Volatility, Information Feedback and Market Microstructure Noise: A Tale of Two Regimes

Torben G. Andersen
Northwestern University

Gökhan Cebiroglu
University of Vienna

Nikolaus Hautsch
University of Vienna

Introduction

- Major topic in financial econometrics over the last decade:

  How can we optimally use financial high-frequency data to construct efficient volatility estimators for aggregated periods (intraday, day, week, ...)?

- Typical starting point:

$$p_{t_i} = p^*_{t_i} + \epsilon_i, \quad \epsilon_i \sim (0, \sigma^2_\epsilon), \quad i = 1, \ldots, n,$$
$$p^*_t = p^*_0 + \int_0^t \sigma^{*2}(s) dB_s \, , t \in [0, T],$$

  where p_t denotes the (observed) log price, B_t is a standard Brownian motion B_t, and ϵ_i is microstructure "noise".

- Object of interest: $\int_0^T \sigma^{*2}(s) ds$, corresponding to the variance of a T-period return.

- If p_t are discretely observed with $p_{i/n}, i = 0, \ldots, n$, a natural estimator for $\int_0^T \sigma^2(s) ds$ is given by the *realized variance*,

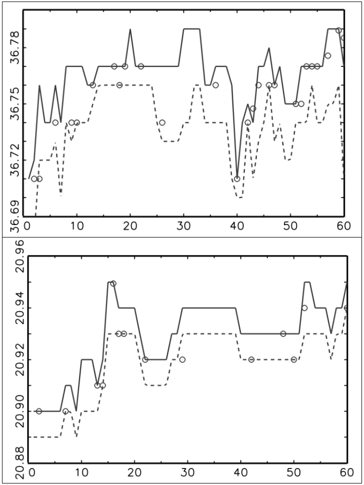
$$\mathrm{RV}^n = \sum_{i=1}^{n} (p_{i/n} - p_{(i-1)/n})^2,$$

which is consistent and efficient with

$$n^{1/2} \left(\mathrm{RV}^n - \int_0^1 \sigma^{*2}(s)\, ds \right) \xrightarrow{\mathcal{L}} \mathbf{N}\Big(0, 2 \int_0^1 \sigma^{*4}(s)\, ds\Big) .$$

- Suggestion: Sampling on highest possible frequency!

Real Intraday Price Path



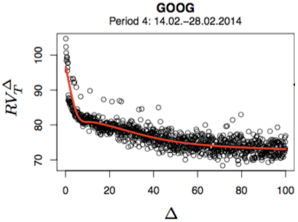
- <u>Problem</u>: HF prices are subject to noise, i.e., we only observe

$$p_i = p_i^* + \epsilon_i, \quad i = 1, \ldots, n,$$

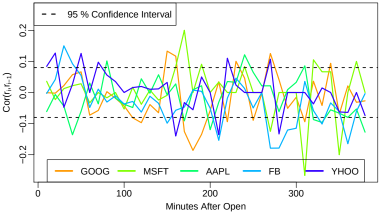
where ϵ_i is associated with market microstructure "noise" (MMN).

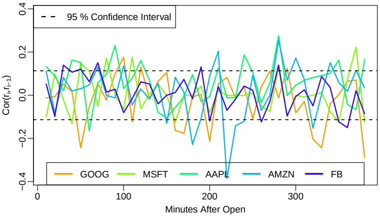
⇒ If we let $n \to \infty$, RV becomes biased and inconsistent.

⇒ If MMN is i.i.d., (log) returns $(p_i - p_{i-1})$ are negatively autocorrelated.



1-sec and 2-sec autocorrelations over 10min windows





$\Rightarrow$ MMN cannot be i.i.d.!

$\Rightarrow$ Noise properties change locally!

- Huge literature on efficiently estimating $\sigma^2_{\varepsilon^*}$
  - kernel estimators (Barndorff-Nielsen et al, 2008), pre-averaging (Jacod et al, 2009), MLE (Ait-Sahalia et al 2005), multi-scale estimators (Zhang 2006), spectral estimators (Reiss, 2011), ...

But:

- Assumptions on noise statistically motivated!
- Missing link to microstructure theory and trading behavior
  - Exceptions: Diebold/Strasser (2013), Chaker (2013), Li/Xie/Zeng (2016)
- All approaches rely on the classical RW+Noise decomposition

$$p_i = p_i^* + \epsilon_i, \quad i = 1, \dots, n.$$

A Model With Information/Trading Feedback

- Idea: Model with mis-pricing component:

$$p_i = p_{i-1} - \alpha(p_{i-1} - p^*_{i-1}) + \epsilon_i$$

- Changes in observed prices are caused by two sources:
  - Market microstructure noise
  - Mis-pricing component due to deviations between observed prices and efficient prices ("error correction")
- Reasoning:
  - "Non-informational" shocks cause "mis-pricing"
  - Prices are permanently in dis-equilibrium
- Speed by which observed prices react to inherent mis-pricing governed by α $\Rightarrow$ Measuring "market efficiency"

- Model captures two fundamental market regimes:
  - Mis-pricing removed by "contrarian behavior"
    $\Rightarrow$ negative autocorrelations in observed returns
  - Mis-pricing enforced by "momentum behavior"
    $\Rightarrow$ positive autocorrelations in observed returns
- State of market driven by relationship between
  - speed of price reversion α,
  - noise-to-signal ratio.

Why important?

- Model opens up channels for market microstructure foundations of HF-based volatility estimation.
- HF-based assessment of market efficiency.
- Statistical implications!

Outline

1. Introduction
2. A Model with Information Feedback
3. Two Market Regimes
4. Estimation
5. Empirical Evidence
6. Model Generalization
7. Conclusions

2. A Model with Information Feedback

Setup

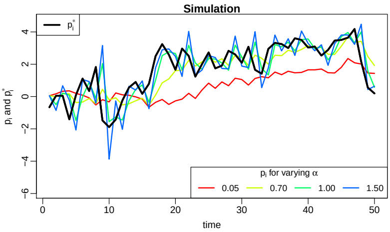
- Model in discrete time, i.e., $i \in \{0, 1, 2, \dots, n\}$ with $n = T/\Delta$.
- Observed log prices p_i are assumed to be driven by

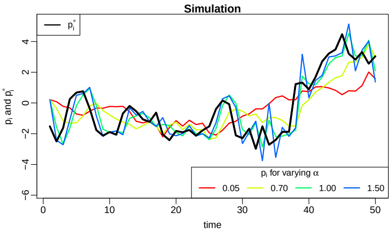
$$p_{i+1} = p_i - \alpha \underbrace{(p_i - p_i^*)}_{:=\mu_i} + \epsilon_{i+1}, \quad 0 < \alpha < 2, \quad \epsilon_{i+1} \overset{iid}{\sim} N(0, \sigma_\epsilon^2),$$

$$p_{i+1}^* = p_i^* + \varepsilon_{i+1}^*, \quad \varepsilon_{i+1}^* \overset{iid}{\sim} N(0, \sigma_{\varepsilon^*}^2), \quad \mathbb{E}[\epsilon_{i+1}\varepsilon_{i+1}^*] = 0.$$

- $\mu_i := p_i - p_i^*$ mis-pricing component.
- α: speed of price reversion

Simulations of p_i and p_i^* for different α





Alternative Representation

- Model can be written as

$$\begin{aligned} p_i &= p_i^* + \mu_i, \\ p_i^* &= p_{i-1}^* + \varepsilon_i^*, \quad \varepsilon_i^* \overset{i.i.d.}{\sim} N(0, \sigma_{\varepsilon^*}^2), \\ \mu_i &= (1-\alpha)\mu_{i-1} + \epsilon_i^\mu, \end{aligned}$$

where $\epsilon_i^\mu := \epsilon_i - \varepsilon_i^* \overset{iid}{\sim} WN(0, \sigma_\mu^2)$ with $\sigma_\mu^2 := \sigma_{\varepsilon^*}^2 + \sigma_\epsilon^2$.

- μ_i follows mean zero AR(1) process with

$$\mathbb{V}[\mu_i] = \frac{\sigma_\mu^2}{\alpha(2-\alpha)}$$

- The error covariance matrix Σ is given by

$$\Sigma := \begin{bmatrix} \mathbb{E}[(\varepsilon_i^{\mu})^2] & \mathbb{E}[\varepsilon_i^{\mu}\varepsilon_i^{*}] \\ \mathbb{E}[\varepsilon_i^{\mu}\varepsilon_i^{*}] & \mathbb{E}[(\varepsilon_i^{*})^2] \end{bmatrix} = \begin{bmatrix} \mathbb{E}[\mu_i^2] & \mathbb{E}[\mu_i\varepsilon_i^{*}] \\ \mathbb{E}[\mu_i\varepsilon_i^{*}] & \mathbb{E}[(\varepsilon_i^{*})^2] \end{bmatrix}$$
$$= \begin{bmatrix} \sigma_{\varepsilon}^2 + \sigma_{\varepsilon^*}^2 & -\sigma_{\varepsilon^*}^2 \\ -\sigma_{\varepsilon^*}^2 & \sigma_{\varepsilon^*}^2 \end{bmatrix}$$

with $\mathbb{E}[\varepsilon_i^{*}\epsilon_{i-h}^{\mu}] = 0 \quad \forall\, h.$

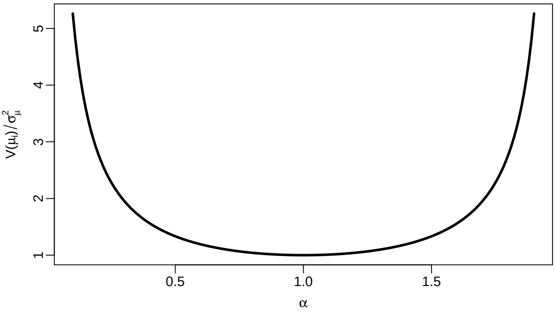
- Observed returns $r_i = p_i - p_{i-1}$ are then given by

$$r_i = -\alpha\mu_{i-1} + \epsilon_i, \quad \epsilon_i \overset{iid}{\sim} N(0, \sigma_{\epsilon}^2),$$

with

$$\mathbb{E}[\epsilon_i\mu_{i+h}] = (1-\alpha)^h\sigma_{\epsilon}^2 \quad \forall h \geq 0$$
$$\mathbb{E}[\epsilon_i\mu_{i-h}] = 0 \quad \forall h > 0$$
$$\mathbb{E}[\epsilon_i\epsilon_i^{\mu}] = \sigma_{\epsilon}^2$$
$$\mathbb{E}[\epsilon_i\epsilon_{i-h}^{\mu}] = 0 \quad \forall h \neq 0$$

Illustration of $\mathbb{V}[\mu_i]/\sigma_\mu^2$ depending on α for $\alpha \in (0, 2)$.



Special Case $\alpha = 1$ (Perfect "Efficiency")

- For $\alpha = 1$ we obtain

$$p_{i+1} = p_i^* + \epsilon_{i+1} = p_{i+1}^* + \epsilon_{i+1}^\mu,$$

where $\epsilon_i^\mu := \epsilon_i - \varepsilon_i^* = p_i - p_i^* = \mu_i$ is iid with

$$\mathbb{V}[\epsilon_i^\mu] = \sigma_\mu^2 = \sigma_\epsilon^2 + \sigma_{\varepsilon^*}^2$$
$$\mathbb{E}[\varepsilon_i^* \epsilon_i^\mu] = \mathbb{E}[(p_i^* - p_{i-1}^*)\epsilon_i^\mu] = -\sigma_{\varepsilon^*}^2$$

$\Rightarrow$ RW plus endogenous iid noise!

$\Rightarrow$ $\mathbb{E}[r_i, r_{i-1}] = -\sigma_\varepsilon^2$

$\Rightarrow$ Endogeneity structurally built into the model!

3. Two Market Regimes

Return Variances

- The return variance is given by

$$\mathbb{V}[r_i] = \sigma_\varepsilon^2 + \alpha^2 \mathbb{V}[\mu_i] = \frac{1}{2-\alpha}(2\sigma_\varepsilon^2 + \alpha\sigma_{\varepsilon^*}^2)\,.$$

implying $\mathbb{V}[r_i] \geq \sigma_\epsilon^2$.

- We define the *noise-to-signal ratio* λ as

$$\lambda = \sigma_\varepsilon^2 / \sigma_{\varepsilon^*}^2\,.$$

- Then, the unconditional return variance is given by

$$\mathbb{V}[r_i] \;=\; \sigma_{\varepsilon^*}^2\,\frac{2\lambda\,+\,\alpha}{2-\alpha}$$

- It follows that

$$
\begin{aligned}
&\mathbb{V}[r_i] \leq \mathbb{V}[r_i^*] \qquad \text{if} \qquad \lambda \leq 1-\alpha, \\
&\mathbb{V}[r_i] > \mathbb{V}[r_i^*] \qquad \text{otherwise.}
\end{aligned}
$$

⇒ If proportion of "informational variance" (λ) is high, and p_i sluggishly follows p_i^*, changes in efficient price are passed over to the observed price in mitigated way.

Regimes in Return Autocovariances

- **Lemma.** Assume $\sigma_\varepsilon^2 > 0$, $0 < \alpha < 2$, and $h \geq 1$. Then,

$$\mathbb{C}\mathrm{ov}[r_i, r_{i-h}] \;=\; \psi(h-1)\;\sigma_{\varepsilon^*}^2\,\frac{(1-\alpha-\lambda)}{2-\alpha},$$

with $\psi(h-1) \;=\; \alpha\,(1-\alpha)^{h-1}$, and $\psi(0) = 1$, if $\alpha = 1$.

- **Corollary.** Assume $\sigma_\varepsilon^2 > 0$, $0 < \alpha < 2$, and $h \geq 1$.
  (i) If $0 < \alpha < 1$, then
  $sgn\{\mathbb{C}\mathrm{ov}[r_i, r_{i-h}]\} \;=\; sgn\{(1-\alpha) - \lambda\}$.
  (ii) If $\alpha = 1$, then $\mathbb{C}\mathrm{ov}[r_i, r_{i-1}] = -\sigma_\varepsilon^2 < 0$, and
  $\mathbb{C}\mathrm{ov}[r_i, r_{i-h}] = 0$, for $h > 1$.
  (iii) If $1 < \alpha < 2$, then $sgn\{\mathbb{C}\mathrm{ov}[r_i, r_{i-h}]\} = sgn\{(-1)^h\}$.

- $\mathbb{C}\mathrm{ov}[r_i, r_{i-h}] = 0$ holds as long as

$$\lambda = 1 - \alpha$$

Implications:

- If $\alpha = 1$ and there is noise ($\lambda > 0$) returns cannot be uncorrelated!
- As long there is noise ($\lambda > 0$), price updating must be sluggish ($\alpha < 1$) to ensure $\mathbb{C}\mathrm{ov}[r_i, r_{i-h}] = 0$!

Implications for the Realized Variance

- Consider log prices, $p_0, p_\Delta, \ldots, p_{i\Delta}, \ldots, p_T$ at equidistant points $i = 1, \ldots, T/\Delta - 1, T/\Delta$, with grid size Δ and $n = T/\Delta$ and $r_{i\Delta} = p_{i\Delta} - p_{(i-1)\Delta}$.
- The realized return variance measure at time T is given as

$$RV_T^\Delta = \sum_{i=1}^{T/\Delta} r_{i\Delta}^2.$$

Theorem. For $0 < \alpha < 1$, the expected time-T realized variance sampled at calendar time grid size Δ equals,

$$\langle p \rangle_T^\Delta = \; T \cdot \sigma^2_{\varepsilon^*} \; + \; T \cdot \sigma^2_{\varepsilon^*} \cdot \phi(\Delta) \; \frac{\lambda - (1-\alpha)}{(2-\alpha)} \, ,$$

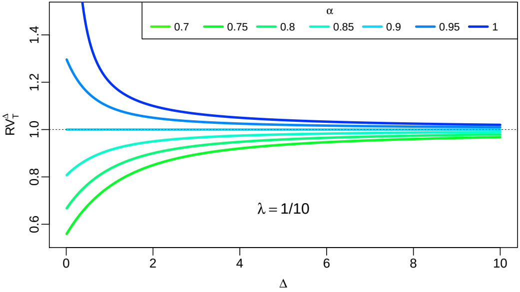
$$\text{with} \quad \phi(\Delta) = \; \frac{2}{\alpha \Delta} \left(1 - (1-\alpha)^\Delta \right) .$$

The mapping $\Delta \mapsto \phi(\Delta)$, from $\mathbb{R}_+$ into $\big(0, -\frac{2}{\alpha} \ln(1-\alpha)\big)$ is strictly decreasing with

(i) $\lim_{\Delta \to 0} \phi(\Delta) \; = \; -\frac{2}{\alpha} \ln(1-\alpha)$,
(ii) $\lim_{\Delta \to \infty} \phi(\Delta) \; = \; 0.$

Volatility Signature Plots

$$\text{(i) if} \quad \lambda > (1-\alpha), \quad \text{then} \quad T \cdot \sigma^2_{\varepsilon^*} < \langle p \rangle^{\Delta}_{T} \,,$$
$$\text{(ii) if} \quad \lambda < (1-\alpha), \quad \text{then} \quad T \cdot \sigma^2_{\varepsilon^*} > \langle p \rangle^{\Delta}_{T} \,.$$



4. Estimation

State-Space Representation

- Denote X_i as a state vector at i with $X_i := (\mu_i \quad \mu_{i-1} \quad \epsilon_i)$.
- Then, r_i can be written as

$$
\begin{aligned}
r_i &= F X_i, \\
X_i &= G X_{i-1} + w_i
\end{aligned}
$$

with $F = (0 \quad -\alpha \quad 1)$ and

$$
G = \begin{pmatrix} (1-\alpha) & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; w_i = \begin{pmatrix} \epsilon_i^{\mu} \\ 0 \\ \epsilon_i \end{pmatrix}, \; \Sigma_w = \begin{pmatrix} \sigma_\epsilon^2 + \sigma_{\varepsilon^*}^2 & 0 & \sigma_\epsilon^2 \\ 0 & 0 & 0 \\ \sigma_\epsilon^2 & 0 & \sigma_\epsilon^2 \end{pmatrix}.
$$

- Parameters can be estimated by maximum likelihood using the Kalman filter.

Alternative: Moment Estimation

- We can employ the unconditional moment restrictions

$$\phi_1(r_i;\alpha;\sigma_\varepsilon^2,\sigma_{\varepsilon^*}^2) = \sigma_{\varepsilon^*}^2 n - \sigma_\mu^2 n\phi(\Delta)\frac{1-\alpha-\lambda}{(2-\alpha)(\lambda+1)} - \sum_{i=1}^{n} r_i^2$$

$$\phi_2(r_i;\alpha;\sigma_\varepsilon^2,\sigma_{\varepsilon^*}^2) = r_i^2 - \frac{1}{2-\alpha}(2\sigma_\varepsilon^2 + \alpha\sigma_{\varepsilon^*}^2),$$

$$\phi_{2,h}(r_i;\alpha;\sigma_\varepsilon^2,\sigma_{\varepsilon^*}^2) = r_i r_{i-h} - \psi(h-1)\sigma_{\varepsilon^*}^2 \frac{1-\alpha-\lambda}{2-\alpha},$$

with $\psi(h) = \alpha(1-\alpha)^h \geq 0$ and $h = 1, 2, \ldots$.

- A GMM estimator can be formulated as

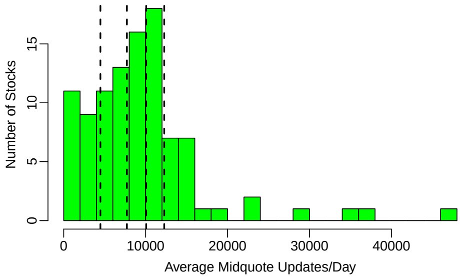
$$\hat{\theta}(\mathcal{W}) = \arg\min_{\theta} \left[\frac{1}{n} \sum_{i=1}^{n} \tilde{m}(r_i; \theta) \right]' \mathcal{W}_n \left[\frac{1}{n} \sum_{i=1}^{n} \tilde{m}(r_i; \theta) \right],$$

where $\theta = (\alpha, \sigma_{\varepsilon}^2, \sigma_{\varepsilon *}^2)'$, while $\tilde{m}(r_i; \theta) = (\phi_1(r_i; \theta), \phi_2(r_i; \theta), \phi_{2,1}(r_i; \theta), \phi_{2,2}(r_i; \theta), \dots)$ represents a set of model implied moment conditions, and $\mathcal{W}_n$ is a conforming positive definite weighting matrix.

5. Empirical Evidence

Data

- Data sampled from LOBSTER (`https://lobsterdata.com/`)
- Mid-quote returns from NASDAQ 100 constituents, first 40 trading days of 2014

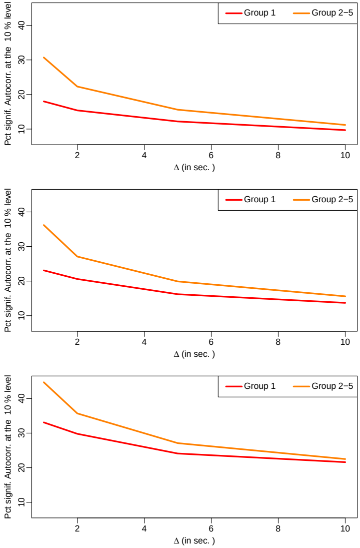


Empirical distribution of per-stock averages of daily mid-quote revisions

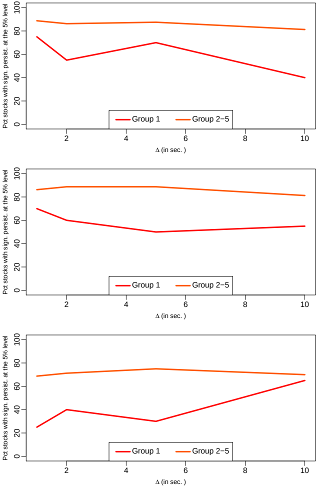
Significant first-order return autocorrelations

	Two-Sided			One-Sided $\widehat{Cor}(r_i, r_{i-1}) < 0$			One-Sided $\widehat{Cor}(r_i, r_{i-1}) > 0$			
$T = 30min \quad (N = 41,600)$										
Δ	1%	5%	10%	1%	5%	10%	1%	5%	10%	n
$1sec$	25.8	37.2	44.7	18.8	27.0	32.6	11.3	17.7	22.3	1800
$2sec$	15.7	27.3	35.7	9.5	17.3	22.8	10.3	18.4	23.8	900
$5sec$	9.4	19.7	27.1	4.5	9.8	14.4	8.4	17.3	24.4	360
$10sec$	6.4	15.4	22.5	2.3	6.6	10.9	6.6	15.9	23.1	180
$T = 10min \quad (N = 124,800)$										
Δ	1%	5%	10%	1%	5%	10%	1%	5%	10%	n
$1sec$	17.9	28.7	36.2	13.0	20.8	26.2	8.9	15.4	20.3	600
$2sec$	10.3	19.9	27.1	6.6	13.2	18.5	6.8	14.0	19.8	300
$5sec$	5.0	13.0	19.9	2.7	7.6	12.3	4.8	12.2	18.6	120
$10sec$	3.0	9.3	15.6	1.5	5.3	9.2	3.3	10.3	17.1	60
$T = 5min \quad (N = 249,600)$										
Δ	1%	5%	10%	1%	5%	10%	1%	5%	10%	n
$1sec$	14.1	23.9	30.7	10.1	17.2	22.2	7.5	13.6	18.2	300
$2sec$	7.9	15.9	22.3	5.1	10.9	15.7	5.4	11.4	16.7	150
$5sec$	3.5	9.6	15.6	1.9	6.0	10.4	3.3	9.6	15.5	60
$10sec$	1.5	6.0	11.2	0.8	3.8	7.5	1.8	7.3	13.6	30

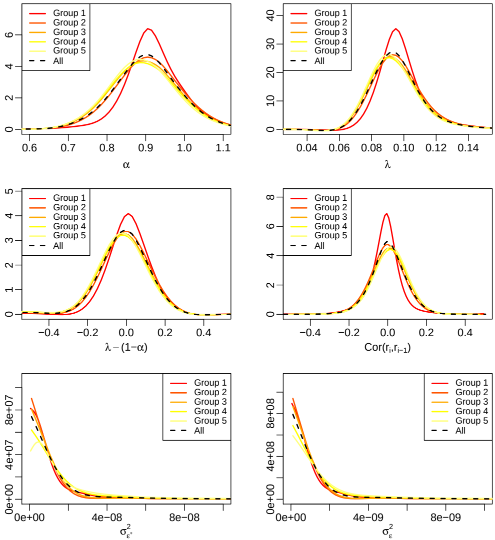
Prop. of sign. ACFs (10% level), $T \in \{5, 10, 30\}$min



Proportion of stocks with significant window-to-window ACF, $T \in \{5, 10, 30\}$min



Distribution of Parameter Estimates



Summary Statistics of Estimates

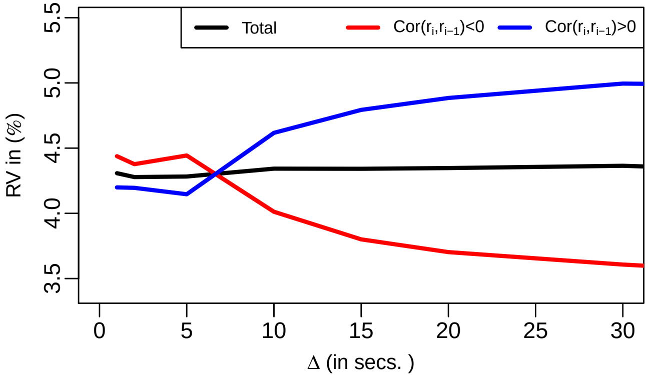
$T = 10min, \Delta = 1sec$

	$q5$	$q25$	$Median$	$Mean$	$q75$	$q95$	SD
$\hat{\alpha}$	0.778	0.864	0.909	0.909	0.963	1.052	0.105
$\hat{\lambda}$	0.077	0.088	0.096	0.102	0.106	0.129	0.073
$\hat{\lambda} - (1 - \hat{\alpha})$	−0.150	−0.048	0.005	0.011	0.069	0.181	0.124
$\widehat{\mathbb{C}or}(r_i, r_{i-1})$	−0.144	−0.056	−0.000	−0.008	0.041	0.121	0.081
$\hat{\sigma}^2_{\varepsilon^*}$ $(\cdot 10^{-8})$	0.060	0.140	0.254	0.563	0.517	1.850	1.219
$\hat{\sigma}^2_{\varepsilon}$ $(\cdot 10^{-8})$	0.005	0.013	0.025	0.055	0.050	0.180	0.125

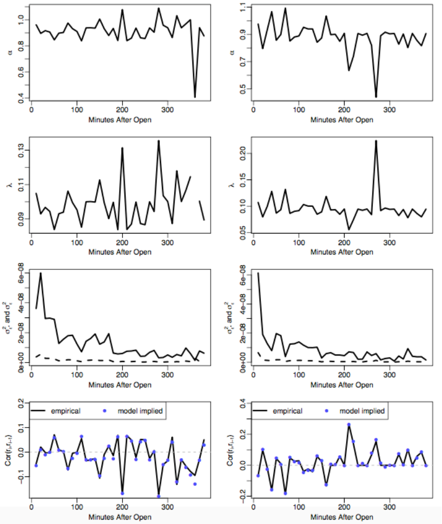
$T = 10min, \Delta = 2sec$

	$q5$	$q25$	$Median$	$Mean$	$q75$	$q95$	SD
$\hat{\alpha}$	0.726	0.846	0.905	0.891	0.961	1.056	0.138
$\hat{\lambda}$	0.072	0.086	0.095	0.104	0.106	0.138	0.101
$\hat{\lambda} - (1 - \hat{\alpha})$	−0.210	−0.070	−0.001	−0.004	0.066	0.188	0.158
$\widehat{\mathbb{C}or}(r_i, r_{i-1})$	−0.149	−0.053	0.000	0.002	0.058	0.151	0.091
$\hat{\sigma}^2_{\varepsilon^*}$ $(\cdot 10^{-8})$	0.116	0.279	0.516	1.177	1.058	3.875	2.670
$\hat{\sigma}^2_{\varepsilon}$ $(\cdot 10^{-8})$	0.009	0.026	0.048	0.118	0.101	0.394	0.287

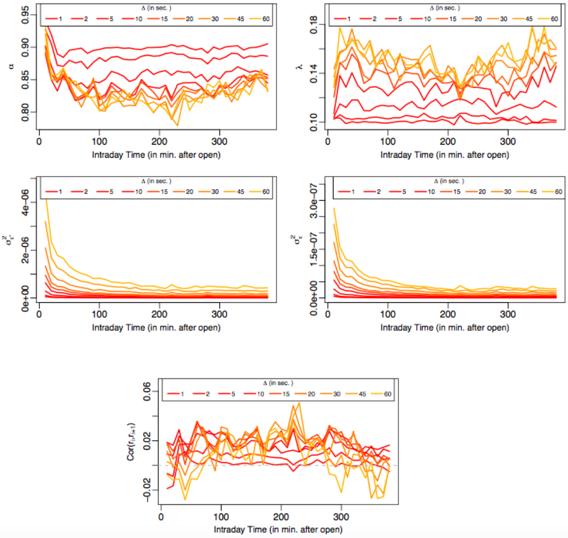
Volatility Signature Plots



TS Plots of Estimates of Yahoo and Microsoft



Intraday Seasonalities



Temporal Aggregation

- Let p_i, sampled at step size $\Delta \geq 1$, with observations for $i = 0, \Delta, 2 \cdot \Delta, \dots T - \Delta$, governed by

  $$p_{i+\Delta} = p_i - \alpha_\Delta (p_i - p_i^*) + \varepsilon_{i+\Delta,\Delta}, \quad p_{i+\Delta}^* = p_i^* + \hat{\varepsilon}_{i+\Delta,\Delta}^*$$

  where $\varepsilon_{i+\Delta,\Delta} \overset{iid}{\sim} \mathcal{N}(0, \sigma_{\varepsilon,\Delta}^2)$ and $\hat{\varepsilon}_{i+\Delta,\Delta}^* \overset{iid}{\sim} \mathcal{N}(0, \sigma_{\varepsilon^*,\Delta}^2)$ for all $i \in \{k \cdot \Delta, k = 0, 1, 2, 3, \dots\}$.

- Then, for $\Delta \geq 1$ and $0 < \alpha < 1$, we have,

  $$\alpha_\Delta = \ 1 - (1-\alpha)^\Delta, \quad \sigma_{\varepsilon,\Delta}^2 = \ g_\varepsilon \sigma_\varepsilon^2 + g_{\varepsilon^*} \sigma_{\varepsilon^*}^2, \quad \sigma_{\varepsilon^*,\Delta}^2 = \ \Delta \cdot \sigma_{\varepsilon^*}^2,$$

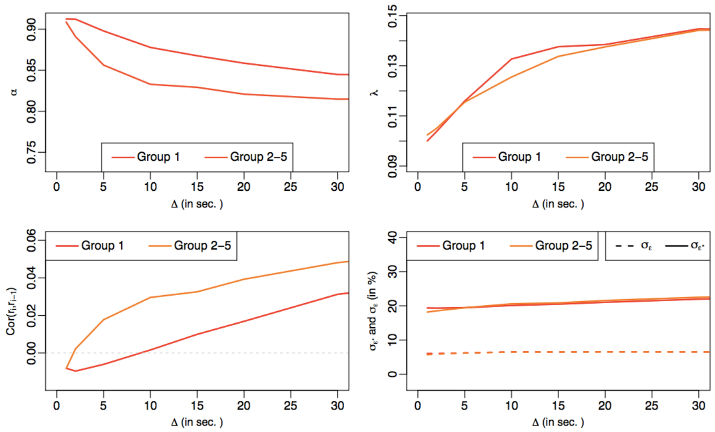
  with g_ε and g_{ε^*} denoting two functions depending on α and Δ.

- α_Δ is strictly increasing in Δ, with $\lim_{\Delta\to\infty} \alpha_\Delta = 1$
- For the noise-to-signal ratio λ_Δ for models estimated at lower frequencies we have

$$\lambda_\Delta = \frac{\sigma^2_{\varepsilon,\Delta}}{\sigma^2_{\varepsilon^*,\Delta}} = \frac{1}{\Delta}\Big(g_\varepsilon \lambda + g_{\varepsilon^*}\Big)$$

with $\lim_{\Delta\to\infty} \lambda_\Delta = 1$.

Temporal Aggregation



6. Model Generalization

- Assume the model

$$p_{i+1} = p_i - \alpha(p_i - p_i^*) + \epsilon_{i+1}, \qquad \epsilon_{i+1} \overset{iid}{\sim} N(0, \sigma_\epsilon^2),$$
$$p_{i+1}^* = p_i^* + \varepsilon_{i+1}^*, \quad \varepsilon_{i+1}^* \overset{iid}{\sim} N(0, \sigma_{\varepsilon^*}^2),$$

with $\mathbb{E}[\epsilon_{i+1}\varepsilon_{i+1}^*] = \gamma \neq 0$.

- For $\gamma = \alpha\sigma_{\varepsilon^*}^2$: Model by Amihud & Mendelson (1987):

$$p_i = p_{i-1} - \alpha(p_{i-1} - p_i^*) + \tilde{\varepsilon}_i,$$

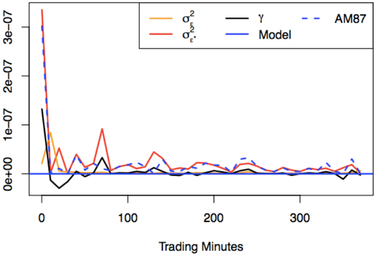
with $\tilde{\varepsilon}_i := \varepsilon_i - \alpha\varepsilon_i^*$ and $\mathbb{E}[\tilde{\varepsilon}_i\varepsilon_i^*] = 0$.

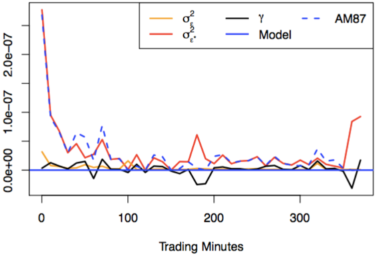
- For $\gamma = \sigma_{\varepsilon^*}^2$: RW-plus-iid-noise model

$$p_i = p_i^* + \varepsilon_i \qquad \text{with } \mathbb{E}[\varepsilon_i\varepsilon_i^*] = 0.$$

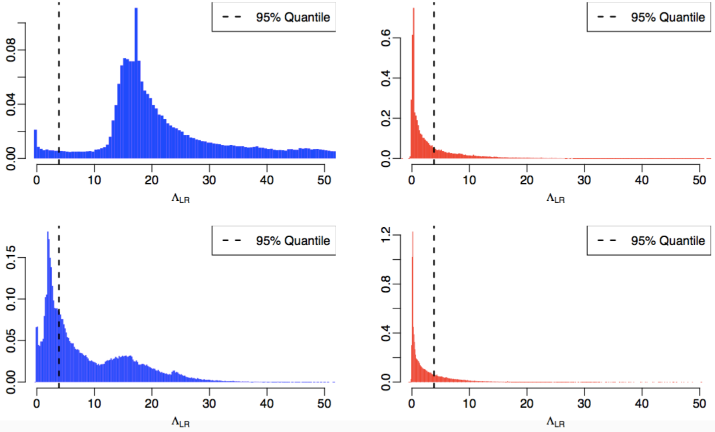
- For $\gamma = 0$, we obtain the original model.

Estimates for 2 days for AAPL





LR Tests for $H_0 : \gamma = \alpha\sigma^2_{\varepsilon*}$ (left) and $H_0 : \gamma = 0$ (right)



Top: $T = 10min$, $\Delta =$ 2secs; Bottom: $T = 10min$, $\Delta = 5$ secs

8. Conclusions

Conclusions

- Evidence for a model with information feedback
- Show mostly sluggish price updating due to mis-pricing
- Extent of market efficiency varies over time $\Rightarrow$ Identification of local states of "contrarian trading" and "momentum trading"
- Strong intraday and cross-sectional variation

Implications

- Channels for bridging the gap between high-frequency statistics and market microstructure theory
- New implications for volatility estimation
- Can be extended in various directions