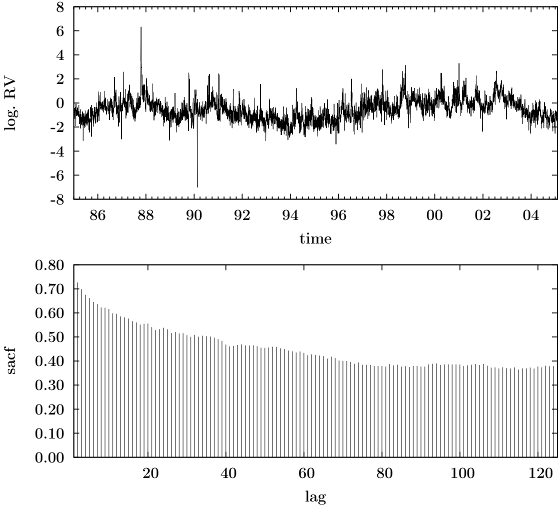


Figure 3: log RV and its sample autocorrelation function.

Localized realized volatility

LAR(1) model with parameter set $\theta_t = (\theta_{1t}, \theta_{2t}, \theta_{3t})^\top$:

$$\log RV_t = \theta_{1t} + \theta_{2t} \log RV_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim \mathrm{N}(0, \theta_{3t}^2). \tag{1}$$

Suppose $\theta_t \equiv \theta^*$ for $t \in I = [1, T]$

$$\begin{aligned} \tilde{\theta}_t &= \operatorname{argmax}_{\theta \in \Theta} L(\log RV; \theta) \\ &= \operatorname{argmax}_{\theta \in \Theta} \left\{ -\frac{T}{2} \log 2\pi - T \log \theta_3 \right. \\ &\qquad \left. - \frac{1}{2\theta_3^2} \sum_{t=1}^{T} (\log RV_t - \theta_1 - \theta_2 \log RV_{t-1})^2 \right\}. \end{aligned}$$

Goal: identify a local homogeneous interval I_τ for time point τ.

Identify local homogeneity

At time point τ, choose a local homogeneous interval from

$$\{I_\tau^k\}_{k=1}^K = \{I_\tau^1, I_\tau^2, \cdots, I_\tau^K\}$$

where $I_\tau^k = [\tau - s_k, \tau)$ with $0 < s_k < \tau$, which leads to the best possible accuracy of estimation.

- Under **local homogeneity** $\theta_\tau \equiv \theta_\tau^*$ within $I_\tau^k = [\tau - s_k, \tau)$:

$$\tilde{\theta}_\tau^{(k)} \text{ estimates } \theta_\tau^* \text{ at rate } 1/\sqrt{s_k}$$

- The modeling bias of approximating LAR(1) increases w.r.t. k.

The optimal choice $\hat{I}_\tau$: balances the bias and variation.

Estimation under local homogeneity

Given $I_\tau = [\tau - s, \tau)$, the local MLE is:

$$\tilde{\theta}_\tau = \operatorname{argmax}_{\theta \in \Theta} L(\log RV; I_\tau, \theta)$$

Under **local homogeneity**: $\theta_\tau \equiv \theta^*_\tau$, the fitted likelihood ratio measures the estimation risk:

$$LR(I_\tau, \tilde{\theta}_\tau, \theta^*_\tau) = L(I_\tau, \tilde{\theta}_\tau) - L(I_\tau, \theta^*_\tau). \tag{2}$$

Estimation under local homogeneity

The estimation risk $LR(I_\tau, \tilde{\theta}_\tau, \theta^*_\tau)$ is stochastically bounded:

$$\mathsf{E}_{\theta^*_\tau} \left| LR(I_\tau, \tilde{\theta}_\tau, \theta^*_\tau) \right|^r \leq \xi_r$$

with $\xi_r = 2r \int_{\xi \geq 0} \xi^{r-1} e^{-\xi} d\xi = 2r\Gamma(r)$.
It leads to the **confidence set**:

$$\mathcal{E}(\varepsilon) = \{\theta : LR(I_\tau, \tilde{\theta}_\tau, \theta^*_\tau) \leq \varepsilon\}$$

in the sense that $\mathrm{P}_{\theta^*}\left\{\mathcal{E}(\varepsilon) \not\ni \theta^*\right\} \leq \alpha$, Polzehl and Spokoiny (2006).

Localized AR(1) model

Interval set $\{I_\tau^k\}$ for $k = 1, \cdots, K$:

$$
\begin{array}{cccc}
I_\tau^1 = [\tau - 1\text{w}, \tau) & I_\tau^2 = [\tau - 1\text{m}, \tau) & \cdots & I_\tau^K = [\tau - 5\text{y}, \tau) \\
\downarrow & \downarrow & \downarrow & \downarrow \\
\tilde{\theta}_\tau^{(1)} & \tilde{\theta}_\tau^{(2)} & \ldots & \tilde{\theta}_\tau^{(K)}
\end{array}
$$

- The interval is growing in length.
- Local homogeneity is assumed at I_τ^1.
- Final estimate $\hat{\theta}_\tau$ is based on a sequential testing.

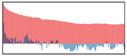
Sequential testing

Suppose that I_τ^{k-1} is a homogeneous interval: $\hat{\theta}_\tau^{(k-1)} = \tilde{\theta}_\tau^{(k-1)}$. The null hypothesis at step k:

$$H_0 : I_\tau^k \text{ is an homogeneous interval.}$$

Test homogeneity of I_τ^k: $\hat{\theta}_\tau^k = \tilde{\theta}_\tau^k$ or terminates at I_τ^{k-1}

Test: $\left| LR(I_\tau^k, \tilde{\theta}_\tau^k, \hat{\theta}_\tau^{k-1}) \right|^r \leq \zeta_k$, where ζ_k is critical value (CV).



Adaptive procedure

1. Initialization: $\hat{\theta}_\tau^1 = \tilde{\theta}_\tau^1$.
2. $k = 1$
   while $\left| LR(I_\tau, \tilde{\theta}_\tau^{k+1}, \hat{\theta}_\tau^k) \right|^r \leq \zeta_{k+1}$ and $k < K$,

$$\begin{aligned} k &= k+1 \\ \hat{\theta}_\tau^k &= \tilde{\theta}_\tau^k \end{aligned}$$

3. Final estimate: $\hat{\theta}_\tau = \hat{\theta}_\tau^k$

Parameter choice: Interval set

- $\{I^k\}_{k=1}^{13}$ for every τ with the following interval lengths:

$$\begin{aligned}\{s_k\}_{k=1}^{13} \quad = \quad & \{1\text{w}, 1\text{m}, 3\text{m}, 6\text{m}, 1\text{y}, 1.5\text{y}, \\ & 2\text{y}, 2.5\text{y}, 3\text{y}, 3.5\text{y}, 4\text{y}, 4.5\text{y}, 5\text{y}\},\end{aligned}$$

where w denotes a week (5 days), m refers to one month (21 days) and y to one year (252 days).

Parameter choice: CV

- Monte Carlo simulation: generate AR(1) processes with $\theta_t = \theta^* = (\theta_1^*, \theta_2^*, \theta_3^*)^\top$ for all t
  - $y_t = \theta_1^* + \theta_2^* y_{t-1} + \varepsilon_t$, $\varepsilon_t \sim \mathrm{N}(0, \theta_3^{*2})$.
  - $y_0 = \theta_1^* / (1 - \theta_2^*)$
  - 100 000 paths, each including 1261 observations
- Choice of critical values:

  Parametric case $\theta_t \equiv \theta^*$:

$$\mathrm{E}_{\theta^*} \left| LR\left(I^K, \tilde{\theta}_t^K, \hat{\theta}^K_{t(\zeta_1,\ldots,\zeta_K)} \right) \right|^r \leq \xi_r \tag{3}$$

$$\mathrm{E}_{\theta^*} \left| LR\left(I^k, \tilde{\theta}_t^k, \hat{\theta}^k_{t(\zeta_1,\ldots,\zeta_k)} \right) \right|^r \leq \frac{k-1}{K-1} \xi_r \tag{4}$$

Parameter choice: CV

Sequential choice of critical values

- Choice of $\zeta_1 = \infty$ initializing the procedure $\hat{\theta}_t^1(\zeta_1) = \tilde{\theta}_t^1$
- Choice of ζ_2 leading to $\hat{\theta}_t^k(\zeta_1, \zeta_2)$ by setting $\zeta_3 = \ldots = \zeta_K = \infty$:

$$\mathrm{E}_{\theta^*} \left| LR\left(I^k, \tilde{\theta}_t^k, \hat{\theta}_t^k(\zeta_1, \zeta_2)\right) \right|^r \leq \frac{1}{K-1} \xi_r , \qquad k = 2, \ldots, K.$$

- Choice of ζ_3 leading to $\hat{\theta}_t^k(\zeta_1, \zeta_2, \zeta_3)$ by setting $\zeta_4 = \ldots = \zeta_K = \infty$:

$$\mathrm{E}_{\theta^*} \left| LR\left(I^k, \tilde{\theta}_t^k, \hat{\theta}_t^k(\zeta_1, \zeta_2, \zeta_3)\right) \right|^r \leq \frac{2}{K-1} \xi_r , \qquad k = 3, \ldots, K.$$

Parameter choice: CV

Sequential choice of critical values

- Choice of ζ_k leading to $\hat{\theta}_t^l(\zeta_1, \cdots, \zeta_k)$ by setting $\zeta_{k+1} = \ldots = \zeta_K = \infty$:

$$\mathrm{E}_{\theta^*} \left| LR\left(I^k, \tilde{\theta}_t^l, \hat{\theta}_t^l(\zeta_1, \ldots, \zeta_k)\right) \right|^r \leq \frac{k-1}{K-1} \xi_r \,, \qquad l = k, \ldots, K.$$

Parameter choice: CV and r

- The critical values (CV) depend on θ^*_τ used in the Monte Carlo simulation:
  - local global CV: global parameter $\theta^*_\tau = \theta^*$ over the whole sample
  - local adaptive CV: time varying parameter θ^*_τ using a moving window with fixed size.
- r: default choice $1/2$

Critical values

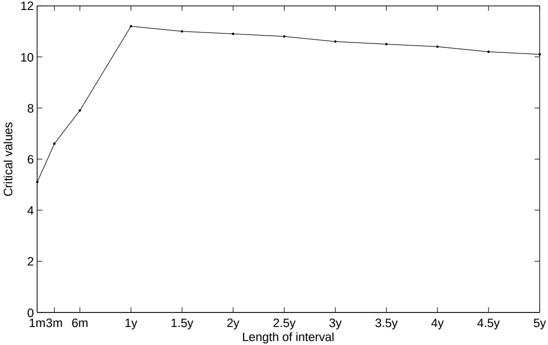
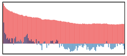


Figure 4: CV for $r = 1/2$ and $\theta^* = (-0.1197, 0.7754, 0.5634)^\top$ (global). Data source: log RV of the S&P500 index futures.



Simulation

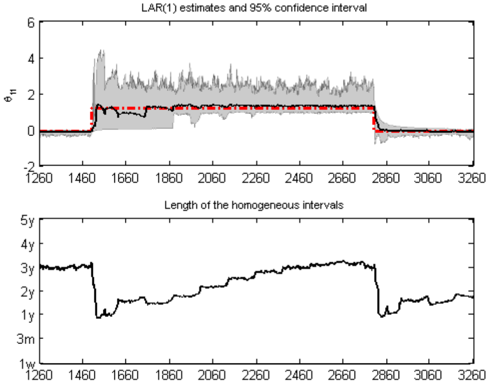


Figure 5: The average values for $\theta^*_{1t} \in \{-0.120, 1.197\}$.

Simulation

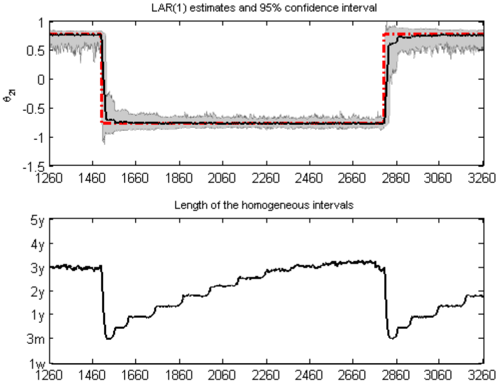


Figure 6: The average values for $\theta^*_{2t} \in \{-0.775, 0.775\}$.

Simulation

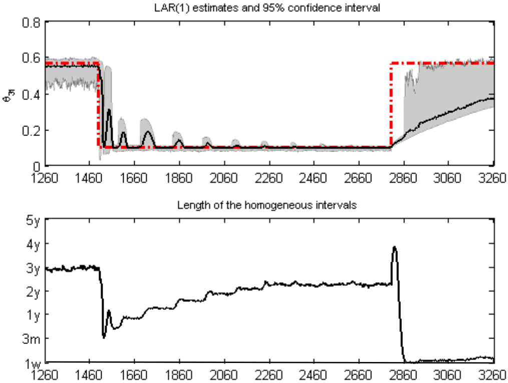
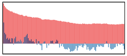


Figure 7: The average values for $\theta^*_{3t} \in \{0.100, 0.563\}$.



The ARFIMA model

ARFIMA(p, d, q), model:

$$\phi(L)(1-L)^d(\log RV_t - \mu) = \psi(L)u_t,$$

with $\phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p$, $\psi(L) = 1 + \psi_1 L + \cdots + \psi_q L^q$
L denoting the lag operator, $d \in (0, 0.5)$, $u_t \overset{iid}{\sim} \mathrm{N}(0, \sigma^2)$.
(Andersen, Bollerslev, Diebold and Labys, 2003)

The HAR model

Heterogeneous autoregressive model:

$$\log RV_t = \alpha_0 + \alpha_d \log RV_{t-1} + \alpha_w \log RV_{t-5:t-1} + \alpha_m \log RV_{t-21:t-1} + u_t$$

with the multiperiod realized volatility components

$$RV_{t+1-k:t} = \frac{1}{k} \sum_{j=1}^{k} RV_{t-j}.$$

Empirical evidence

- The HAR model is *no* long memory model, but provides an approximation.
- The HAR and ARFIMA models exhibit *similar* in-sample and out-of-sample performance.
- Both strongly *outperform* conventional volatility models.

Forecast setup

- The first five years (1985-1989) of the S&P 500 index futures data serve as *training set*.
- The remaining data serve as *forecast evaluation period* (January 2, 1990 to February 4, 2005).

Forecast setup

- Consider 5 sets of critical values: the global ones and the adaptive ones based on a sample period of 1 month, 6 months, 1 year and 2.5 years.
- The interval of homogeneity is always selected based on the following set of interval lengths:

$$
\begin{aligned}
\{s_k\}_{k=1}^{13} &= \{1\text{w}, 1\text{m}, 3\text{m}, 6\text{m}, 1\text{y}, 1.5\text{y}, \\
&\quad 2\text{y}, 2.5\text{y}, 3\text{y}, 3.5\text{y}, 4\text{y}, 4.5\text{y}, 5\text{y}\},
\end{aligned}
$$

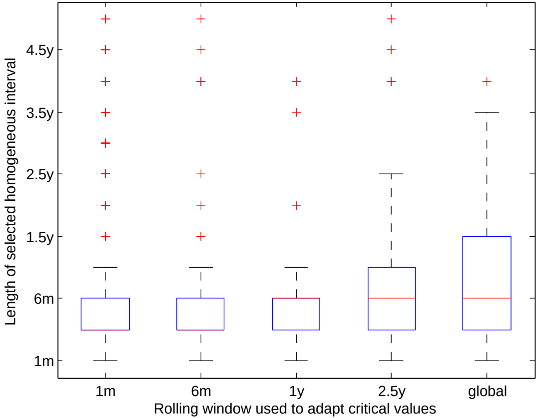


Figure 8: Boxplot of the homogenous intervals selected by the LAR(1) procedure based on different sets of critical values.

Forecast setup

- ARFIMA forecasts are based on an ARFIMA(2,d,0) model (as indicated by the AIC and BIC using the full sample period).
- Estimation and prediction is performed using a rolling window scheme with different window sizes, i.e. $\{3\text{m}, 6\text{m}, 1\text{y}, 1.5\text{y}, 2\text{y}, 2.5\text{y}, 3\text{y}, 3.5\text{y}, 4\text{y}, 4.5\text{y}, 5\text{y}\}$.
- Same setup is used to assess the predictability of the HAR model and a constant AR(1) model, i.e. $\log RV_t = \alpha_0 + \alpha_1 \log RV_{t-1} + u_t,\ u_t \overset{iid}{\sim} N(0, \sigma^2)$.

crit. values	LAR(1)	sample size	AR(1)	ARFIMA	HAR
local adaptive 1m	0.4858	3m	**0.5149**	0.5328	0.5381
local adaptive 6m	**0.4811**	6m	0.5288	0.5225	0.5240
local adaptive 1y	0.4876	1y	0.5398	0.5178	0.5185
local adaptive 2.5y	0.4916	1.5y	0.5462	0.5143	0.5172
local global	0.5014	2y	0.5509	0.5133	0.5158
		2.5y	0.5555	0.5132	**0.5153**
		3y	0.5574	**0.5123**	0.5155
		4y	0.5649	0.5129	0.5171
		5y	0.5712	0.5129	0.5176

Table 2: Root mean square forecast errors.

crit. values	LAR(1)	sample size	AR(1)	ARFIMA	HAR
local adaptive 1m	0.3667	3m	**0.3900**	0.3978	0.4025
local adaptive 6m	**0.3654**	6m	0.3987	0.3902	0.3862
local adaptive 1y	0.3704	1y	0.4057	0.3860	0.3857
local adaptive 2.5y	0.3748	1.5y	0.4103	0.3836	0.3843
local global	0.3824	2y	0.4136	0.3826	**0.3836**
		2.5y	0.4157	0.3816	0.3839
		3y	0.4177	**0.3814**	0.3843
		4y	0.4238	0.3817	0.3851
		5y	0.4300	0.3819	0.3858

Table 3: Mean absolute forecast error.

Empirical results

Mincer–Zarnowitz regression:

Evaluate the predictive performance of the different models based on Mincer–Zarnowitz regressions:

$$\log RV_t = \alpha + \beta \widehat{\log RV}_{t,i} + \nu_t$$

with $\widehat{\log RV}_{t,i}$ denoting the log realized volatility forecast of model i.

- Assess R^2 of the regression.
- Test on *unbiasedness* of the different forecasts:
  $H_0: \ \alpha = 0$ and $\beta = 1$.

Model	α	β	p-value	R^2
global LAR	-0.0130	1.0128	0.1007	0.6959
adaptive LAR, 1y	0.0025	1.0014	0.9780	0.7117
1y AR(1)	-0.0010	1.0117	0.6002	0.4669
5y AR(1)	0.0221	1.0367	0.2216	0.6052
1y ARFIMA	0.0008	1.0011	0.9962	0.6747
5y ARFIMA	0.0009	1.0154	0.4907	0.6811
1y HAR	-0.0076	0.9907	0.7509	0.6742
5y HAR	0.0145	1.0237	0.2036	0.6756

Table 4: Mincer–Zarnowitz regression results.

Empirical results

- **Test on equal forecast performance:**
  Diebold–Mariano test on *equal MSFEs*:

$$e^2_{t,LAR} - e^2_{t,i} = \mu + v_t$$

with $e_{t,i}$ denoting the forecast error of model i.

  - $H_0: \ \mu = 0.$

global LAR compared to	t-values for $H_0 : \mu = 0$	adaptive LAR,1y compared to	t-values for $H_0 : \mu = 0$
1y AR(1)	-4.1667	1y AR(1)	-5.5154
5y AR(1)	-1.2935	5y AR(1)	-4.9189
1y ARFIMA	-1.5412	1y ARFIMA	-2.8148
5y ARFIMA	-1.0825	5y ARFIMA	-2.3211
1y HAR	-1.6827	1y HAR	-3.0097
5y HAR	-1.5865	5y HAR	-2.5048

Table 5: Diebold–Mariano test results.

Conclusion

- Dual view on the long memory diagnosis of volatility.
- Long memory phenomenon can be reproduced by localized short memory.
- Identification of homogeneity by a localized realized volatility model.

Conclusion

- The localized approach outperforms long memory-type models and constant AR(1) models in terms of predictability.
- An adaptive choice of the critical values (and a decrease in the underlying sample period) improves the estimation and forecast accuracy of the localized approach.

References

Andersen, T. G. and Bollerslev, T.
Answering the skeptics: yes, standard volatility models do provide accurate forecasts
International Economic Review, 39, 885-905, 1998.

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P.
The distribution of realized exchange rate volatility
Journal of the American Statistical Association, 96, 42-55, 2001.

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P.
Modeling and forecasting realized volatility
Econometrica, 71, 579-625, 2003.

References

Barndorff-Nielsen, O. E. and Shephard, N.
Estimating quadratic variation using realized variance
Journal of Applied Econometrics, 17, 457-477, 2002.

Barndorff-Nielsen, O.E., Hansen, P. R., Lunde, A. and Shephard, N.
Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise
Econometrica, 76, 1481-1536, 2008.