Chapter 2
Noise-Adjusted Variance Estimators

High-Frequency Econometrics
National University Singapore, August 2019

Nikolaus Hautsch
University of Vienna
Outline 1

1. Market Microstructure Noise
   1.1 Sampling RV Under Noise
   1.2 ML Estimation

2. Kernel-Based Estimators
   2.1 Autocorrelation Adjustments
   2.2 Realized Kernels

3. Noise & Local Mispricing
   3.1 HF Dynamics with Mispricing
   3.2 Estimation & Empirics
Idealized Intraday Price Path
Real Intraday Price Path
Problem in practice: Market microstructure frictions!

In reality, we can only observe

\[ Y_{i\Delta_n} = X_{i\Delta_n} + U_{i\Delta_n}, \]

where

- \( Y_{i\Delta_n} \): observed (log) transaction price or quote at \( i\Delta_n \),
- \( X_{i\Delta_n} \): ”efficient” (fundamental) latent (log) price at \( t_{i\Delta_n} \),
- \( U_{i\Delta_n} \sim WN \) captures so-called market microstructure ”noise” (bid-ask spread, price discreteness ...).
Intuition

Under discrete sampling with noise:

\[ r_i := Y_{i \Delta_n} - Y_{(i-1) \Delta_n} := r_i^* + \varepsilon_i \]

with

\[ r_i^* := X_{i \Delta_n} - X_{(i-1) \Delta_n} \]
\[ \varepsilon_i := U_{i \Delta_n} - U_{(i-1) \Delta_n}. \]

Then, \( \lim_{\Delta_n \to 0} \mathbb{E}[(r_i^*)^2] = 0 \) while

\[ \lim_{\Delta_n \to 0} \mathbb{E}[\varepsilon_i^2] > 0! \]

Noise term dominates for \( \Delta_n \to 0! \)
Volatility Signature Plots

▶ In general, for \( n = \Delta_n^{-1} \), the bias of RV can be decomposed as

\[
\mathbb{E}[RV^{(n)}] = \mathbb{E} \left[ \sum_{i=1}^{n} r_i^* \right] + 2\mathbb{E} \left[ \sum_{i=1}^{n} \varepsilon_i r_i^* \right] + \mathbb{E} \left[ \sum_{i=1}^{n} \varepsilon_i^2 \right]
\]

\( \sigma^2 = \text{IV} \)
price/noise correlation bias
noise bias

\Rightarrow \text{Can be positive or negative!}

▶ If noise is exogenous, i.e., \( \mathbb{E}[\varepsilon_i r_i^*] = 0 \), and serially uncorrelated:

\[
\mathbb{E}[RV^{(n)}] = \sigma^2 + 2 \frac{\omega^2}{\Delta_n},
\]

where \( \omega^2 = \mathbb{E}[U_i^2 \Delta_n] \).
Volatility Signature Plots

Side note: Little $o$ and Big $\mathcal{O}$ Notation

- Write $X_n = \mathcal{O}(a_n)$ if there exist positive real numbers $M$ and $N$ such that
  
  $$|X_n| \leq M|a_n| \quad \forall \ n \geq N.$$

- Write $X_n = o(a_n)$ if

  $$\lim_{n \to \infty} X_n/a_n = 0.$$
Order in Probability $o_p$ and $O_p$

- $\{X_n\}$: sequence of random variables.
- Write $X_n = O_p(1)$ if, for every $\epsilon > 0$, there exists a constant $M$ and an integer $N$, such that
  \[ \mathbb{P}[|X_n| > M] \leq \epsilon \quad \forall \ n \geq N. \]
  \[ \Rightarrow \text{Write } X_n = O_p(a_n) \text{ if } X_n/a_n = O_p(1). \]
- Write $X_n = o_p(1)$ if
  \[ \lim_{n \to \infty} X_n = 0. \]
  \[ \Rightarrow \text{Write } X_n = o_p(a_n) \text{ if } X_n/a_n = o_p(1). \]
- Note: $X_n = o_p(1) \Rightarrow X_n = O_p(1)$ but $X_n = O_p(1) \not\Rightarrow X_n = o_p(1)$.
**Bias and Variance of RV**

- **Theorem (e.g., Aït-Sahalia et al, 2005).** For i.i.d. Gaussian noise and $\Delta_n = [T/n]$, the mean and variance of $RV^{(n)}$ are given by

\[
\mathbb{E}[RV^{(n)}] = \sigma^2 + \frac{2\omega^2}{\Delta_n},
\]

\[
\text{Var}[RV^{(n)}] = \frac{2(\sigma^4\Delta_n^2 + 4\sigma^2\Delta_n\omega^2 + 6\omega^4)}{T\Delta_n} - \frac{4\omega^4}{T^2}.
\]

- **Tradeoff:** Higher $\Delta_n$ decreases bias but increases variance.

  $\Rightarrow$ Minimize RMSE by balancing $\Delta_n$ (“sparse sampling”) for given $T$: e.g., Ait-Sahalia et al (2005), Bandi & Russell (2008)

- **For $\Delta_n \to 0$ it follows that $\mathbb{E}[RV^{(n)}] = \frac{2n\omega^2}{T} + o(n)$ and $\text{Var}[RV^{(n)}] = \frac{12\omega^4n}{T^2} + o(n)$.

  $\Rightarrow$ Hence, $RV^{(n)}/(2n)$ becomes an estimator of $\mathbb{E}[U^2] = \omega^2$!
Sub-Sampling RV

Consider sub-samples of prices \( Y \) given by

\[ \text{SS 1: } \{ Y_{1\Delta_n}, Y_{(K+1)\Delta_n}, Y_{(2K+1)\Delta_n}, \cdots, Y_{(n_1K+1)\Delta_n} \} \]

\[ \text{SS 2: } \{ Y_{2\Delta_n}, Y_{(K+2)\Delta_n}, Y_{(2K+2)\Delta_n}, \cdots, Y_{(n_2K+2)\Delta_n} \} \]

\[ \text{SS K: } \{ Y_{K\Delta_n}, Y_{2K\Delta_n}, Y_{3K\Delta_n}, \cdots, Y_{(n_KK+1)\Delta_n} \} \]

The sub-sampled RV is obtained by

\[ SRV_{Y}^{(n,K)} := \frac{1}{K} \sum_{k=1}^{K} RV_{Y}^{(n_k)} = \frac{1}{K} \sum_{k=1}^{K} \sum_{j=1}^{n_k} r_{j,n_k}^2, \]

where \( r_{j,n_k} := Y_{(jK+k)\Delta_n} - Y_{((j-1)K+k)\Delta_n} \) for \( k = 1 \ldots, K \).

Sub-sampling reduces the variance as all observations can be used.

In practice, sub-sampled 5min RV estimator is performing well! (Liu, Patton, Sheppard, 2015)
Outline I

1. Market Microstructure Noise
   1.1 Sampling RV Under Noise
   1.2 ML Estimation

2. Kernel-Based Estimators
   2.1 Autocorrelation Adjustments
   2.2 Realized Kernels

3. Noise & Local Mispricing
   3.1 HF Dynamics with Mispricing
   3.2 Estimation & Empirics
Maximum Likelihood Estimation

- Under i.i.d. normal noise we have

\[ Y_{i\Delta_n} = X_{i\Delta_n} + U_{i\Delta_n}, \quad U_{i\Delta_n} \overset{i.i.d.}{\sim} N(0, \omega^2), \]

with

\[ r^*_i = X_{i\Delta_n} - X_{(i-1)\Delta_n} \overset{i.i.d.}{\sim} N(0, \sigma^2 \Delta_n) \]

independent of \( U_{i\Delta_n} \).

- Then, \( r_i = r^*_i + U_{i\Delta_n} - U_{(i-1)\Delta_n} \) can be re-parameterized as a MA(1) process,

\[ r_i := \mu_i + \eta \mu_{i-1}, \]

where \( \mu_i \sim (0, \gamma^2) \).
The parameters $\eta$ and $\gamma^2$ can be identified by

$$\gamma^2(1 + \eta^2) = \nabla[r_i] = \sigma^2 \Delta_n + 2\omega^2$$
$$\gamma^2 \eta = \text{Cov}[r_i, r_{i-1}] = -\omega^2$$

Equivalently, we have

$$\gamma^2 = \frac{1}{2} \left\{ 2\omega^2 + \sigma^2 \Delta_n + \sqrt{\sigma^2 \Delta_n(4\omega^2 + \sigma^2 \Delta_n)} \right\}$$
$$\eta = \frac{1}{2\omega^2} \left\{ -2\omega^2 - \sigma^2 \Delta_n + \sqrt{\sigma^2 \Delta_n(4\omega^2 + \sigma^2 \Delta_n)} \right\}.$$
Consequently, the daily variance $\sigma^2$ as well as the microstructure noise can be estimated by

$$\hat{\sigma}^2 = \hat{I}V = \Delta_n^{-1}\hat{\gamma}^2(1 + \hat{\eta})^2,$$

$$\hat{\omega}^2 = -\hat{\gamma}^2\hat{\eta}$$

where $\hat{\gamma}^2$ and $\hat{\eta}$ are ML estimates based on a MA(1) process using high-frequency returns.

Proportion of the total return variance that is market microstructure-induced is

$$\pi = \frac{2\omega^2}{\sigma^2\Delta_n + 2\omega^2}$$

with

$$-\pi/2 = -\omega^2/(\sigma^2\Delta_n + 2\omega^2) = Cor [r_i, r_{i-1}]$$

corresponding to the first order autocorrelation.
Asymptotic Properties

- For $\Delta_n \to 0$ and $n = \Delta_n^{-1}$, we have

$$\left\{ \left( \frac{n^{1/4}}{\hat{\sigma}^2 - \sigma^2} \right) \right\} \overset{L}{\to} \mathcal{N}\left(0, \left( \begin{array}{cc} 8\omega^3 & 0 \\ 0 & 2\omega^4 \end{array} \right) \right)$$

- $n^{1/4}$ is fastest rate unless additional assumptions are made for $U$.
- Loss of efficiency relative to no-noise case.
- In no-noise case, i.e., the true value of $\omega^2$ is $\omega^2 = 0$:

$$n^{1/2}(\hat{\sigma}_{ML}^2 - \sigma^2) \overset{d}{\to} \mathcal{N}(0, 6\sigma^4)$$

- When $\omega^2$ is known a priori to be zero, then

$$n^{1/2}(\hat{\sigma}_{ML}^2 - \sigma^2) \overset{d}{\to} \mathcal{N}(0, 2\sigma^4).$$
Outline I

1. Market Microstructure Noise
   1.1 Sampling RV Under Noise
   1.2 ML Estimation

2. Kernel-Based Estimators
   2.1 Autocorrelation Adjustments
   2.2 Realized Kernels

3. Noise & Local Mispricing
   3.1 HF Dynamics with Mispricing
   3.2 Estimation & Empirics
An Autocorrelation-Based Estimator

- Zhou (1996)
- Augment RV by (realised) autocovariances, i.e.,

\[ RV_{AC}^{(n)} := \sum_{i=1}^{n} \left( r_i^2 + r_{i-1}r_i + r_ir_{i+1} \right) . \]

Under exogenous and serially uncorrelated noise we obtain:

\[ \mathbb{E}[r_i^2] = \mathbb{E}[r_i^{*2}] + \mathbb{E}[\varepsilon_i^2] = \sigma^2 \Delta_n + 2\omega^2, \]
\[ \mathbb{E}[r_{i-1}r_i] = \mathbb{E}[\varepsilon_{i-1}\varepsilon_i] = -\omega^2, \]
\[ \mathbb{E}[r_ir_{i+1}] = \mathbb{E}[\varepsilon_i\varepsilon_{i+1}] = -\omega^2. \]

\[ \Rightarrow \mathbb{E}[RV_{AC}^{(n)}] = \sigma^2 = IV. \]
Under exogenous and serially uncorrelated noise, it can be shown (see, e.g., Hansen/Lunde, 2006) that

\[
\mathbb{V}[RV_{AC}^{(n)}] = 8\omega^4 n + 8\omega^2 \sum_{i=1}^{n} \sigma_i^2 - 6\omega^4 + 6 \sum_{i=1}^{n} \sigma_i^4 + O(n^{-2}),
\]

with \( \sigma_i^2 := \int_{t(i-1)\Delta}^{t_i\Delta} \sigma^2(s)ds \) for \( i = 1, \ldots, n \), and

\[
\frac{RV_{AC}^{(n)} - IV}{\sqrt{8\omega^4 n}} \overset{d}{\rightarrow} \mathcal{N}(0, 1), \quad \text{as} \quad n \rightarrow \infty.
\]

Bias-corrected estimator, \( RV_{AC}^{(n)} \) has smaller asymptotic variance than \( RV^{(n)} \)!

But: Asymptotic variance is increasing in \( n \) \( \Rightarrow RV_{AC}^{(n)} \) is inconsistent!
In the absence of noise we have

\[ \mathbb{V}[RV_{AC}^{(n)}] \approx 6 \sum_{i=1}^{n} \sigma_i^4, \]

\[ \Rightarrow \text{Variance of } RV_{AC}^{(n)} \text{ about three times larger than that of } RV^{(n)} \text{ when } \omega^2 = 0. \]
Outline I

1. Market Microstructure Noise
   1.1 Sampling RV Under Noise
   1.2 ML Estimation

2. Kernel-Based Estimators
   2.1 Autocorrelation Adjustments
   2.2 Realized Kernels

3. Noise & Local Mispricing
   3.1 HF Dynamics with Mispricing
   3.2 Estimation & Empirics
Realized Autocovariation

Define for any process $Y$ and $Z$,

$$
\gamma_h(Z_{\Delta_n}, Y_{\Delta_n}) = \sum_{j=1}^{n} (Z_j \Delta_n - Z_{(j-1)\Delta_n})(Y_{(j-h)\Delta_n} - Y_{(j-h-1)\Delta_n}),
$$

$$
h = -H, \ldots, -1, 0, 1, 2, \ldots, H
$$

as the $h$-order cross-autocovariance between $\Delta_n$-increments in $Z$ and $Y$.

Then, $\gamma_h(Y_{\Delta_n}) = \gamma_h(Y_{\Delta_n}, Y_{\Delta_n})$ is the realized autocovariation with

$$
\gamma_h(Y_{\Delta_n}) = \gamma_h(X_{\Delta_n}) + \gamma_h(X_{\Delta_n}, U_{\Delta_n}) + \gamma_h(U_{\Delta_n}, X_{\Delta_n}) + \gamma_h(U_{\Delta_n})
$$
The realized kernel (Barndorff-Nielsen et al, 2008) is given by

\[ K(Y_{\Delta n}) = \gamma_0(Y_{\Delta n}) + \sum_{h=1}^{H} k \left( \frac{h - 1}{H} \right) \{ \gamma_h(Y_{\Delta n}) + \gamma_{-h}(Y_{\Delta n}) \}, \]

with \( k(x) \) for \( x \in [0, 1] \) being a weight function that is two times continuously differentiable with \( k(0) = 1 \) and \( k(1) = 0 \), and

\[ K(Y_{\Delta n}) = K(X_{\Delta n}) + K(X_{\Delta n}, U_{\Delta n}) + K(U_{\Delta n}, X_{\Delta n}) + K(U_{\Delta n}). \]
Asymptotics for a Simple Kernel ...

Assume that $U = 0$ and consider the simple kernel $K(\Delta_n) = K(Y_{\Delta_n}) = \gamma_0(X_{\Delta_n})$. Then, we have

$$\Delta_n^{-1/2} \left( \gamma_0(X_{\Delta_n}) - \int_0^T \sigma_u^2 du \right) \bigg| \rightarrow^d N(0, 2Z)$$

and

$$\Delta_n^{-1/2} \left( \gamma_0(X_{\Delta_n}) - \int_0^T \sigma_u^2 du \right) \xrightarrow{LX} MN(0, 2Z)$$

with $Z = \int_0^T \sigma_u^4 du$,

where $\xrightarrow{LX}$ denotes stable convergence and $MN()$ denotes a mixed-normal distribution.
Concept of Stable Convergence

Let $\mathcal{X}_n$ denote a random variate on $(\Omega, \mathcal{F}, P)$ and let $\mathcal{G}$ be a sub-$\sigma$-field of $\mathcal{F}$.

Then, $\mathcal{X}_n$ converges $\mathcal{G}$-stably in law to $\mathcal{X}$, written $\mathcal{X}_n \overset{L_\mathcal{G}}{\rightarrow} \mathcal{X}$, if and only if $(\mathcal{X}_n, \mathcal{Z}_n) \overset{L_\mathcal{G}}{\rightarrow} (\mathcal{X}, \mathcal{Z})$ for all $\mathcal{G}$-measurable random variables $\mathcal{Z}$ and some random variate $\mathcal{X}$.

When $\mathcal{G} = \sigma(Y)$, we will write $\overset{L_Y}{\rightarrow}$ in place of $\overset{L_\mathcal{G}}{\rightarrow}$. 
Consider the simple example where

$$\mathcal{X}_n = \Delta_n^{-1/2} \left( \gamma_0(X_{\Delta_n}) - \int_0^T \sigma_u^2 du \right) \overset{L^X}{\Rightarrow} MN(0, 2Z)$$

and

$$Z = \int_0^T \sigma_u^4 du$$

(1)

If our focus is on $\mathcal{X}_n/\sqrt{Z}$, and if $Z$ is $G$-measurable, then convergence $G$-stably in law implies that

$$\Delta_n^{-1/2} \left( \gamma_0(X_{\Delta_n}) - \int_0^T \sigma_u^2 du \right) / \sqrt{\int_0^T \sigma_u^4 du} \overset{L^X}{\Rightarrow} \mathcal{N}(0, 2).$$

This cannot be deduced from the convergence in law to a mixed Gaussian variable in (1) without stable convergence.
Zhou (1996) as Kernel Estimator

- Consider the Zhou (1996) estimator written as

\[ K(\Delta Y_n) = \gamma_0(\Delta Y_n) + \gamma_1(\Delta Y_n) + \gamma_{-1}(\Delta Y_n). \]

- Then, in the case of no noise we have

\[
\Delta_n^{-1/2} \left( \gamma_0(\Delta Y_n) + \gamma_1(\Delta Y_n) + \gamma_{-1}(\Delta Y_n) - \int_0^T \sigma_u^2 du \right) \\
\xrightarrow{L^\gamma} \text{MN} \left(0, 6 \int_0^T \sigma_u^4 du\right)
\]

- Recall: The main impact of the noise is through \( \gamma_0(U_{\Delta_n}) \) with \( \gamma_0(U_{\Delta_n}) \to \infty \) as \( n \to \infty \).
Note:

\[
\mathbb{E}[\gamma_0(U_{\Delta n}) + \gamma_1(U_{\Delta n}) + \gamma_{-1}(U_{\Delta n})] = 0 \\
\mathbb{V}[\gamma_0(U_{\Delta n}) + \gamma_1(U_{\Delta n}) + \gamma_{-1}(U_{\Delta n})] = 4\omega^2(2n - 1.5),
\]

i.e., autocovariance adjustments make estimator unbiased but do not remove inconsistency.

In general, higher-order autocovariances, \( h \geq 2 \), are noisy estimates of zero as

\[
\mathbb{E}[\gamma_h(U_{\Delta n}) + \gamma_{-h}(U_{\Delta n})] = 0 \\
\mathbb{V}[\gamma_h(U_{\Delta n}) + \gamma_{-h}(U_{\Delta n})] \propto n
\]

but including them (adquately weighted) can reduce the variance!

Idea: Higher-order autocovariances serve as control variables, similarly as in estimates of long-run variances.
E.g., one can show that

\[ \nabla [K(U_{\Delta n})] \simeq \frac{(n/H^2)}{8} \omega^4 \]

when the Bartlett kernel is employed.

Hence, increasing \( H \) with \( n \) makes it possible to reduce the variance induced by the noise!

Idea of realized kernels: "Correcting" \( \gamma_0 \) (= RV) by \( H \) autocovariances \( \gamma_h \), weighted by kernels \( K\left(\frac{h-1}{H}\right) \).

Bandwidth \( H \) determines magnitude of the correction.
Asymptotics

- For certain noise properties, Barndorff-Nielsen et al (2008) provide asymptotic distribution and convergence rates depending on choice of kernel (polynomial, Parzen, Tukey-Hanning) and bandwidth.
- Best rate of convergence that can be achieved is $n^{1/4}$.
- Optimal choice of bandwidth depends on noise-to signal ratio

\[ \xi^2 = \frac{\omega^2}{\sqrt{T \int_0^T \sigma_u^4 du}}. \]
Bandwidth Selection in Practice

- ξ can be estimated by

\[ \hat{\xi}^2 = \hat{\omega}^2 / \hat{IV}, \]

where \( \hat{IV} \) is a preliminary estimate of \( IV = \int_0^T \sigma_u^2 du \).

- Idea: It is rather easier to obtain a precise estimate of \( IV \) than of \( \sqrt{T \int_0^T \sigma_u^4 du} \).

⇒ Exploiting that \( IV^2 \approx T \int_0^T \sigma_u^4 du \) when \( \sigma_u^2 \) does not vary too much over the interval \([0, t]\).

- Hence, \( IV \) can be, e.g., estimated by

\[ \hat{IV} = RV_{\text{sparse}}, \]

which is a subsampled realised variance based on 20 minute returns.
\[ \hat{\omega}^2 \text{ can be, e.g., estimated by computing the realised variance using every } q\text{-th trade.} \]

\[ \text{By varying the starting point, we obtain } q \text{ distinct realised variances, } RV_{\text{dense}}^{\{1\}}, \ldots, RV_{\text{dense}}^{\{q\}} \text{ say.} \]

\[ \text{Hence,} \]

\[ \hat{\omega}^2_{(i)} = \frac{RV_{\text{dense}}^{\{i\}}}{2n^{\{i\}}}, \quad i = 1, \ldots, q, \]

where \( n^{\{i\}} \) is the number of non-zero returns that were used to compute \( RV_{\text{dense}}^{\{i\}} \).

\[ \text{Then, } \hat{\omega}^2 \text{ is the average of the } q \text{ estimates,} \]

\[ \hat{\omega}^2 = \frac{1}{q} \sum_{i=1}^{q} \omega^2_{\{i\}}. \]
Confidence intervals for daily increments to $[Y]$ for GE.

Alternative estimators

- Multi-scale estimators (Zhang et al, 2005)
- Pre-averaging estimators (Jacod et al, 2009)
- Spectral estimators (Reiss, 2011, Bibinger et al, 2014)
- Likelihood estimators (e.g., Xiu, 2010)
- Jump-robust estimators (Andersen et al, 2012)
- ...
Outline I

1. Market Microstructure Noise
   1.1 Sampling RV Under Noise
   1.2 ML Estimation

2. Kernel-Based Estimators
   2.1 Autocorrelation Adjustments
   2.2 Realized Kernels

3. Noise & Local Mispricing
   3.1 HF Dynamics with Mispricing
   3.2 Estimation & Empirics
Motivation

- HF-based volatility estimation mostly rely on Semi-Martingale + Noise decomposition
  \[ p_i = p_i^* + \varepsilon_i, \quad i = 1, \ldots, n \]

- Presumes that all relevant information is impounded *instantaneously*.

- Market microstructure noise \( \varepsilon_i \) is seen as (a mostly exogenous) nuisance factor.

- Noise properties result from statistical assumptions but are not economically motivated.
  \[ \Rightarrow \text{Missing connections to microstructure theory!} \]
Some evidence on Volatility Signature Plots
Implications for Noise

- IID noise $\Rightarrow$ downward signature plots; returns follow MA(1)
- Serially correlated noise and/or endogenous noise:
  $\Rightarrow$ upward and downward signature plots
- If noise is correlated and/or endogenous, the classical Semi-Martingale+Noise decomposition

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

is too simplistic!
- Missing link to microstructure theory: miscpricing!
A Model with Mispricing

Andersen, Archakov, Cebiroglu & Hautsch (2019, WP)

- Consider log returns $r_i := p_i - p_{i-1}$ over equidistant grid $i \in \{0, 1, 2, \ldots, T\}$.

- "Efficient" log price:
  \[ p_i^* = p_{i-1}^* + \varepsilon_i^*, \quad \varepsilon_i^* \overset{i.i.d.}{\sim} (0, \sigma_*^2), \]
  where $r_i^* = \varepsilon_i^*$ is the "efficient" return.

- Return dynamics:
  \[ r_i = -\alpha (p_{i-1} - p_{i-1}^*) + (\gamma \varepsilon_i^* + \varepsilon_i), \quad 0 < \alpha < 2, \]
  where $\varepsilon_i$ is i.i.d. with $\varepsilon_i^*$, $\mathbb{E}[\varepsilon_i] = \mathbb{E}[\varepsilon_i \varepsilon_i^*] = 0$ and $\mathbb{V}[\varepsilon_i] = \sigma_\varepsilon^2$. 
Uncorrelated Endogenous Pricing Errors

For $\alpha = 1$, we have

$$p_i = p_i^* + (\gamma - 1) \varepsilon_i^* + \varepsilon_i$$

$\gamma \neq 1$: "endogeneity"; temporary periods of over- or under-reaction

Return variance:

$$\nabla[r_i] = \nabla[p_i - p_{i-1}] = \sigma^2_* + 2 \left[ \gamma(\gamma - 1) + \lambda \right] \sigma_*^2,$$

where $\lambda := \sigma^2_\varepsilon/\sigma_*^2$ denotes the noise-to-signal ratio.

First-order return auto-covariance:

$$\text{Cov}[r_i, r_{i-1}] = \left[ \gamma(1 - \gamma) - \lambda \right] \sigma_*^2$$
Same values of $|\gamma - 1/2|$ imply same degree of smoothing; lead to identical return variances and auto-covariances

$\gamma = 1$ and $\gamma = 0$ imply both classical martingale plus noise model

$$p_i = p_i^* + \varepsilon_i$$

or

$$p_i = p_{i-1}^* + \varepsilon_i.$$

Observationally equivalent
3. Noise & Local Mispricing | 3.1 HF Dynamics with Mispricing

Error Correction Dynamics

▶ For $\gamma = 1$ and $\alpha \neq 1$, we have

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

▶ Risk-averse agents with incomplete information form unbiased, yet not error-free, expectations about efficient price $\Rightarrow$ mispricing!

▶ Feedback effects similarly to Kyle (1985) and Vives (1995)

▶ Second moments:

$$\nabla [r_i] = \sigma_\star^2 + \frac{2 \lambda}{2 - \alpha} \sigma_\star^2,$$

$$\text{Cov}[r_i, r_{i-h}] = - (1 - \alpha)^{h-1} \frac{\alpha \lambda}{2 - \alpha} \sigma_\star^2.$$  

$\Rightarrow$ Autocorrelations negative!

$\Rightarrow$ Variance inflated but increasing in $\alpha$ (faster information transmission and thus less return smoothing)
Random-Walk-plus-Noise Representation

- Define

\[ \mu_i := p_i - p_i^*, \quad \epsilon_i^{\mu} := (\gamma - 1) \epsilon_i^* + \epsilon_i \]

with

\[ \nabla[\epsilon_i^{\mu}] =: \sigma_{\mu}^2 = (\gamma - 1)^2 \sigma_*^2 + \sigma_\epsilon^2. \]

- Then, the model can be written as

\[ p_i = p_i^* + \mu_i, \]
\[ \mu_i = (1 - \alpha) \mu_{i-1} + \epsilon_i^{\mu} \]

with \( \mu_i \) serving as "standard" noise following AR(1) with

\[ \nabla[\mu_i] = \frac{\sigma_{\mu}^2}{\alpha(2 - \alpha)} = \frac{(\gamma - 1)^2 + \lambda}{\alpha(2 - \alpha)} \sigma_*^2, \quad 0 < \alpha < 2, \]

minimized for \( \alpha = 1. \)
3. Noise & Local Mispricing | 3.1 HF Dynamics with Mispricing

**Covariance structure:**

\[
\Sigma = \begin{bmatrix}
\mathbb{E}[(\varepsilon_i^\mu)^2] & \mathbb{E}[\varepsilon_i^\mu \varepsilon_i^*] \\
\mathbb{E}[\varepsilon_i^\mu \varepsilon_i^*] & \mathbb{E}[(\varepsilon_i^*)^2]
\end{bmatrix} = \begin{bmatrix}
(\gamma - 1)^2 + \lambda & \gamma - 1 \\
\gamma - 1 & 1
\end{bmatrix} \sigma_*^2.
\]

**Special case** \(\gamma = 1\):

\[
\Sigma = \begin{bmatrix}
\mathbb{E}[(\varepsilon_i^\mu)^2] & \mathbb{E}[\varepsilon_i^\mu \varepsilon_i^*] \\
\mathbb{E}[\varepsilon_i^\mu \varepsilon_i^*] & \mathbb{E}[(\varepsilon_i^*)^2]
\end{bmatrix} = \begin{bmatrix}
\lambda & 0 \\
0 & 1
\end{bmatrix} \sigma_*^2
\]

**Then, for any stationary noise dynamics, we have**

\[
r_i = p_i - p_{i-1} = \varepsilon_i^* + \Delta \mu_i, \quad \Delta \mu_i = \mu_i - \mu_{i-1}
\]

with

\[
\mathbb{E}[r_i^2] = \sigma_*^2 + \mathbb{E}[(\Delta \mu_i)^2] \geq \sigma_*^2.
\]
Lemma. Assume $\sigma^2 > 0$, $0 < \alpha \leq 1$, $0 \leq \gamma < 2$. Then,

$$\mathbb{V}[r_i] = \alpha^2 \mathbb{V}[\mu_i] + (\gamma^2 + \lambda)$$

$$= \sigma^2_* + \frac{2}{2 - \alpha} F_\alpha(\gamma, \lambda) \sigma^2_* ,$$

where $F_\alpha(\gamma, \lambda) = \gamma^2 + (1 - \gamma) \alpha + \lambda - 1$.

Corollary. Assume $\sigma^2 > 0$, and $0 < \alpha \leq 1$.

$$\mathbb{V}[r_i] \leq \mathbb{V}[r_i^*] \quad \text{if} \quad F_\alpha(\gamma, \lambda) \leq 0,$$
$$\mathbb{V}[r_i] > \mathbb{V}[r_i^*] \quad \text{otherwise}.$$
Lemma. Assume $\sigma_\epsilon^2 > 0$, $0 < \alpha \leq 1$, $h \geq 1$. Then,

$$\text{Cov}[r_i, r_{i-h}] = \psi(h-1) \sigma_*^2 \left[ 1 - \alpha (1 - \gamma) - (\gamma^2 + \lambda) \right]$$

$$= -F_\alpha(\gamma, \lambda) \psi(h-1) \sigma_*^2,$$

with $\psi(h-1) = \frac{\alpha}{2^\alpha} (1 - \alpha)^{h-1}$, and $\psi(0) = 1$, if $\alpha = 1$.

Corollary. Assume $\sigma_\epsilon^2 > 0$, $0 < \alpha \leq 1$, and $h \geq 1$.

$$\text{Cov}[r_i, r_{i-h}] \geq 0 \quad \text{if} \quad F_\alpha(\gamma, \lambda) \leq 0,$$

$$\text{Cov}[r_i, r_{i-h}] \leq 0 \quad \text{otherwise}.$$

- If $\gamma = 1$, positive return autocorrelation is infeasible.
- For $\gamma < 1$ (even $\gamma = 0$), positive return autocorrelation attainable.
- Zero autocorrelation for low values of $\gamma = 0$: $1 - \alpha \approx \lambda$
  $\Rightarrow$ Consistent with some noise, as long as impact is mitigated by sluggishness in prices!
Special Cases

- **γ = 1 and α = 0**: "Classical" model with \( E[\varepsilon_i \varepsilon_i^*] = 0 \) and,
  \[
  p_i = p_i^* + \varepsilon_i.
  \]
  yielding
  \[
  \nabla[r_i] = \sigma_*^2(1 + 2\lambda)
  \]
  \[
  \text{Cov}[r_i, r_{i-1}] = -\lambda \sigma_*^2.
  \]

- **Amihud-Mendelson ('87) / Information Delay Model with**

  \[
  p_i = (1 - \alpha) p_{i-1} + \alpha p_i^* + \varepsilon_i \quad \text{for } \gamma = \alpha
  \]

  \[
  p_i = (1 - \alpha) p_{i-1} + \alpha p_{i-1}^* + \varepsilon_i \quad \text{for } \gamma = 0,
  \]

  and

  \[
  \Sigma = \begin{bmatrix}
  \mathbb{E}[(\varepsilon_i^{\mu})^2] & \mathbb{E}[\varepsilon_i^{\mu} \varepsilon_i^*] \\
  \mathbb{E}[\varepsilon_i^\mu \varepsilon_i^*] & \mathbb{E}[(\varepsilon_i^*)^2]
  \end{bmatrix} = \sigma_*^2 \begin{bmatrix} 1 + \lambda & -1 \\
  -1 & 1 \end{bmatrix}
  \]
Realized Variance

Realized variance for interval $[0, T]$, using equidistant log-price observations, $p_{i\Delta}, i = 0, \ldots, n = n(\Delta) = \lceil T/\Delta \rceil$, is obtained as

$$RV_T(\Delta) = \sum_{i=1}^{n(\Delta)} (p_{i\Delta} - p_{(i-1)\Delta})^2 = \sum_{i=1}^{n(\Delta)} (r_{i\Delta})^2$$

$\alpha = \gamma = 1, \lambda = 0$: $\mathbb{E} [RV_T(\Delta)] = T \sigma^2$
$\gamma = 1$: $\mathbb{E} [RV_T(\Delta)] = T \left[1 + 2\lambda \right] \sigma^2$
$\gamma \neq 1, \alpha \neq 1$: $\mathbb{E} [RV_T(\Delta)] = T \left[1 + \frac{2}{2 - \alpha} F_\alpha(\gamma, \lambda) \right] \sigma^2$

$\Rightarrow$ Volatility signature plots can be up- or downward biased.
Outline I

1. Market Microstructure Noise
   1.1 Sampling RV Under Noise
   1.2 ML Estimation

2. Kernel-Based Estimators
   2.1 Autocorrelation Adjustments
   2.2 Realized Kernels

3. Noise & Local Mispricing
   3.1 HF Dynamics with Mispricing
   3.2 Estimation & Empirics
Statistical Identification

Problem: \( \lambda \) and \( \gamma \) are jointly only partially identified!

- Any combinations of \( \gamma \) and \( \lambda \) yielding same value for \( F_\alpha(\gamma, \lambda) = \gamma^2 - \alpha \gamma + \alpha + \lambda - 1 \) imply identical second-order return moments.

- Fixing \( \gamma \) provides unique identification of \( \lambda \).

- Conversely, fixing \( \lambda \) does not yield unique identification of \( \gamma \).

\( \Rightarrow \) Scenarios for which \( |\gamma - \alpha/2| \) are identical, are observationally equivalent.

- \( \alpha \) and \( \sigma^2 \) are (uniquely) identifiable.
3. Noise & Local Mispricing | 3.2 Estimation & Empirics

Regime I: Positive return autocorrelation
Regime II: Negative return autocorrelation
Estimation

Restate the model as

\[ r_i = -\alpha \mu_{i-1} + \tilde{\varepsilon}_i, \]
\[ \mu_i = (1 - \alpha) \mu_{i-1} + \varepsilon^\mu_i, \]

where \( \tilde{\varepsilon}_i = \gamma \varepsilon^*_i + \varepsilon_i \) and \( \varepsilon^\mu_i = \tilde{\varepsilon}_i - \varepsilon^*_i = (\gamma - 1) \varepsilon^*_i + \varepsilon_i. \)

Covariance structure:

\[ \mathbb{E}[\tilde{\varepsilon}_i \mu_{i+h}] = (1 - \alpha)^h (\gamma^2 + \lambda - \gamma) \sigma_*^2, \quad \forall h \geq 0, \]
\[ \mathbb{E}[\tilde{\varepsilon}_i \mu_{i-h}] = 0, \quad \forall h > 0, \]
\[ \mathbb{E}[\tilde{\varepsilon}_i \varepsilon^\mu_i] = [\gamma(\gamma - 1) + \lambda] \sigma_*^2, \]
\[ \mathbb{E}[\tilde{\varepsilon}_i \varepsilon^\mu_{i-h}] = 0, \quad \forall h \neq 0. \]
State-Space Representation

Denote $X_i$ as a state vector at $i$ with $X_i = (\mu_i, \mu_{i-1}, \tilde{\varepsilon}_i)'$.

Then, $r_i$ can be written as

$$r_i = FX_i, \quad X_i = GX_{i-1} + w_i$$

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

$$G = \begin{pmatrix}(1 - \alpha) & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad w_i = \begin{pmatrix} \varepsilon_i' \\ 0 \\ \tilde{\varepsilon}_i \end{pmatrix},$$

and

$$\Sigma_w = \begin{pmatrix}(\gamma - 1)^2 \sigma_*^2 + \sigma_{\tilde{\varepsilon}}^2 & 0 & \sigma_*^2[\gamma(\gamma - 1) + \lambda] \\ 0 & 0 & 0 \\ \sigma_*^2[\gamma(\gamma - 1) + \lambda] & 0 & \gamma^2 \sigma_*^2 + \sigma_{\tilde{\varepsilon}}^2 \end{pmatrix}.$$ 

Parameters can be estimated by maximum likelihood using Kalman filter (Kalman, 1960, 1963).
3. Noise & Local Mispricing | 3.2 Estimation & Empirics

Data

- Data sampled from LOBSTER database: https://lobsterdata.com/
- Nasdaq 100 constituents
- Mid-quote returns based on time grid $\Delta=2$ secs
- Estimates based on intraday windows of length of 10 min
- Total estimation period: first 61 trading days of 2014
- Identification restriction $\gamma = 0$ ($\gamma = \alpha$)
Parameter Estimates

- $\hat{\alpha} \approx 0.9$
- For majority of intervals, $\gamma < 1$.
- For non-trivial proportion of intervals, $\widehat{\text{Cov}}[r_i, r_{i-1}] < 0$ and $\gamma > 1$
- $\lambda_{\text{max}}$ typically below 0.4
- Positive return autocorrelation common for more liquid stocks
- For less liquid stocks, lower dependence, lower volatility and more rapid error correction
Implications

- Error correction implies that (calendar) time correlation becomes integral part of model specification.
- Noise is neither exogenous from efficient price nor independent of sampling frequency
  ⇒ Evidence against setting for in-fill asymptotics.
- Model naturally cast in discrete time.
- Evidence for non-martingale behavior for returns
- Interplay between
  (i) temporal correction of pricing errors due to noise
  (ii) incomplete incorporation of news ("endogeneity")
  (iii) noise-to-signal ratio
- Channels for bridging the gap between high-frequency statistics and market microstructure theory