

An Anatomy of Asset Returns

Roberto Renò ¹ **Shuping Shi** ²

¹ESSEC Business School

²Macquarie University

IIKI 2026 and AI-DEFIT 2026 — Bled, Slovenia — 30 September 2026

The Ubiquitous Itô Semimartingale

High frequency econometrics rests on one decomposition of log price changes:

$$\underbrace{dX_t}_{\text{return}} = \underbrace{b_t dt}_{\text{drift}} + \underbrace{\sigma_t dW_t}_{\text{diffusion}} + \underbrace{dJ_t}_{\text{jumps}}$$

The Ubiquitous Itô Semimartingale

High frequency econometrics rests on one decomposition of log price changes:

$$\underbrace{dX_t}_{\text{return}} = \underbrace{b_t dt}_{\text{drift}} + \underbrace{\sigma_t dW_t}_{\text{diffusion}} + \underbrace{dJ_t}_{\text{jumps}}$$

- Volatility σ_t dominates at first order.

The Ubiquitous Itô Semimartingale

High frequency econometrics rests on one decomposition of log price changes:

$$\underbrace{dX_t}_{\text{return}} = \underbrace{b_t dt}_{\text{drift}} + \underbrace{\sigma_t dW_t}_{\text{diffusion}} + \underbrace{dJ_t}_{\text{jumps}}$$

- **Volatility** σ_t dominates at first order.
- **Drift** b_t and **leverage** (σ_t co-varying with W_t) appear only at second order, entangled with each other.

The Ubiquitous Itô Semimartingale

High frequency econometrics rests on one decomposition of log price changes:

$$\underbrace{dX_t}_{\text{return}} = \underbrace{b_t dt}_{\text{drift}} + \underbrace{\sigma_t dW_t}_{\text{diffusion}} + \underbrace{dJ_t}_{\text{jumps}}$$

- **Volatility** σ_t dominates at first order.
- **Drift** b_t and **leverage** (σ_t co-varying with W_t) appear only at second order, entangled with each other.
- **Jumps** J_t confound the continuous/discontinuous split.

Related Literature: Components

Four strands of high-frequency econometrics, each tackling one component in isolation:

- **Realized variance** (jump-robust):

(Andersen et al., 2003; Barndorff-Nielsen and Shephard, 2003, 2004; Andersen et al., 2012, 2023a)

- **Jumps**: truncation, testing, identification

(Mancini, 2001; Barndorff-Nielsen and Shephard, 2006; Bollerslev et al., 2008; Mancini, 2009; Lee and Mykland, 2008; Aït-Sahalia and Jacod, 2012; Jacod and Todorov, 2009; Li et al., 2017, 2019; Zhao et al., 2025)

- **Drift** at high frequency

(Bandi and Phillips, 2003; Christensen et al., 2020; Laurent and Shi, 2020; Mancini, 2023; Andersen et al., 2024; Laurent et al., 2026; Kolokolov et al., 2025; Shi and Phillips, 2025; Benvenuti and Renò, 2025; Li et al., 2026)

- **Leverage** from high-frequency data

(Bandi and Renò, 2012; Aït-Sahalia et al., 2013; Wang and Mykland, 2014; Sanfelici et al., 2015; Kalnina and Xiu, 2017; Aït-Sahalia et al., 2017; Christensen et al., 2022; Li et al., 2022; Toscano et al., 2024; Todorov, 2021; Bandi and Renò, 2025)

Gap

No unified framework jointly identifies all **four components**.

Our Contributions

- A new framework: realized autocovariances of power transformed semi-returns.

Our Contributions

- A new framework: realized autocovariances of power transformed semi-returns.
- The half power ($c = 1/2$) transformation unpacks the return into:
 - the continuous component of the log return;
 - the jump component, with no truncation threshold;
 - the integrated leverage coefficient;
 - a jump robust integrated volatility.

Related Literature

- Semi-returns, realized semicovariances

(Barndorff-Nielsen et al., 2010; Patton and Sheppard, 2015; Bollerslev et al., 2020a; Jacod and Meddahi, 2024)

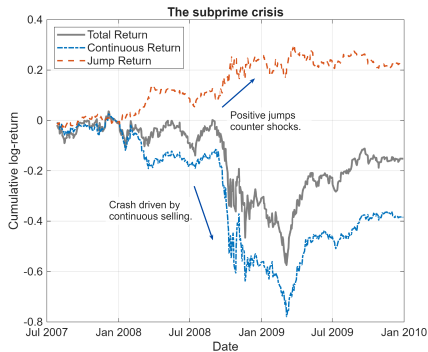
- High-frequency autocovariance

(Li and Linton, 2022; Andersen et al., 2023b; Laurent et al., 2024; Kolokolov et al., 2024; Li and Yang, 2025)

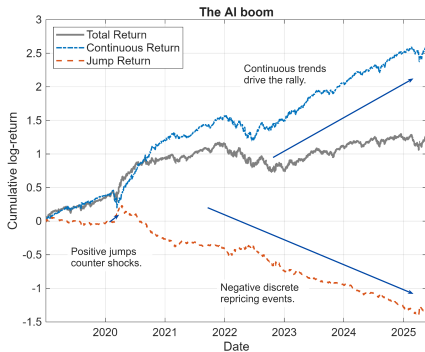
- Non-central limit theorems

(Kinnebrock and Podolskij, 2008; Jacod and Protter, 2012; Bollerslev et al., 2020b)

What If We Could Unpack the Return?



(a) S&P 500 E-mini: subprime crisis



(b) Nasdaq E-mini: COVID and AI boom

Furthermore, we leverage decomposed returns to study overnight and intraday dynamics.

Methodology and Theory

The Data Generating Process

Log price X_t is an Itô semimartingale:

Log price

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + J_t$$

Volatility σ_t is itself an Itô semimartingale:

Stochastic volatility

$$\sigma_t = \sigma_0 + \int_0^t \tilde{b}_s ds + \int_0^t \tilde{\sigma}_s dW_s + \tilde{M}_t + \sum_{s \leq t} \Delta \sigma_s \mathbf{1}_{\{\Delta \sigma_s \neq 0\}}$$

- $\tilde{\sigma}_s$ loads volatility on price shocks dW_s : the **leverage effect**.
- $\tilde{M}_t$ orthogonal to W ; finite activity jumps for tractability.

Semi-Returns

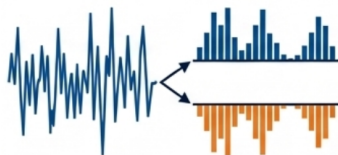
Decompose each return into positive and negative parts:

$$\Delta_i^n X = X_{i\Delta_n} - X_{(i-1)\Delta_n}$$

$$p(x) = \max\{x, 0\},$$

$$n(x) = \min\{x, 0\}$$

$$x = p(x) + n(x)$$



Two consecutive returns produce four sign combinations:

$$p-p, \quad n-n, \quad p-n, \quad n-p.$$

Power Transformed Autocovariances

The realized c -power autocovariances, fixed lag k , exponent $c > 0$:

Concordant terms

$$p(\Delta_i^n X)$$

Power Transformed Autocovariances

The realized c -power autocovariances, fixed lag k , exponent $c > 0$:

Concordant terms

$$p(\Delta_i^n X)^c$$

Power Transformed Autocovariances

The realized c -power autocovariances, fixed lag k , exponent $c > 0$:

Concordant terms

$$p(\Delta_i^n X)^c p(\Delta_{i-k}^n X)^c$$

Power Transformed Autocovariances

The realized c -power autocovariances, fixed lag k , exponent $c > 0$:

Concordant terms

$$\Delta_n^{1-c} \sum_{i=k+1}^n p(\Delta_i^n X)^c p(\Delta_{i-k}^n X)^c$$

Power Transformed Autocovariances

The realized c -power autocovariances, fixed lag k , exponent $c > 0$:

Concordant terms

$$\hat{P}^{(c)} = \Delta_n^{1-c} \sum_{i=k+1}^n p(\Delta_i^n X)^c p(\Delta_{i-k}^n X)^c$$

Power Transformed Autocovariances

The realized c -power autocovariances, fixed lag k , exponent $c > 0$:

Concordant terms

$$\hat{P}^{(c)} = \Delta_n^{1-\textcolor{red}{c}} \sum_{i=k+1}^n \textcolor{blue}{p}(\Delta_i^n X)^{\textcolor{red}{c}} \textcolor{blue}{p}(\Delta_{i-k}^n X)^{\textcolor{red}{c}}$$

$$\hat{N}^{(c)} = \Delta_n^{1-\textcolor{red}{c}} \sum_{i=k+1}^n [-\textcolor{brown}{n}(\Delta_i^n X)]^{\textcolor{red}{c}} [-\textcolor{brown}{n}(\Delta_{i-k}^n X)]^{\textcolor{red}{c}}$$

Power Transformed Autocovariances

Mixed (discordant) terms

$$\widehat{M}_1^{(c)} = \Delta_n^{1-c} \sum_{i=k+1}^n p(\Delta_i^n X)^c [-n(\Delta_{i-k}^n X)]^c$$

$$\widehat{M}_2^{(c)} = \Delta_n^{1-c} \sum_{i=k+1}^n [-n(\Delta_i^n X)]^c p(\Delta_{i-k}^n X)^c$$

- $\widehat{P}$: concordant positive ($p-p$); $\widehat{N}$: concordant negative ($n-n$).
- $\widehat{M}_1, \widehat{M}_2$: mixed signs ($p-n$ and $n-p$).
- All four are strictly positive measures.

First Order Asymptotics

Theorem (first order convergence)

Under regularity conditions, if $c \in (0, 2)$, as $\Delta_n \rightarrow 0$:

$$\left(\widehat{P}^{(c)}, \widehat{N}^{(c)}, \widehat{M}_1^{(c)}, \widehat{M}_2^{(c)} \right) \xrightarrow{P} \underbrace{\frac{2^{c-2}}{\pi} \Gamma\left(\frac{c+1}{2}\right)^2}_{\bar{\Lambda}^{(c)}} \int_0^T \sigma_s^{2c} ds \cdot \mathbf{1}'.$$

- All four estimators are first order *equivalent*.
- Local volatility σ_s dominates and enters *symmetrically* across the four sign combinations.
- To separate drift, leverage, and jumps we must go to **second order**.

Second Order: Concordant Pairs

$$\Delta_n^{-1/2} \begin{pmatrix} \hat{P}^{(c)} - \bar{\Lambda}^{(c)} \\ \hat{N}^{(c)} - \bar{\Lambda}^{(c)} \end{pmatrix} \xRightarrow{st} \underbrace{k_1(c) \begin{pmatrix} B^{(c)} \\ -B^{(c)} \end{pmatrix}}_{\text{drift bias}}$$

$$B^{(c)} \equiv \int_0^T b_u \sigma_u^{-1+2c} du \quad (\text{drift bias})$$

Second Order: Concordant Pairs

$$\Delta_n^{-1/2} \begin{pmatrix} \hat{P}^{(c)} - \bar{\Lambda}^{(c)} \\ \hat{N}^{(c)} - \bar{\Lambda}^{(c)} \end{pmatrix} \xRightarrow{st} \underbrace{k_1(c) \begin{pmatrix} B^{(c)} \\ -B^{(c)} \end{pmatrix}}_{\text{drift bias}} + \underbrace{k_2(c) \begin{pmatrix} L^{(c)} \\ -L^{(c)} \end{pmatrix}}_{\text{leverage bias}}$$

$$B^{(c)} \equiv \int_0^T b_u \sigma_u^{-1+2c} du \quad (\text{drift bias})$$

$$L^{(c)} \equiv \int_0^T \sigma_u^{2c-1} \tilde{\sigma}_u du \quad (\text{leverage bias})$$

Second Order: Concordant Pairs

$$\Delta_n^{-1/2} \begin{pmatrix} \hat{P}^{(c)} - \bar{\Lambda}^{(c)} \\ \hat{N}^{(c)} - \bar{\Lambda}^{(c)} \end{pmatrix} \xRightarrow{st} \underbrace{k_1(c) \begin{pmatrix} B^{(c)} \\ -B^{(c)} \end{pmatrix}}_{\text{drift bias}} + \underbrace{k_2(c) \begin{pmatrix} L^{(c)} \\ -L^{(c)} \end{pmatrix}}_{\text{leverage bias}} \\ + \underbrace{k_1(c) \begin{pmatrix} \zeta^{(c)} \\ -\zeta^{(c)} \end{pmatrix}}_{\text{diffusive risk}} + \underbrace{k^*(c) \begin{pmatrix} \zeta_P^{(c)} \\ \zeta_N^{(c)} \end{pmatrix}}_{\text{independent noise}}$$

$$B^{(c)} \equiv \int_0^T b_u \sigma_u^{-1+2c} du \quad (\text{drift bias})$$

$$L^{(c)} \equiv \int_0^T \sigma_u^{2c-1} \tilde{\sigma}_u du \quad (\text{leverage bias})$$

$$\zeta^{(c)} \equiv \int_0^T \sigma_u^{2c} dW_u \quad (\text{diffusive risk})$$

Second Order: Mixed Pairs

Mixed pairs $\widehat{M}_1$ and $\widehat{M}_2$ isolate the **leverage** term:

$$\Delta_n^{-1/2} \begin{pmatrix} \widehat{M}_1^{(c)} - \bar{\Lambda}^{(c)} \\ \widehat{M}_2^{(c)} - \bar{\Lambda}^{(c)} \end{pmatrix} \xRightarrow{st} \underbrace{k_4(c) \begin{pmatrix} -L^{(c)} \\ L^{(c)} \end{pmatrix}}_{\text{leverage bias}} + \underbrace{k^\dagger(c) \begin{pmatrix} \zeta_{M_1}^{(c)} \\ \zeta_{M_2}^{(c)} \end{pmatrix}}_{\text{independent noise}}$$

- No drift bias and no **diffusive risk** term: $\widehat{M}_1$ and $\widehat{M}_2$ deliver a clean leverage signal.

Why Jumps Do Not Dominate

Two mechanisms work together:

- **Power transformation.** For $c < 1$, the size of a jump after $(\cdot)^c$ is no longer the dominant element (Aït-Sahalia and Jacod, 2012).

Why Jumps Do Not Dominate

Two mechanisms work together:

- **Power transformation.** For $c < 1$, the size of a jump after $(\cdot)^c$ is no longer the dominant element (Aït-Sahalia and Jacod, 2012).
- **Bipower structure.** A jump is multiplied by its *neighboring* return, which is continuous with probability tending to one (Barndorff-Nielsen and Shephard, 2004).

Why Jumps Do Not Dominate

Two mechanisms work together:

- **Power transformation.** For $c < 1$, the size of a jump after $(\cdot)^c$ is no longer the dominant element (Aït-Sahalia and Jacod, 2012).
- **Bipower structure.** A jump is multiplied by its *neighboring* return, which is continuous with probability tending to one (Barndorff-Nielsen and Shephard, 2004).
- Combining the two: the jump $\times$ continuous interaction is $O_p(\Delta_n^{1-c/2})$, asymptotically negligible.

Bias Coefficients in c

- **Drift:**

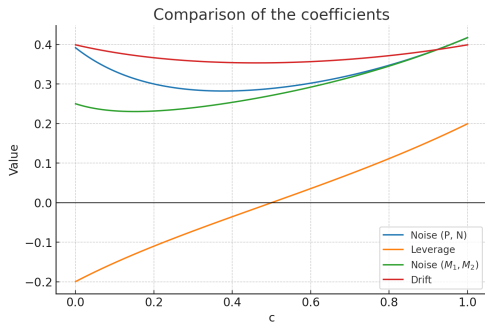
$$k_1(c) = \frac{c\Gamma(c)}{\sqrt{2\pi}}$$

- **Leverage:**

$$k_2(c) = \frac{(2c-1)c\Gamma(c)}{2\sqrt{2\pi}}$$

- **Noise (P, N):** $k^*(c)$

- **Noise (M_1, M_2):** $k^\dagger(c)$



At $c = 1/2$: **Leverage** = 0, while **drift** and noise remain.

The Magic of the Half Power ($c = 1/2$)

Setting $c = 1/2$ (geometric mean of semi-returns) collapses the bias terms:

- Drift bias becomes pure integrated drift:

$$B^{(1/2)} = \int_0^T b_u \sigma_u^0 du = \int_0^T b_u du.$$

- Concordant leverage coefficient vanishes:

$$k_2(1/2) = \frac{(2 \cdot \frac{1}{2} - 1) c \Gamma(c)}{2\sqrt{2\pi}} = 0.$$

- Mixed terms isolate pure integrated leverage:

$$L^{(1/2)} = \int_0^T \sigma_u^0 \tilde{\sigma}_u du = \int_0^T \tilde{\sigma}_u du.$$

The Four Key Estimators

Simple differences of the $c = 1/2$ measures cancel the dominant volatility term $\bar{\Lambda}^{(1/2)}$.

1. Continuous return

$$\begin{aligned}\hat{R}_T^C &= \sqrt{2} \Delta_n^{-1/2} \left(\hat{P}^{(1/2)} - \hat{N}^{(1/2)} \right) \\ &= \sqrt{2} \sum_i \left[\sqrt{p(\Delta_i^n X) p(\Delta_{i-k}^n X)} - \sqrt{[-n(\Delta_i^n X)][-n(\Delta_{i-k}^n X)]} \right]\end{aligned}$$

- **Conditional** on the volatility path: asymptotically unbiased for R_T^C .
- **Unconditionally**: an estimator of the **integrated drift** $\int_0^T b_u du$.

The Four Key Estimators

2. Integrated leverage coefficient

$$\begin{aligned}\hat{L}_T &= 4\sqrt{2} \Delta_n^{-1/2} \left(\widehat{M}_2^{(1/2)} - \widehat{M}_1^{(1/2)} \right) \\ &= 4\sqrt{2} \sum_i \left[\sqrt{-n(\Delta_i^n X) p(\Delta_{i-k}^n X)} - \sqrt{p(\Delta_i^n X) [-n(\Delta_{i-k}^n X)]} \right]\end{aligned}$$

- Estimates $L_T = \int_0^T \tilde{\sigma}_u du$.

The Four Key Estimators

3. Discontinuous (jump) return

$$\hat{R}_T^J = (X_T - X_0) - \hat{R}_T^C$$

4. Robust integrated volatility

$$\widehat{IV}_T = \frac{1}{0.676} \left(\hat{P}^{(1/2)} + \hat{N}^{(1/2)} + \widehat{M}_1^{(1/2)} + \widehat{M}_2^{(1/2)} \right)$$

- Continuous and **jump** separated **without any truncation threshold**.
- $\widehat{IV}_T$ equals a bipower variation with powers $[\frac{1}{2}, \frac{1}{2}]$.

Summary of Asymptotics

Corollary

As $\Delta_n \rightarrow 0$ (conditionally on the volatility path; unconditionally for drift/leverage):

$$(1) \quad \widehat{IV}_T \xrightarrow{p} \int_0^T \sigma_s ds,$$

$$(2) \quad \widehat{R}_T^C \xRightarrow{st} \textcolor{blue}{R}_T^C + 0.419 \int_0^T \sigma_u d\tilde{W}_u,$$

$$(3) \quad \widehat{R}_T^J \xRightarrow{st} \textcolor{red}{J}_T - 0.419 \int_0^T \sigma_u d\tilde{W}_u,$$

$$(4) \quad \widehat{L}_T \xRightarrow{st} \int_0^T \textcolor{green}{\tilde{\sigma}}_u du + 1.248 \int_0^T \sigma_u d\tilde{W}'_u.$$

$\widehat{IV}_T$ is consistent; $\widehat{R}_T^C, \widehat{R}_T^J, \widehat{L}_T$ carry a conditionally Gaussian noise term. Errors in $\widehat{R}_T^C$ and $\widehat{R}_T^J$ are perfectly negatively correlated.

Consistency via long-span averaging $\rightarrow$

Bias Correction and Variance Reduction

Trimming: finite-sample bias correction

Problem

A jump at time i^* inflates $\pi_{i^*}^{(\cdot)}$ and $\pi_{i^*+k}^{(\cdot)}$, biasing each of $\hat{P}, \hat{N}, \hat{M}_1, \hat{M}_2$.

Solution: rank and drop

- Let $c_i = \max\{\pi_i^P, \pi_i^N, \pi_i^{M_1}, \pi_i^{M_2}\}$ with

$$\pi_i^P = \sqrt{p_i p_{i-k}}, \pi_i^N = \sqrt{n_i n_{i-k}}, \pi_i^{M_1} = \sqrt{p_i n_{i-k}}, \pi_i^{M_2} = \sqrt{n_i p_{i-k}}.$$

- Find the **largest n^* values of c_i and drop them:**

$$\hat{P}_{n^*}^{(1/2)} = \Delta_n^{1/2} \sum_{i \notin \mathcal{I}_{n^*}} \pi_i^P,$$

where $\mathcal{I}_{n^*}$ collects the indices of the n^* largest values of c_i .

Lag Averaging: variance reduction

At lag k , the trimmed continuous-return estimator is $\hat{R}_T^{C,n*}(k)$.

Average over $k = 1, \dots, K_{\max}$ to diversify the lag-specific noise:

$$\hat{R}_T^{C,\text{avg}} = \frac{1}{K_{\max}} \sum_{k=1}^{K_{\max}} \hat{R}_T^{C,n*}(k), \quad \hat{R}_T^{J,\text{avg}} = (X_T - X_0) - \hat{R}_T^{C,\text{avg}}.$$

(idem for $\hat{IV}, \hat{L}$).

Lag Averaging: variance reduction

Theorem

As $\Delta_n \rightarrow 0$, with $K_{\max} \geq 1$:

$$\begin{aligned}\Delta_n^{-1/2} \left(\hat{R}_T^{C, \text{avg}} - R_T^C \right) &\xRightarrow{st} \sqrt{\bar{V}_{RC}} \int_0^T \sigma_u d\tilde{W}_u, \\ \Delta_n^{-1/2} \left(\hat{R}_T^{J, \text{avg}} - R_T^J \right) &\xRightarrow{st} -\sqrt{\bar{V}_{RJ}} \int_0^T \sigma_u d\tilde{W}_u, \\ \Delta_n^{-1/2} \left(\hat{L}_T^{\text{avg}} - L_T \right) &\xRightarrow{st} \sqrt{\bar{V}_L} \int_0^T \sigma_u d\tilde{W}'_u.\end{aligned}$$

Variance constants:

$$\bar{V}_{RC} = \bar{V}_{RJ} = \underbrace{\frac{V_C}{K_{\max}}}_{\rightarrow 0} + \underbrace{\frac{K_{\max}-1}{K_{\max}} C_{RC}}_{\rightarrow C_{RC} > 0}, \quad \bar{V}_L = \frac{V_L}{K_{\max}} \rightarrow 0.$$

$$V_C = 0.419^2 \approx 0.176, \quad V_L = 1.248^2 \approx 1.557; \quad C_{RC} \approx 0.079 \text{ (cross-lag covariance; } C_L = 0).$$

Simulation Studies

Monte Carlo Design

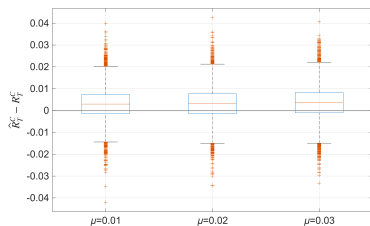
DGP

$$dX_t = b_t dt + f_t \sigma_t dW_t + dJ_{X,t}$$

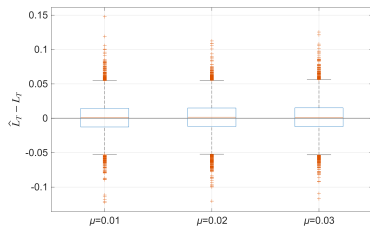
- Heston-type stochastic volatility, co-jumps in price and volatility.
- Correlated shocks $\mathbb{E}(dW_t dW_t^V) = \rho dt$, $\rho = -0.767$ (leverage).
- Intraday periodicity f_t (U shape); additive microstructure noise ($\xi = 0.5$).
- One-second latent grid over a day, $n_{\text{sec}} = 23,400$.
- Estimation at the five-minute frequency (78 returns per day).
- 10,000 Monte Carlo replications.

Robustness to the Drift

Varying the constant drift $b_t \in \{0.005, 0.01, 0.02\}$.¹



(a) $\hat{R}_T^C - R_T^C$



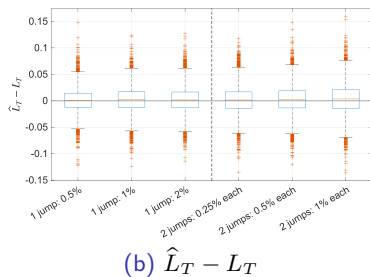
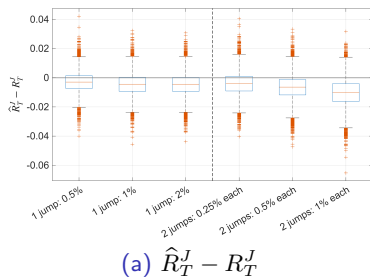
(b) $\hat{L}_T - L_T$

The continuous return and leverage estimators are essentially invariant to the drift level.

¹Jump design held fixed: one price jump per day of size 0.5%, co-jumping with volatility.

Robustness to Jumps

Left three boxes vary jump size; right three vary jump frequency:²

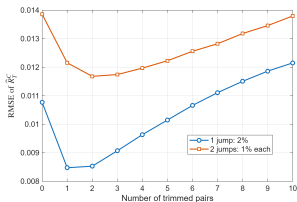


A mild bias in $\hat{R}_T^J$ grows slowly with jump size (jumps interacting with neighbors); leverage is insensitive. Dispersion ordering: $\hat{L}_T > \hat{R}_T^C \approx \hat{R}_T^J > \hat{IV}_T$.

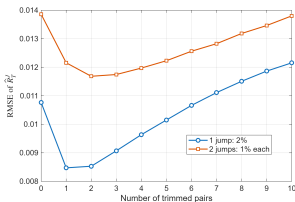
²Drift held fixed at $b_t = 0.01$; variance co-jumps with price in every scenario.

Trimming

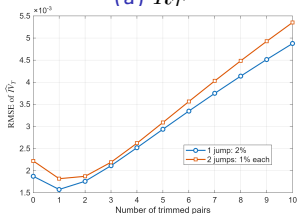
RMSE against the number of trimmed local products n^* :



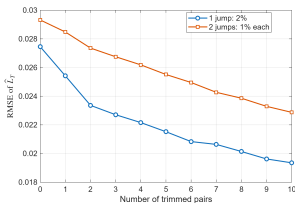
(a) $\hat{R}_T^C$



(b) $\hat{R}_T^J$



(c) $\hat{I}_T$

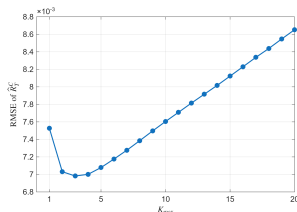


(d) $\hat{L}_T$

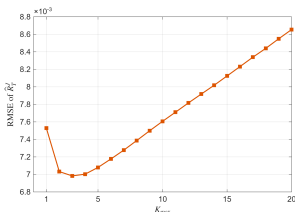
U shape with minimum at $n^* = 2$ for $\hat{R}^C$, $\hat{R}^J$, $\hat{I}_T$; monotone for $\hat{L}_T$. One 2% jump (circles); two 1% jumps (squares).

Trimming and Averaging

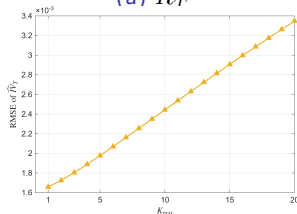
RMSE against the maximum lag $K_{\max}$ (trimming fixed at $n^* = 2$):



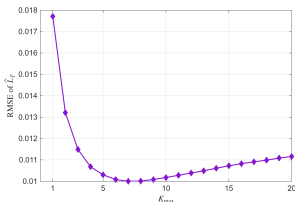
(a) $\hat{R}_T^C$



(b) $\hat{R}_T^J$



(c) $\hat{IV}_T$



(d) $\hat{L}_T$

Returns: U shape, optimum near $K_{\max} = 4$. Leverage: keeps falling. $\hat{IV}$ already consistent, so averaging is counterproductive: trim only.

Simulation Takeaways

- The four estimators are close to unbiased and informative at five-minute sampling.
- Robust to drift; only mildly sensitive to large jumps.

Recommended setting

- **Trim** at $n^* = 2 \Rightarrow$ removes jump-contamination bias.
- **Lag-average** up to $K_{\max} = 4 \Rightarrow$ reduces variance of $\hat{R}^C, \hat{R}^J, \hat{L}$.
- For $\widehat{IV}$: trim only (averaging is counterproductive).

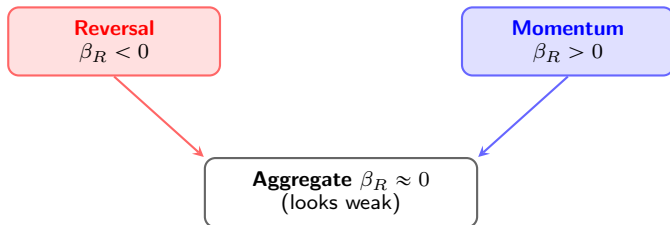
Used throughout the empirical analysis.

Empirical Application

The Overnight vs Intraday Puzzle

Two strands of literature, two opposite predictions for R_d^{Day} on R_d^{ON} :

- **Reversal**. Clientele rotation; constrained liquidity provision; uncertainty at the open.
(Lou et al., 2019; Akbas et al., 2022; Lu et al., 2023; Liu et al., 2025)
- **Momentum**. Gradual information assimilation; “frog in the pan.”
(Da et al., 2014; Gao et al., 2018)



Gap: aggregate returns mix two opposing forces. We *decompose* and test them separately.

Two Applications

Index Futures (time series)

- S&P 500 and Nasdaq 100 E-mini.
- Overnight and day sessions.
- Decompose: $\hat{R}^C, \hat{R}^J, \hat{IV}, \hat{L}$.
- Test **momentum** vs **reversal** channels.

U.S. Equities (cross section)

- 937 stocks from KIBOT/CRSP/Compustat.
- Firm-by-firm decomposition.
- Test the same channels stock-by-stock.

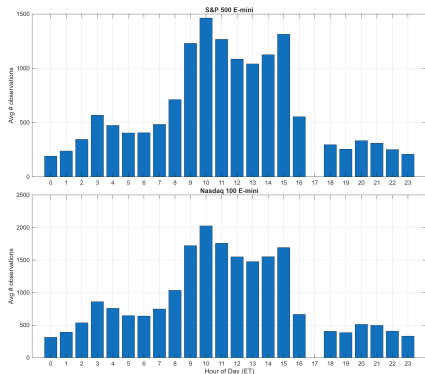
Two parallel exercises, two angles on the same overnight/intraday puzzle.

Index Futures: Time Series

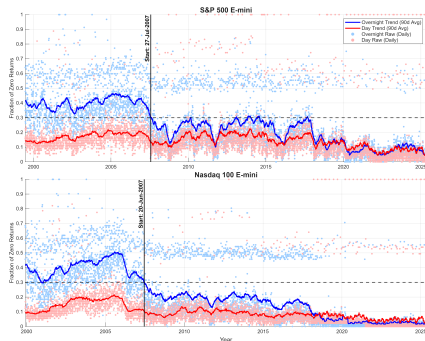
Futures Data

- **Assets:** S&P 500 and Nasdaq 100 E-mini futures.
- **Source:** Refinitiv Tick History, one-second frequency.
- **Sample:** 2007 to 31 May 2025.
- **Sessions:**
 - **Overnight:** 6:00 pm to 9:30 am, sampled every 10 minutes.
 - **Day:** 9:30 am to 4:00 pm, sampled every five minutes.
- Liquidity filters on zero-return fractions define the start date and remove stale days.

Liquidity Across the Clock



(a) Hourly trading activity

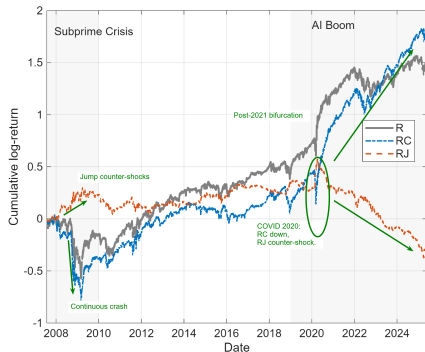


(b) Zero-return fraction over time

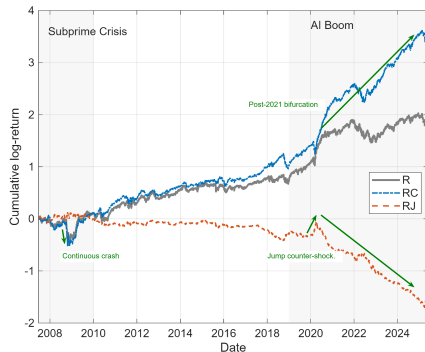
- We sample every 10 minutes overnight and every five minutes during the day.

(a) one-second obs per hour; (b) zero-return fraction at the 10/5-min sampling (90-day MA).

Cumulative Decomposition (Round the Clock)



(a) S&P 500

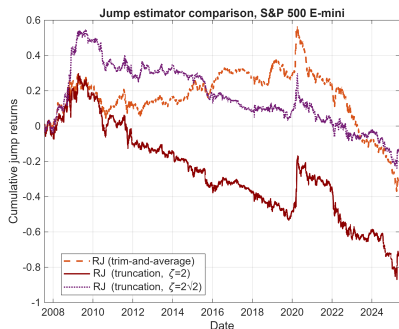


(b) Nasdaq 100

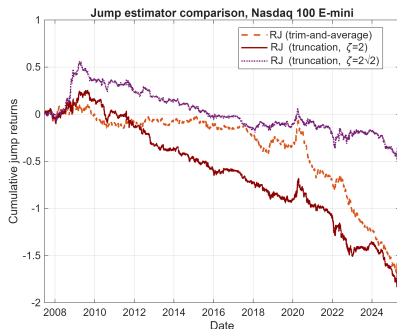
- Crises (2008, March 2020) are **continuous** driven, not jump driven.
- Positive **jumps** counter shocks (rescue and stimulus news).
- From 2020 the components **bifurcate**: continuous up, cumulative jump down.

Robustness: Jump Estimator vs Truncation

Cumulative jump returns: **half-power** estimator (no tuning) vs truncation (Mancini, 2009) at $\zeta \in \{2, 2\sqrt{2}\}$.



(a) S&P 500



(b) Nasdaq 100

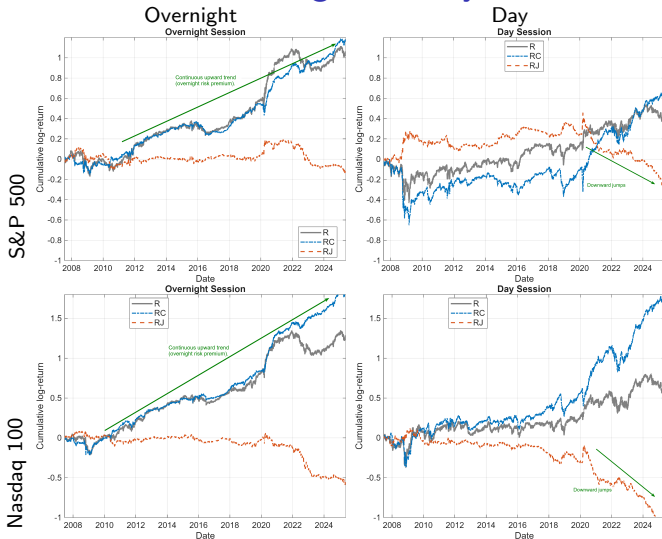
The bifurcation is **not** a construction artifact.

Truncation: Jump iff $|\Delta X_{d,i} - \Delta X_{d,i-1}| > \zeta \sqrt{\text{Med}_{d-1}} \Delta_n^{1/2}$ (Mancini, 2009; Andersen et al., 2023a).

Why the Drift-Jump Divergence?

- Basele et al. (2025): **speculative bubbles** in Nasdaq and the Magnificent Seven during the recent AI boom.
- Greenwood et al. (2019): Rapid industry run-ups **raise crash probabilities**; discrete drawdowns of large magnitudes are common in the late-bubble phase.
- Daniel et al. (2023): **optimistic investors** keep prices elevated for winners while **corrective beliefs** accumulate, generating large negative repricings as disagreement resolves.
- Both mechanisms imply **positive diffusive drift** coexisting with **cumulative negative jumps** during a concentrated, technology-led rally.

Cumulative Returns: Overnight vs Day



Overnight: drift driven overnight risk premium.

(Boyarchenko et al., 2023; Bondarenko and Muravyev, 2023)

Day: jumps mostly occur during the day.

Momentum or Reversal?

Decompose $R = R^C + R^J$ and test the transmission channels separately.

Baseline

$$R_d^{Day} = \alpha_R + \beta_R R_d^{ON} + u_d^R$$

Component channels (efficient GMM, errors-in-variables corrected)

$$\tilde{\hat{R}}_d^{C,Day} = \alpha_{CC} + \beta_{CC} \tilde{\hat{R}}_d^{C,ON} + u_d^{CC} \quad (\text{diffusive momentum?})$$

$$\tilde{\hat{R}}_d^{J,Day} = \alpha_{JJ} + \beta_{JJ} \tilde{\hat{R}}_d^{J,ON} + u_d^{JJ} \quad (\text{jump reversal?})$$

$$\tilde{\hat{R}}_d^{J,Day} = \alpha_{JC} + \beta_{JC} \tilde{\hat{R}}_d^{C,ON} + u_d^{JC}$$

$$\tilde{\hat{R}}_d^{C,Day} = \alpha_{CJ} + \beta_{CJ} \tilde{\hat{R}}_d^{J,ON} + u_d^{CJ}$$

Estimation: GMM details →

Regression Results: Index Futures

	S&P 500	Nasdaq 100
<i>Panel A: total returns (baseline)</i>		
R^{Day} on R^{ON}	0.083*	0.024
<i>Panel B: component on component</i>		
$\hat{R}^{C,Day}$ on $\hat{R}^{C,ON}$	0.149**	0.113**
$\hat{R}^{J,Day}$ on $\hat{R}^{J,ON}$	-0.105***	-0.087***
$\hat{R}^{J,Day}$ on $\hat{R}^{C,ON}$	-0.032	-0.067**
$\hat{R}^{C,Day}$ on $\hat{R}^{J,ON}$	0.142	0.083
<i>Panel C: components on total overnight</i>		
$\hat{R}^{C,Day}$ on R^{ON}	0.139***	0.092**
$\hat{R}^{J,Day}$ on R^{ON}	-0.057**	-0.068***

Result: a positive **continuous** (momentum) channel and a negative **jump** (reversal) channel. They offset, so the total relation looks weak.

U.S. Equities: Cross Section

Cross Sectional Setup

Data

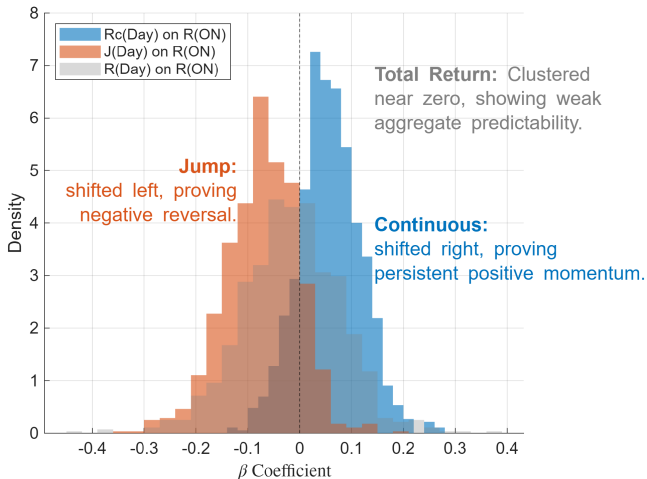
- Balanced panel of 937 U.S. common stocks, 2007 to 2021.
- High-frequency prices from KIBOT; characteristics from CRSP and Compustat.
- Five-minute intraday sampling, 9:30 am to 4:00 pm.

Stock-by-stock predictive regressions on $R_{d,i}^{ON}$

$$\begin{aligned}R_{d,i}^{Day} &= \alpha_i^R + \beta_i^R R_{d,i}^{ON} + \varepsilon_{d,i}^R, \\ \hat{R}_{d,i}^{C,Day} &= \alpha_i^C + \beta_i^C R_{d,i}^{ON} + \varepsilon_{d,i}^C, \\ \hat{R}_{d,i}^{J,Day} &= \alpha_i^J + \beta_i^J R_{d,i}^{ON} + \varepsilon_{d,i}^J.\end{aligned}$$

Cross-sectional decomposition (4-panel) →

Cross Sectional Evidence: Two Channels



What Explains the Heterogeneity?

$$\hat{\beta}_i^k = a_k + b_{S,k} \text{LogSize}_i + b_{D,k} \log(\text{DollarVol}_i) + b_{L,k} \text{Lev}_i + e_{k,i}$$

	β^R (total)	β^C (cont.)	β^J (jump)	$\beta^D = \beta^C - \beta^J $
LogSize	-0.023***	0.004	-0.027***	0.030***
log(DollarVol)	0.022***	0.001	0.021***	-0.021***
Leverage	0.040***	0.040***	0.000	0.032**

Drivers of the two channels

- **Size:** larger firms $\Rightarrow$ stronger **jump reversal**.
- **Trading activity:** less traded firms $\Rightarrow$ stronger **jump reversal**.
- **Financial leverage:** amplifies the **continuous momentum** channel.

Conclusions

Methodological contribution

A new high-frequency framework based on power-transformed semi-returns. At $c = 1/2$ we extract, in one step, four components of the return process:

- the **continuous** return (jump-free, no truncation);
- the **jump** return (no truncation threshold);
- the **integrated leverage** coefficient;

Economic insight

Overnight-to-intraday is *not* a simple reversal:

- **continuous** returns generate **momentum**;
- **jump** components generate **reversal**;
- strength varies with regime and with firm size, trading volume, and leverage.

Thank You! Questions?

Appendix – Estimation: GMM (1/2): Setup and Problem

Component regression in observables, with measurement error $\mathcal{E}_d^{k,s} = \bar{R}_d^{k,s} - R_d^{k,s}$, $k \in \{C, J\}$, $s \in \{Day, ON\}$:

$$\underbrace{\bar{R}_d^{k,Day}}_{R_d^{k,Day} + \mathcal{E}_d^{Day}} = \alpha_k + \beta_k \underbrace{\bar{R}_d^{k,ON}}_{R_d^{k,ON} + \mathcal{E}_d^{ON}} + u_d^k.$$

By Theorem (variance): $\mathcal{E}_d^{C,s} = -\mathcal{E}_d^{J,s}$ and $\text{Var}(\mathcal{E}_d^{C,s} | \mathcal{F}) \approx \bar{V}_{RC} \int_{\mathcal{I}_d^s} \sigma_u^2 du$.

Problem: attenuation in OLS

- $\mathcal{E}_d^{Day}$ orthogonal to $\bar{R}_d^{k,ON}$ (sessions disjoint) – only inflates residual variance.
- $\mathcal{E}_d^{ON}$ correlated with $\bar{R}_d^{k,ON} \Rightarrow$ OLS **attenuates** β_k ; also heteroskedastic by day.

Appendix – Estimation: GMM (2/2): Two-step Weighted GMM

Notation

$Z_d^{(k)} = [1, \tilde{R}_d^{k,ON}]^\top$, $\theta^{(k)} = [\alpha_k, \beta_k]^\top$, $\tilde{IV}_d^{(2),s} = \text{trimmed } \int \sigma_u^2 du \text{ over session } s$;

$$\hat{\Omega}_d = \text{diag}(0, \bar{V}_{RC} \tilde{IV}_d^{(2),ON}).$$

Step 1: initial bias-corrected (unweighted) GMM

$$\sum_d Z_d^{(k)} \left(\tilde{R}_d^{k,Day} - Z_d^{(k)\top} \tilde{\theta}^{(k)} \right) + \hat{\Omega}_d \tilde{\theta}^{(k)} = 0.$$

Consistent first-pass $\tilde{\theta}^{(k)}$; residuals $\tilde{v}_d^k$ deliver $\hat{\sigma}_{u,k}^2$.

Step 2: efficient reweighting

Inverse-variance weights from Step 1:

$$w_d^{-1} = \hat{\sigma}_{u,k}^2 + \bar{V}_{RC} \left(\tilde{IV}_d^{(2),Day} + \tilde{\beta}_k^2 \tilde{IV}_d^{(2),ON} \right).$$

Solve $\sum_d w_d \left[Z_d^{(k)} \left(\tilde{R}_d^{k,Day} - Z_d^{(k)\top} \hat{\theta}_W^{(k)} \right) + \hat{\Omega}_d \hat{\theta}_W^{(k)} \right] = 0$; HAC (Newey–West).

Appendix – Consistency via Long-Span Averaging

Average the daily estimators over N non-overlapping windows:

$$\hat{R}_N^C = \frac{1}{N} \sum_{j=1}^N \hat{R}_j^C, \quad \hat{R}_N^J = \frac{1}{N} \sum_{j=1}^N \hat{R}_j^J, \quad \hat{L}_N = \frac{1}{N} \sum_{j=1}^N \hat{L}_j.$$

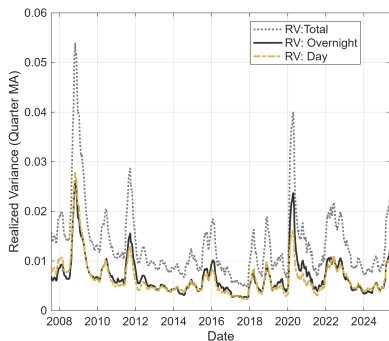
Corollary

Under ergodicity, as $\Delta_n \rightarrow 0$ and $N \rightarrow \infty$ jointly,

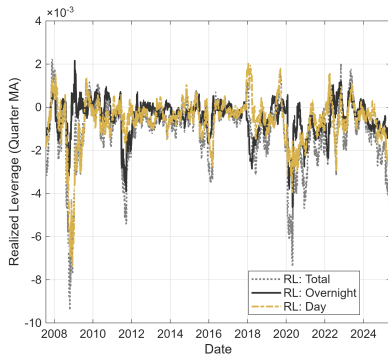
$$\left(\hat{R}_N^C, \hat{R}_N^J, \hat{L}_N \right) \xrightarrow{P} (\mu_b, \mu_J, \mu_{\tilde{\sigma}}).$$

Intuition: joint infill and long-span asymptotics deliver consistent estimates of average drift, jumps, and leverage.

Appendix – Realized Volatility and Leverage



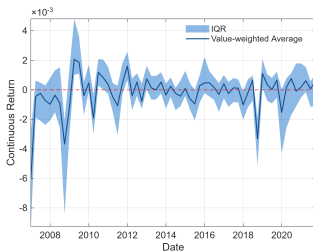
(a) S&P 500: realized volatility



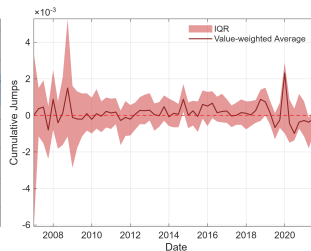
(b) S&P 500: realized leverage

Overnight and day volatilities are highly synchronized. Realized leverage is mostly **negative**, with deeper daytime excursions during turbulent periods.

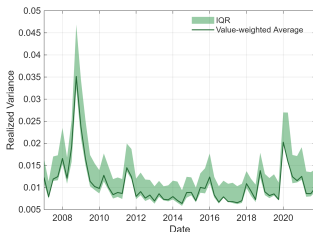
Appendix – Cross Sectional Decomposition



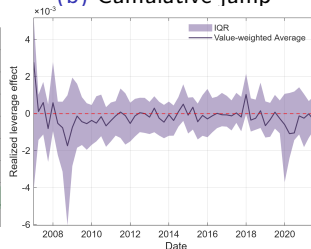
(a) Continuous return



(b) Cumulative jump



(c) Realized volatility



(d) Realized leverage

Interquartile band (cross section) and market-cap-weighted average.