Small Jumps in Noisy High Frequency Data: A Local to Continuity Theory

Jia Li
Department of Economics & Bendheim Center for Finance
Princeton University

April 24, 2010

The theme of this paper

- Intuition: Jumps “appear” to be small in a noisy environment.
- Econometric formalization of this view: local asymptotic model.
- Study the asymptotics of the pre-averaging estimator in the local model.

The benchmark (non-local) model

- The price or log price

$$
\begin{aligned}
X_t &= X_0 + \underbrace{\int_0^t b_s \, ds}_{\text{drift}} + \underbrace{\int_0^t \sigma_s dW_s}_{\text{Brownian}} \\
&\quad + \underbrace{\delta 1_{\{|\delta| \leq 1\}} * (\mu - \nu)_t + \delta 1_{\{|\delta| > 1\}} * \mu_t}_{\text{Jumps}}. \\
\mu &: \text{ jump measure, } \nu : \text{compensator, } \delta : \text{jump size}
\end{aligned}
$$

- Data are discretely sampled: $X_0, X_\Delta, X_{2\Delta}, ..., \Delta \to 0$
- High frequency data is noisy: $Z = X + \text{noise}$.

The "local to continuity" model

- The true price at stage n:

$$X_t^n = X_0 + \int_0^t b_s\,ds + \int_0^t \sigma_s dW_s + \underbrace{c_n}_{\substack{\downarrow \\ 0}} \underbrace{\Delta X}_{\left(\substack{\text{Jumps of Ito semimartingale} \\ \text{source of jumps} \\ \text{controlling the ``direction''} \\ \text{in which price deviates from continuity}}\right)}$$

- The observed process $Z^n = X^n + noise$.

Summary of results

- Asymptotics for the $\boxed{\text{preaverging estimator}}$ in the local-to-continuity model
- Convergences in $\mathbb{P}$ and in law (stably).
- Asymptotics depend on the rate at which $c_n \to 0$.

1. **Nearly continuous case**: (jump size $\searrow 0$ fast)
    1. Same rate of convergence as in the continuous case
    2. Non-central limiting law
2. **Small jump case**: (jump size $\searrow 0$ slow)
    1. Weak identification problem in the estimation of (small) jumps

Pre-averaging estimator

- Noise-robust estimation (LLN) and inference (CLT) for

$$\begin{cases} \int_0^t \sigma_s^p \, ds & \text{when } X \text{ is continuous} \\ \sum_{s\le t} |\Delta X_s|^p & \text{when } X \text{ jumps.} \end{cases}, \text{ for even } p.$$

- The design
  - Shrink noise by smoothing
  - Bias correction
- The pre-averaging estimator $\bar{V}(g, k_n, p)$
  - g : smoothing kernel
  - k_n : smoothing window
  - p : as in $\int_0^t \sigma_s^p \, ds$ or $\sum_{s\le t} |\Delta X_s|^p$

Motivating the local model: An numerical example

- To fix idea, let $p = 4$.
- Under the standard asymptotics (Jacod, Podolskij and Vetter 2009),
  - When X is continuous

$$\bar{V}(g, k_n) \xrightarrow{\mathbb{P}} C_1(g) \int_0^T |\sigma_s|^4 \, ds,$$

  - When X jumps

$$\Delta_n^{1/2} \bar{V}(g, k_n) \xrightarrow{\mathbb{P}} C_2(g) \sum_{s \leq T} |\Delta X_s|^4 .$$

- Heuristically, combine the above

$$\bar{V}(g, k_n) \approx \underbrace{C_1(g) \int_0^T |\sigma_s|^4 \, ds}_{\text{negligible asymptotically}} + \underbrace{\Delta_n^{-1/2}}_{\to\infty} C_2(g) \sum_{s \leq T} |\Delta X_s|^4 .$$

- Left→Right: no jump → 1 very large jump

The Relative Bias

Jump Size ($\times\sigma\Delta^{1/2}, \Delta = 5$ sec)	Standard Preaveraging (benchmark model)
10	39183%
20	2502%
30	515%
50	75%
100	7%

Table: Relative bias for estimating $\sum_{s\leq T} |\Delta X_s|^4$.

Next,

- Assumptions
- Theorems
- Applications

Assumptions

Summary of Assumptions

- X is an Ito semimartingale with regular (i.e. càdlàg or càglàd) and locally bounded coefficients.
- σ is also an Ito semiartingale with regular and locally bounded coefficients.
- Conditional on X, noise is, IID, mean 0, and locally bounded.

Assumption 1: The underlying

$$X_t = X_0 + \int_0^t b_s\, ds + \int_0^t \sigma_s dW_s$$
$$+\delta 1_{\{|\delta|\leq 1\}} * (\mu - \nu)_t + \delta 1_{\{|\delta|>1\}} * \mu_t.$$

a) The process (b_t) is optional and locally bounded;
b) The process (σ_t) is càdlàg (i.e., right-continuous with left limits) and adapted;
c) The function δ is predictable, and there is a bounded function γ in $\mathbb{L}^2(E, \mathcal{E}, \lambda)$ such that the process $\sup_{z\in E}(|\delta(\omega^{(0)}, t, z)| \wedge 1)/\gamma(z)$ is locally bounded;
d) We have almost surely $\int_0^t \sigma_s^2\, ds > 0$ for all $t > 0$.

Assumption 2: The volatility

We have Assumption 1 and σ_t is also an Itô semimartingale which can be written as

$$\sigma_t = \sigma_0 + \int_0^t \tilde{b}_s ds + \int_0^t \tilde{\sigma}_s dW_s + M_t + \sum_{s \leq t} \Delta\sigma_s \, 1_{\{|\Delta\sigma_s| > v\}},$$

where M is a local martingale orthogonal to W and with bounded jumps and $\langle M, M \rangle_t = \int_0^t a_s ds$, and the compensator of $\sum_{s \leq t} 1_{\{|\Delta\sigma_s| > v\}}$ is $\int_0^t a'_s ds$, and where $\tilde{b}_t$, a_t, and a'_t are optional locally bounded processes, and $\tilde{\sigma}_t$ is optional and càdlàg, as well as b_t. Furthermore, we suppose that the processes $\tilde{b}_t$, a_t, a'_t are locally bounded, whereas the processes b_t and $\tilde{\sigma}_t$ are left-continuous with right limits.

Assumption 3: The noise

For each $q > 0$ there is a sequence of $(\mathcal{F}_t^{(0)})$-stopping times $(T_{q,n})_{n\geq 1}$ increasing to ∞, such that $\int Q_t(\omega^{(0)}, dz)\,|z|^q \leq n$ whenever $t < T_{q,n}(\omega^{(0)})$. We write

$$\beta(q)_t(\omega^{(0)}) \;=\; \int Q_t(\omega^{(0)}, dz)\, z^q, \qquad \alpha_t \;=\; \sqrt{\beta(2)_t},$$

and we assume that the processes α and $\beta(3)$ are càdlàg, and that

$$\beta(1) \;\equiv\; 0.$$

Theorems

Nearly continuous case

Theorem

Let $p \geq 2$ be an even integer. Let $c_n = \Delta_n^{\frac{1}{4}-\frac{1}{4p}}$. Assume that jump process has finite variation. The standardized process

$$\tilde{V}(g,p)_t^n = \frac{1}{\Delta_n^{1/4}}\left(\Delta_n^{1-p/4}\bar{V}(Z,g,p)_t^n - C_1(g)\int_0^t |\sigma_s|^p \, ds\right),$$

converges stably in law (in functional sense) to the process

$$M_t + C_2(g)\sum_{s\leq t}|\Delta X_s|^p,$$

where M_t is an $\mathcal{F}$-conditional centered Gaussian martingale with known variance.

The small jump case: LLN

Theorem

Let $p \geq 4$. *Let* $c_n = \Delta_n^{\frac{1}{4} - \frac{1}{2p}}$. *We have*

$$\Delta_n^{1-\frac{p}{4}} \bar{V}\left(Z^n, g, p\right)_t^n \xrightarrow{\mathbb{P}} C_1\left(g\right) \int_0^t \sigma_s^p \, ds + C_2\left(g\right) \sum_{s \leq t} \left|\Delta X_s\right|^p .$$

The small jump case: CLT

Theorem

Let $p \geq 4$. *Let* $c_n = \Delta_n^{\frac{1}{4}-\frac{1}{2p}}$. *The variables*

$$\Delta_n^{-\frac{1}{2p}} \left(\Delta_n^{1-\frac{p}{4}} \bar{V}\left(Z^n, g, p\right)_t^n - C_1\left(g\right) \int_0^t \sigma_s^p \, ds - C_2\left(g\right) \sum_{s \leq t} \left|\Delta X_s\right|^p \right)$$

converge stably in law to a $\mathcal{F}$*-conditional centered Gaussian variable with* $\mathcal{F}$*-conditional variance*

$$\sum_{s \leq t} \left|\Delta X_s\right|^{2(p-1)} \left(\underbrace{C_- \sigma_{s-}^2 + C_+ \sigma_s^2}_{\text{Due to vol}} + \underbrace{C'_- \alpha_{s-}^2 + C'_+ \alpha_s^2}_{\text{Due to noise}} \right).$$

Remark: joint convergence available for multiple g*'s.*

Applications

Application 1: The nearly continuous asymptotics

- Aït-Sahalia, Jacod and Li (2010) proposes a noise-robust jump test based on the pre-averaging method
- Compute local power under the continuous null
- Provide a guideline for choosing smoothing parameters (to maximize local power)

Application 2: The small jump asymptotics

- Estimate $\sum_{s\leq t}|\Delta X_s|^p$, for $p \geq 4$ $\boxed{\textit{even integer}}$
- The preaveraging estimator $\bar{V}(g, k_n, p)$ is biased.
- Bias correction using multiple weighting functions
- The "small jump" LLN

$$\bar{V}(Z^n, g)_t^n \xrightarrow{\mathbb{P}} C_1(g)\int_0^t \sigma_s^p\,ds + C_2(g)\sum_{s\leq t}|\Delta X_s|^p,$$

$$\bar{V}(Z^n, h)_t^n \xrightarrow{\mathbb{P}} C_1(h)\int_0^t \sigma_s^p\,ds + C_2(h)\sum_{s\leq t}|\Delta X_s|^p.$$

- Solving the 2×2 system,

$$\begin{pmatrix} C_1(g) & C_2(g) \\ C_1(h) & C_2(h) \end{pmatrix}^{-1} \begin{pmatrix} \bar{V}(Z^n, g)_t^n \\ \bar{V}(Z^n, h)_t^n \end{pmatrix} \xrightarrow{\mathbb{P}} \begin{pmatrix} \int_0^t \sigma_s^p\,ds \\ \sum_{s\leq t}|\Delta X_s|^p \end{pmatrix}.$$

More simulation results

Jump Size $\left(\times\sigma\Delta^{1/2}\right)$ $\Delta = 5$ sec	Corrected Preaveraging (motivated by local model) Number of Jumps				Standard Preaveraging (motivated by non-local model) Number of Jumps			
	1	3	5	10	1	3	5	10
20	-9%	-3%	-1%	-1%	2502%	894%	567%	327%
30	-3%	-1%	-2%	-1%	515%	195%	132%	85%
50	-2%	-2%	-1%	-1%	75%	33%	25%	19%
100	-1%	-1%	-1%	-1%	7%	5%	4%	4%

Table: Relative bias for estimating $\sum_{s\leq T} |\Delta X_s|^4$.

- Generalization

$$\bar{V}\left(Z^n, g_1\right)_t^n \xrightarrow{\mathbb{P}} C_1\left(g_1\right) \int_0^t \sigma_s^p \, ds + C_2\left(g_1\right) \sum_{s \leq t} \left|\Delta X_s\right|^p,$$

$$\bar{V}\left(Z^n, g_2\right)_t^n \xrightarrow{\mathbb{P}} C_1\left(g_2\right) \int_0^t \sigma_s^p \, ds + C_2\left(g_2\right) \sum_{s \leq t} \left|\Delta X_s\right|^p,$$

$$\ldots$$

$$\bar{V}\left(Z^n, g_m\right)_t^n \xrightarrow{\mathbb{P}} C_1\left(g_m\right) \int_0^t \sigma_s^p \, ds + C_2\left(g_m\right) \sum_{s \leq t} \left|\Delta X_s\right|^p.$$

- Overidentified system $\Rightarrow$OLS/GLS/Minimum Distance.
- CLT available from joint convergence of $\left(\bar{V}\left(Z^n, g_j\right)_t^n\right)_{1 \leq j \leq m}$

Conclusion

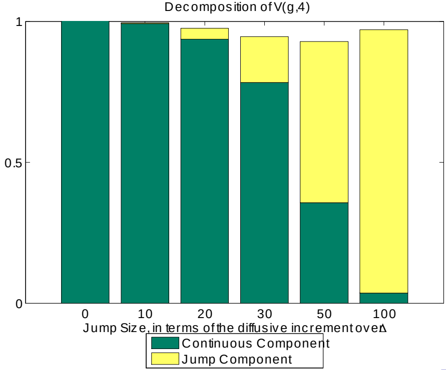
- It is useful to introduce local asymptotics to study the behavior of the preaveraging estimators.

Thanks!

- Simulation:

$$Z_t = \sigma W_t + \text{Jump Size} * 1\{t \geq \tau\} + \text{Noise}_t$$
$$\text{Noise}_t \overset{iid}{\sim} \text{scale factor} \times t_{2.5} \text{ (very heavy tail)}$$

- Relative magnitude: continuous part V.S. jump part



More simulation results

Jump Size $\left(\times\sigma\Delta^{1/2}\right)$ $\Delta = 5$ sec	Corrected Preaveraging (motivated by local model) Number of Jumps				Standard Preaveraging (motivated by non-local model) Number of Jumps			
	1	3	5	10	1	3	5	10
10	-121%	-49%	-41%	-19%	39183%	13283%	8121%	4230%
20	-9%	-3%	-1%	-1%	2502%	894%	567%	327%
30	-3%	-1%	-2%	-1%	515%	195%	132%	85%
50	-2%	-2%	-1%	-1%	75%	33%	25%	19%
100	-1%	-1%	-1%	-1%	7%	5%	4%	4%

Table: Relative bias for estimating $\sum_{s\leq T} |\Delta X_s|^4$.