Assessment of Uncertainty in High Frequency Data: The Observed Asymptotic Variance

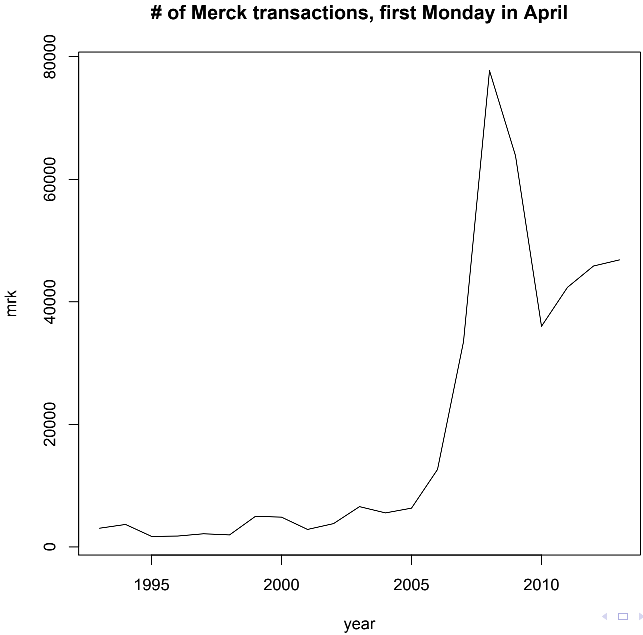
Per A. Mykland and Lan Zhang

Conference in Honor of Ole Barndorff-Nielsen
Aarhus, June 2015

Outline

1. **High frequency data**
   - The data
   - Why high frequency data?
2. **Observed AVAR**
   - The Problem
   - Heuristics
   - Theory
3. **Conclusions**

Evolution of Data Size



High Frequency Financial Data

- Intra-day data
- Ultra high frequency data (UHF)
  - transactions tick-by-tick, from TAQ, Reuters, etc
  - quotes - bid, ask - same sources
  - limit order books, harder to get but more information
  - stocks, bonds, futures, currencies, ...
  - low latency data
- Main feature:
  - almost continuous observation,
  - has microstructure noise
  - observation times can be irregular, and asynchronous for multivariate data
- HF data can also be found in neuroscience, climate recordings, wind measurements, turbulence, fish, ...

Why are High Frequency Data Interesting?

- Modern quantitative finance uses high frequency constructions in stochastic processes:
  - to price assets, underlying and derivative
  - to construct trading strategies
- The high frequency data are the empirical realization of the same processes:
  - What was theory is now observable, testable
  - ... and perhaps even improvable using the data
- The data open a new angle on quantitative finance:
  - better estimates of parameters in models
  - better, more empirically based, models?
  - an opportunity to study rapid change in the market
  - a complement to cross-sectional data and to low frequency time series data.
  - unification of econometrics, risk mgmt, and quantitative finance?

The Standard High Frequency Setup

- Y_{t_i} intra day observables, $Y_{t_i} = X_{t_i}$+microstructure
- X_t: semimartingale which carries information about θ_t
- θ_t: a spot parameter process, e.g. volatility, semi-variance, leverage effect, high frequency β, etc.
- A typical goal: to estimate integrated parameter process,

$$\Theta_{(S,T]} = \int_S^T \theta_t dt$$

over some time period (a day, five minutes, etc)

  - popular example: when $\theta_t = \sigma_t^2$, $\widehat{\Theta}_{(S,T]}$ is a variance estimator.
  - Integrals are also used when estimating the spot θ:

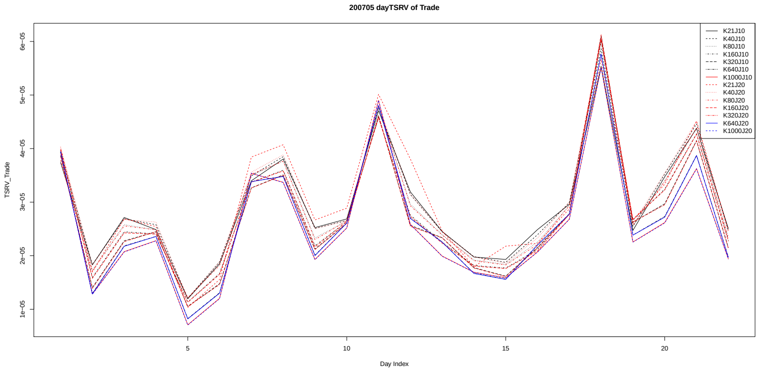
$$\hat{\theta}_t = \frac{1}{h}\widehat{\Theta}_{(t-h,t]}$$

Our question: Uncertainty of $\widehat{\Theta}_{(S,T]}$

This paper is not about developing an estimator $\widehat{\Theta}_{(S,T]}$, but about the uncertainty in $\widehat{\Theta}_{(S,T]}$. In other words,

- Our goal is to estimate the asymptotic variance (AVAR) of $\widehat{\Theta}_{(S,T]}$
- Estimating the AVAR (or s.e.) of $\widehat{\Theta}_{(S,T]}$ is important,
  - to assess the precision of estimators in the form of confidence intervals
  - to create feasible "statistics" for testing
  - to build forecasting models
  - to optimize tuning parameters in finite sample problems

Example: which TSRV?



Estimate $\mathrm{AVAR}\,(\widehat{\Theta})$ in High Frequency Setting

The usual way that leads to AVAR :

(1) An estimator $\widehat{\Theta}_n$

(2) Asymptotics:

$$n^{\alpha}(\hat{\Theta}_n - \Theta) \xrightarrow{\mathcal{L}} N(0, \mathrm{AVAR}) \text{ stably, for some } \alpha > 0$$

(3) A feasible estimator for AVAR .

But, (2) is only available for some estimators. (3) is a rare reality.

- $\mathrm{AVAR}\,(\widehat{\Theta})$ typically is harder to estimate than Θ itself, such as $\int \sigma_t^4 dt$ *and* the volatility of volatility $[\sigma^2, \sigma^2]$ when $\theta_t = \sigma_t^2$ under microstructure – implementation difficulty in (3)
- need AVAR for semivariance, high frequency beta, rank of the volatility matrix, volatility of volatility, under microstructure – analytical difficulty in (2)
- extra challenge in asynchronicity and irregular sampling cases.

An alternative approach to estimating AVAR

We propose an alternative approach: Observed Asymptotic Variance (Observed AVAR)

- Analogy: estimated expected information *versus* observed information
- Allows us to bypass the analytical form of theoretical AVAR
- Estimate both AVAR , and at the same time, the volatility of parameter process θ
- θ can be any semi-martingale process – continuous or not – so long as its integrals can be estimated.
- Deals with edge effects – common in multivariate data, multi-scale or multi-power estimation.
- Antecedents: Barndorff-Nielsen and Shephard, Kalnina and Linton, Gonçalves and Meddahi

Observed AVAR

- Assume that θ is a semimartingale. Similar results may extend to such processes as fractional Brownian motion.
- B: # of time periods (five minutes, say) with $(T_{i-1}, T_i]$ from $T_0 = 0$ to $T_B = \mathcal{T}$
- We have at hand estimator $\hat{\Theta}_i$ of $\Theta_i = \int_{T_{i-1}}^{T_i} \theta_t dt$
- Intuition behind Observed AVAR: compare the estimators $\widehat{\Theta}$ in adjacent blocks

$$\hat{\Theta}_{i+1} - \hat{\Theta}_i = \underbrace{(\hat{\Theta}_{i+1} - \Theta_{i+1})}_{\text{est. error}} + \underbrace{(\Theta_{i+1} - \Theta_i)}_{\text{parameter behavior}} - \underbrace{(\hat{\Theta}_i - \Theta_i)}_{\text{est. error}}$$

Observed AVAR: A First Order Description

The *apparent q.v.* of Θ_t: [note missing n^α]

$$\sum_i (\hat{\Theta}_{i+1} - \hat{\Theta}_i)^2 = 2\sum_i (\hat{\Theta}_i - \Theta_i)^2 + \sum_i (\Theta_{i+1} - \Theta_i)^2$$

$$+ \text{ martingale and negligible terms}$$

$$= (\underbrace{2\sum_i \text{AVAR}(\hat{\Theta}_i - \Theta_i)}_{\text{cumulative AVAR}} + \underbrace{\text{q.v. of } \Theta_i}_{\text{parameter behavior}})(1 + o_p(1)),$$

SRC, when $\max_i (T_{i+1} - T_i)$ goes to zero

Behavior of $\sum_i (\hat{\Theta}_{i+1} - \hat{\Theta}_i)^2$, Under Continuity of θ_t

- Apparent Quadratic Variation: $\sum_i (\hat{\Theta}_{i+1} - \hat{\Theta}_i)^2$

$$\sum_i (\hat{\Theta}_{i+1} - \hat{\Theta}_i)^2 = \left(2 \sum_i \text{AVAR}(\hat{\Theta}_i - \Theta_i) + \text{ q.v. of } \Theta_i \right) (1 + o_p(1))$$

- If $\Delta T = T_{i+1} - T_i$ is independent of i, and if θ_t is continuous, we have an Integral-to-Spot Relation

$$(\Delta T)^{-2} \sum_i (\Theta_{i+1} - \Theta_i)^2 \xrightarrow{p} \frac{2}{3} \left([\theta, \theta]_{\mathcal{T}} - [\theta, \theta]_0 \right) \text{ as } \Delta T \to 0.$$

  - Approximating spot behavior with averaging the integral information induces bias. $\frac{2}{3}$ reflects such bias.
  - Also, see $\frac{1}{2}$ bias in leverage estimation in Mykland and Zhang (2009), Wang and M. (2014).
  - Related to smoothing bias in Ait-Sahalia, Fan, Li (2013).
  - Related to pre-averaging literature on volatility estimation, with X replacing θ.
  - In the paper, the condition on θ being continuous is dropped.

Estimating Cumulative Asymptotic Variance

- One scale estimator:

$$\sum_i (\hat{\Theta}_{i+1} - \hat{\Theta}_i)^2 = \left(2 \sum_i \text{AVAR}(\hat{\Theta}_i - \Theta_i) + \frac{2}{3} (\Delta T)^2 \left([\theta, \theta]_{\mathcal{T}} - [\theta, \theta]_0 \right) \right) \times (1 + o_p(1))$$

- Subsampling and averaging every two time period:

$$\frac{1}{2} \sum_i \left(\hat{\Theta}_{(T_i, T_{i+2}]} - \hat{\Theta}_{(T_{i-2}, T_i]} \right)^2 \approx \sum_i \text{AVAR}(\hat{\Theta}_{(T_i, T_{i+2}]} - \Theta_{(T_i, T_{i+2}]}) + \frac{2}{3} (2\Delta T)^2 \left([\theta, \theta]_{\mathcal{T}} - [\theta, \theta]_0 \right)$$
$$\approx 2 \sum_i \text{AVAR}(\hat{\Theta}_i - \Theta_i) + \frac{2}{3} (2\Delta T)^2 \left([\theta, \theta]_{\mathcal{T}} - [\theta, \theta]_0 \right)$$

- In combination, a new two scales estimator, now for AVAR :

$$\mathrm{TSAVAR} = \frac{2}{3} \sum_i (\hat{\Theta}_{i+1} - \hat{\Theta}_i)^2 - \frac{1}{12} \sum_i \left(\hat{\Theta}_{(T_i, T_{i+2}]} - \hat{\Theta}_{(T_{i-2}, T_i]} \right)^2$$
$$= \left(\sum_i \text{AVAR}(\hat{\Theta}_i - \Theta_i) \right) (1 + o_p(1)).$$

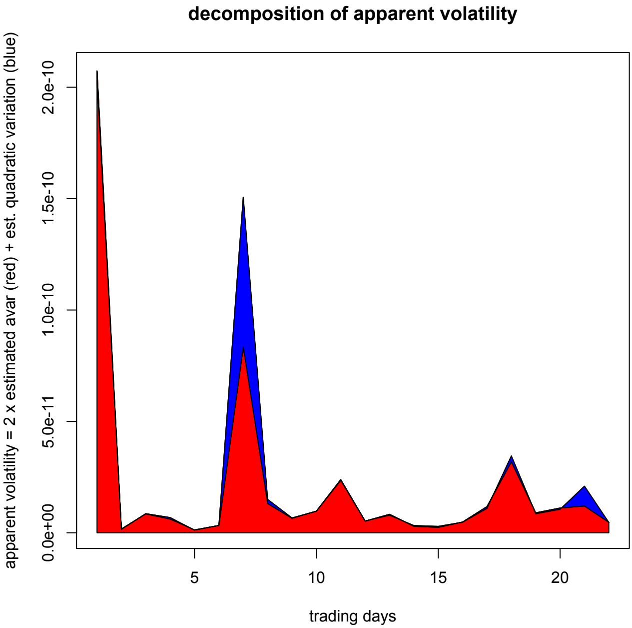
Empirical Decomposition of Apparent Quadratic Variation

- Estimating the Volatility of θ:
  - A different linear combination, a different two scales estimator:

$$\widehat{[\theta,\theta]}_{\mathcal{T}} = \frac{1}{(\Delta T)^2}\left(\frac{1}{4}\sum_i \left(\hat{\Theta}_{(T_i,T_{i+2}]} - \hat{\Theta}_{(T_{i-2},T_i]}\right)^2 - \frac{1}{2}\sum_i (\hat{\Theta}_{i+1} - \hat{\Theta}_i)^2\right)$$
$$\xrightarrow{p} [\theta,\theta]_{\mathcal{T}}$$

  - Earlier work on volatility of σ: Vetter; M. Shephard and Sheppard
- Theoretical and empirical decomposition of apparent quadratic variation:

$$\sum_i (\hat{\Theta}_{i+1} - \hat{\Theta}_i)^2 = \left(2 \times \text{AVAR}(\hat{\Theta} - \Theta) + \frac{2}{3}(\Delta T)^2\left([\theta,\theta]_{\mathcal{T}} - [\theta,\theta]_0\right)\right) \times (1 + o_p(1))$$
$$= 2 \times \text{TSAVAR} + \frac{2}{3}(\Delta T)^2\widehat{[\theta,\theta]}_{\mathcal{T}}$$


decomposition of apparent volatility
apparent volatility = 2 x estimated avar (red) + est. quadratic variation (blue)
0.0e+00
5.0e-11
1.0e-10
1.5e-10
2.0e-10
5
10
15
20
trading days

The Things that go Bump in the Night

We needed

- Convergence of the quadratic variation of big theta:

$$(\Delta T)^{-2} \sum_i (\Theta_{i+1} - \Theta_i)^2 \xrightarrow{p} \frac{2}{3}\left([\theta, \theta]_{\mathcal{T}} - [\theta, \theta]_0\right) \text{ as } \Delta T \to 0$$

  This can fail due to jumps in θ.

- Additivity of the asymptotic variances:

$$\mathrm{AVAR}(\hat{\Theta} - \Theta) = \left(\sum_i \mathrm{AVAR}(\hat{\Theta}_i - \Theta_i)\right)(1 + o_p(1))$$

  This can fail due to Edge effects

There is a way out of both, by going to multiple sampling scales

Resolution for the Process θ

Assumption

θ is a semi-martingale

A general Integral-to-Spot Device:

Theorem

Under our assumption, if $K_n \to \infty$ and $\delta = K\Delta T \to 0$ as $\Delta T \to 0$,

$$\delta^{-2} \frac{1}{K} \sum_{i=K_n}^{B_n - K_n} \left(\Theta_{(T_i, T_{i+K}]} - \Theta_{(T_{i-K}, T_i]}\right)^2 \xrightarrow{p} \frac{2}{3} \left([\theta, \theta]_{\mathcal{T}-} - [\theta, \theta]_0\right)$$

- Result for finite K in the paper
- Robustness to ΔT: ΔT can be arbitrarily small for given δ (integral form in limit), then $\delta \to 0$

The Statistical Setup

- Recall that if $S < T \in [0, \mathcal{T}]$, we set

$$\Theta_{(S,T]} = \int_S^T \theta_t dt$$

- The typical situation: there is a semimartingale M_T and *edge effects* e_S and $\tilde{e}_T$, so that

$$\hat{\Theta}_{(S,T]} - \Theta_{(S,T]} = M_T - M_S + \tilde{e}_T - e_S$$

- The edge effect has a component e_S relating to phasing in the estimator at the beginning of the time interval, and component $\tilde{e}_T$ for the phasing out at T.
- Caveat: All of $\hat{\Theta}_{(S,T]}$, M_T, e_S, and $\tilde{e}_T$ will depend on the number of observations n. For the most part, n is omitted from our notation to avoid an excessive number of subscripts, but we may sometimes write $M_{n,T}$ etc.

The standard asymptotic *result* in the literature

Assumption

- *There is a sigma-field $\mathcal{F}$, representing the underlying processes, including X_t and θ_t, but not necessarily any microstructure noise.*
- *There is a convergence rate n^α so that*

$$L_{n,t} = n^\alpha M_{n,t} \xrightarrow{\mathcal{L}} L_t \text{ as process, stably in law w.r.t. } \mathcal{F}\text{, while}$$

$$\forall S, T : n^\alpha e_{n,S} \xrightarrow{p} 0 \text{ and } n^\alpha \tilde{e}_{n,T} \xrightarrow{p} 0,$$

- *L_t is a local martingale, conditionally Gaussian given $\mathcal{F}$*
- *In addition, it is usuallly required to show results that*

$$\sup_n E \sup_{0 \le t \le T} |\Delta L_{n,t}| < \infty, \text{ or more generally } L_{n,t} \text{ is P-UT}$$

Can handle “large” edge effects, $O_p(n^{-\alpha})$ but not $o_p(n^{-\alpha})$, but more elaborate development

Negligible Edge Effects

Theorem

Under the Assumptions, with the additional condition that average$(e^2_{T_i})$ *and average*$(\tilde{e}^2_{T_i}) = o_p(K\Delta T n^{-2\alpha})$*. Then, as* $\Delta T \to 0$*,*

$$\sum_i (\hat{\Theta}_i - \Theta_i)^2 = \sum_i \text{AVAR}\,(\hat{\Theta}_i - \Theta_i) + o_p(n^{-2\alpha})$$

$$= \text{AVAR}(\hat{\Theta} - \Theta) + o_p(n^{-2\alpha}) = n^{-2\alpha}\left([L,L]_{\mathcal{T}} - [L,L]_0\right) + o_p(n^{-2\alpha})$$

Also, if $\Delta T = o(n^{-\alpha})$*, then with* $\delta = K\Delta T \to 0$*,*

$$\begin{aligned} QV_K(\hat{\Theta}) &= 2[M_n, M_n]_T + \delta^2\left(\frac{2}{3} - \frac{1}{3K}\right)[\theta,\theta]_{\mathcal{T}-} \\ &\quad + o_p(\delta^2) + o_p(n^{-2\alpha}), \\ \text{TSAVAR} &= \text{AVAR}(\hat{\Theta} - \Theta)(1 + o_p(1)) \text{ and} \\ \widehat{[\theta,\theta]}_{\mathcal{T}} &\xrightarrow{p} [\theta,\theta]_{\mathcal{T}} \end{aligned}$$

Guidance: $K\Delta T$ and $n^{-\alpha}$ of same order, in which case need average$(e^2_{T_i})$ and average$(\tilde{e}^2_{T_i}) = o_p(n^{-3\alpha})$ (or go to "hard edge")
If $K\Delta T = o(n^{-\alpha})$, then one scale enough for AVAR

Observed vs. Estimated Expected (tentative)

The estimate is, in fact, of the small sample quantity $[M_n, M_n]_{\mathcal{T}}$

- You do not need to know the AVAR $[L, L]_{\mathcal{T}}$ (convenient)
- You are not taking aim at $[L, L]_{\mathcal{T}}$ (you are not a priori subject to a bad rate of convergence)
- Lesson from likelihood theory: the best estimator of (standard error)^2 may not be the estimated AVAR (papers by Ole Barndorff-Nielsen and others): observed information, r, r^*, etc
- $[M_n, M_n]_{\mathcal{T}}$ is a little like observed information (papers by Ole Barndorff-Nielsen and others)
- In the case of RV, $[M_n, M_n]_{\mathcal{T}}$ reminds you of quarticity (Barndorff-Nielsen and Shephard)
- $[M_n, M_n]_{\mathcal{T}}$ *is* the information in dual likelihood: in asymptotically ergodic case (= when you *don't need* stable convergence), higher order properties of likelihood are mostly inherited
- Another likelihood connection: When you *do need* stable convergence: $n^{-2\alpha}[M_n, M_n]_{\mathcal{T}} \approx [L, L]_{\mathcal{T}}$ is the asymptotic variance if and only if the underlying estimator is asymptotically efficient

Observed vs. Estimated Expected: The Limits to Inference

"parameter"	α in Different Situations			
	Microstructure Absent		Microstructure Present	
	daily	spot	daily	spot
volatility regression ANOVA	$O_p(n^{1/2})$	$O_p(n^{1/4})$	$O_p(n^{1/4})$	$O_p(n^{1/8})$
leverage effect vol of vol	$O_p(n^{1/4})$	$O_p(n^{1/8})$	$O_p(n^{1/8})$	$O_p(n^{1/16})$

Table: n is daily number of transactions/quotes

Hard Edge

- If the edge effects e_T and $\tilde{e}_T$ are of the order $O_p(n^{-\beta})$, $\beta > \alpha$

$$\sum_i (\hat{\Theta}_i - \Theta_i)^2 \approx n^{-2\alpha} \sum_i \left([M_n, M_n]_{T_i} - [M_n M_n]_{T_{i-1}} \right) + \mathrm{Var}(\tilde{e}_{T_i}) + \mathrm{Var}(e_{T_{i-1}}) \Big)$$
$$= n^{-2\alpha}[L, L]_{\mathcal{T}} + O_p(B_n n^{-2\beta})$$

- which is much bigger than

$$\mathrm{AVAR}(\hat{\Theta} - \Theta) = n^{-2\alpha}[L, L]_{\mathcal{T}} + \mathrm{Var}(\tilde{e}_0) + \mathrm{Var}(e_{\mathcal{T}}) + o_p(n^{-2\alpha})$$

- The difference is of order $O_p(B_n n^{-2\beta}) >> \mathrm{AVAR}(\hat{\Theta} - \Theta)$ (potentially)

A Way Out: Subsampling and Averaging

- Consider K-averaged apparent q.v. of Θ
- Decomposition into parameter behavior and estimation error

$$
\begin{aligned}
QV_K = \frac{1}{K}\sum_{i=K}^{B-K}(\hat{\Theta}_{(T_i,T_{i+K}]} - \hat{\Theta}_{(T_{i-K},T_i]})^2 &= \frac{1}{K}\sum_{i=K}^{B-K}(\Theta_{(T_i,T_{i+K}]} - \Theta_{(T_{i-K},T_i]})^2 \\
&+ \frac{1}{K}\sum_{i=K}^{B-K}\Big((\hat{\Theta}_{(T_i,T_{i+K}]} - \Theta_{(T_i,T_{i+K}]}) - (\hat{\Theta}_{(T_{i-K},T_i]} - \Theta_{(T_{i-K},T_i]})\Big)^2 \\
&+ \frac{2}{K}V_1 + O_p(n^{-2\alpha}K_n^{-1}B_n^{1/2}) + o_p(n^{-2\alpha}),
\end{aligned}
$$

where V_1 does not depend on K

- Variance made up of: $\hat{\Theta}_{(T_i,T_{i+K}]} - \Theta_{(T_i,T_{i+K}]} = M_{T_{i+K}} - M_{T_i} + \tilde{\epsilon}_{T_{i+K}} - \epsilon_{T_i}$
- This part resembles a problem where multiscaling works – it's similar to estimating volatility under microstructure

Characterization of the Hard Edge

Assumption

There is an integer J *as a filtration* $(\mathcal{G}_t)$*, and decompositions* $e_{T_i} = e'_{T_i} + e''_{T_i}$ *and* $\tilde{e}_{T_i} = \tilde{e}'_{T_i} + \tilde{e}''_{T_i}$*:*

- $(e'_{T_i}, \tilde{e}'_{T_i})$ *are* $\mathcal{G}_{T_{i+J}}$*-measurable*
- $E(e'_{T_i} \mid \mathcal{G}_{T_{i-J}}) = E(\tilde{e}'_{T_i} \mid \mathcal{G}_{T_{i-J}}) = 0$
- $\sum_i (e''_{T_i})^2 = o_p(K_n n^{-2\alpha})$ *and* $\sum_i (\tilde{e}''_{T_i})^2 = o_p(K_n n^{-2\alpha})$
- *For some* $\beta > \alpha$*:* $\sup_n E\, n^{\beta} \left(\max_{0 \le i \le B_n} |e'_{n,T_i}| + \max |\tilde{e}'_{n,T_i}| \right) < \infty$

Can handle $\beta = \alpha$, but more elaborate development

Benchmark Examples

1. Realized volatility (microstructure absent)
2. Bipower variation (microstructure absent)
3. Two-scale realized volatility from pre-averaged data (microstructure present)
4. Multi-scale and kernel realized volatility (microstructure present)
5. Co-volatility from asynchronous observations
6. Leverage effect (continuous case)
7. Block estimation of higher power variation (microstructure absent)
8. High frequency regression and ANOVA (microstructure absent)
9. Volatility of volatility (microstructure absent)
10. Nearest-neighbor truncation estimator (Andersen, Dobrev, Schaumburg)

Simulation

Case study: Moving Window Estimator of Integral of Volatility

Wish to estimate

$$\Theta = \int_0^{\mathcal{T}} \sigma_t^p \, dt.$$

Block length M:

$$\hat{\Theta}_n^{MW} = \frac{1}{M} \sum_{i=0}^{n-M} c_{M,p}^{-1} \left| \hat{\sigma}_{t_{n,i}} \right|^p ,$$

where

$$\hat{\sigma}_{t_{n,i}}^2 = \frac{1}{\Delta t_n M} \sum_{j=1}^{M} \Delta X_{t_{n,i+j}}^2 ,$$

and $c_{M,p} = \left(\frac{2}{M}\right)^{p/2} \frac{\Gamma\left(\frac{p+M}{2}\right)}{\Gamma\left(\frac{M}{2}\right)}$.

In simulation: $p = 4$ (quarticity). Data from Heston model, observed every second.

Simulation

Correlations of TSAVAR for Moving Window Estimator of Quarticity, and Theoretical AVAR

$\hat{\Theta}_n(M, \Delta T)$	$\hat{\Theta}_n\ (10, 60)$	$\hat{\Theta}_n\ (20, 60)$	$\hat{\Theta}_n\ (30, 60)$	$\hat{\Theta}_n\ (100, 200)$	$\hat{\Theta}_n\ (200, 400)$	$\hat{\Theta}_n\ (300, 600)$	Theo. AVAR
$\hat{\Theta}_n\ (10, 60)$		0.9876	0.9701	0.8149	0.7072	0.6161	0.9259
$\hat{\Theta}_n\ (20, 60)$			0.9917	0.8162	0.7081	0.6166	0.9271
$\hat{\Theta}_n\ (30, 60)$				0.8158	0.7083	0.6167	0.9274
$\hat{\Theta}_n\ (100, 200)$					0.6851	0.5461	0.8439
$\hat{\Theta}_n\ (200, 400)$						0.6730	0.7345
$\hat{\Theta}_n\ (300, 600)$							0.6401
Theo. AVAR							

M is the block length for the underlying estimator. ΔT is the length of the basic block in the AVAR estimator. The fast and slow time scales of the AVAR estimator are set as: $J = 1$ and $K = 2$.

Representation of K Scale Quadratic Variation

Theorem

Under our assumptions, let $K = K_n$ *be a sequence of integers so that* $J \leq K_n \leq B_n$, *with* $K_n \Delta T = O(n^{-\alpha}), K_n \to \infty$ *and* $\Delta T_n = K_n / B_n \to 0$ *as* $n \to \infty$. *Then*

$$
\begin{aligned}
QV_K &= \frac{1}{K} \sum_{i=K}^{B-K} (\hat{\Theta}_{(T_i, T_{i+K}]} - \hat{\Theta}_{(T_{i-K}, T_i]})^2 \\
&= \frac{2}{3} (K \Delta T)^2 \left([\theta, \theta]_{\mathcal{T}-} - [\theta, \theta]_0 \right) + 2n^{-2\alpha} ([L, L]_{\mathcal{T}} - [L, L]_0) \\
&+ \frac{1}{K} V_0 + \textit{ MEE } + o_p(n^{-2\alpha}) + O_p(n^{-2\alpha} K_n^{-1} (B_n - 2K_n + 1)^{1/2}),
\end{aligned}
$$

where V_0 *does not depend on* K_n. *The "Meta Edge Effect" (MEE) is zero unless* $\beta = \alpha$

Multiscale Estimators

- Up to remainder terms:

$$QV_K = \underbrace{\frac{2}{3}(K\Delta T)^2\left([\theta,\theta]_{\mathcal{T}-} - [\theta,\theta]_0\right)}_{\text{variation of parameter}} + \underbrace{2n^{-2\alpha}([L,L]_{\mathcal{T}} - [L,L]_0)}_{\text{main AVAR term}} + \frac{1}{K}V_0$$

- Need to eliminate V_0/K term (which is big, $O_p(n^{-\alpha}B_n^{1/2})$)
- Need to separate red and blue terms
- Solution: m scales: $J \leq K_1 < K_2 < \cdots < K_m$ where $K_l = K_{n,l}$
- Multi-scale estimator:
  $MSQV = \sum_{l=1}^{m} \gamma_l QV_{K_l}$, where $\gamma_l = \gamma_{n,l}$
- To eliminate V_0/K term, impose

$$\sum_{l=1}^{m} \frac{\gamma_l}{K_l} = 0 \tag{1}$$

Theorem

Under our assumptions, let m be given. Assume (1), and also

$$K_{n,m}\Delta T = O(n^{-\alpha}), K_{n,m} \to \infty \text{ as } n \to \infty.$$

Assume that $\sum_{K_l \text{ bounded}} |\gamma_{n,l}| = o(1)$. If $\gamma_n = \sum_{l=1}^m \gamma_{n,l}$, then

$$\begin{aligned} MSQV = & \frac{2}{3}\sum_{l=1}^{m} \gamma_{n,l} K_{n,l}^2 (\Delta T)^2 \left([\theta,\theta]_{\mathcal{T}-} - [\theta,\theta]_0\right) \\ & + \gamma_n \left\{2n^{-2\alpha}([L,L]_{\mathcal{T}} - [L,L]_0) + \mathrm{MEE}_0\right\} \\ & + o_p(\max_l |\gamma_l| n^{-2\alpha}) + O_p\left(n^{-2\alpha}\mathcal{E}_n^{1/2}\right), \end{aligned}$$

where

$$\mathcal{E}_n = \sum_{l=1}^{m} \left(\frac{\gamma_{n,l}}{K_{n,l}}\right)^2 (B_n - 2K_{n,l} + 1).$$

Consistent Estimation of Asymptotic Variance

From the theorem, estimate $\mathrm{AVAR}(\widehat{\Theta}_n)$ by further requiring

$$\sum_{l=1}^{m} \gamma_{n,l} = \frac{1}{2} \text{ and } \sum_{l=1}^{m} \gamma_{n,l} K_{n,l}^2 = 0, \qquad (2)$$

and by adjusting the edge effect. We thus get an estimator

$$\widehat{\mathrm{AVAR}}(\widehat{\Theta}_n) \;=\; MSQV + \frac{1}{2}\widehat{AMEE}.$$

Theorem

Assume the conditions of the Multiscale Theorem, and also that (2) is satisfied. Also suppose that $\widehat{AMEE}$ is formed as described, that $\max_l |\gamma_l| = O(1)$ and $\mathcal{E}_n = o(1)$. Then

$$\widehat{\mathrm{AVAR}}(\widehat{\Theta}_n) \;=\; \mathrm{AVAR}(\widehat{\Theta}_n)\,(1 + o_p(1))$$

This follows from previous Theorem since $\mathrm{AVAR}(\widehat{\Theta}_n) = O_p(n^{-2\alpha})$

Feasible Central Limit Theorem

Recall that given the constraints on weights γ_l in multiscale construction:

$$\sum_{l=1}^{m} \frac{\gamma_l}{K_l} = 0, \quad \sum_{l=1}^{m} \gamma_{n,l} = \frac{1}{2} \text{ and } \quad \sum_{l=1}^{m} \gamma_{n,l} K_{n,l}^2 = 0,$$

and by adjusting the meta edge effect. The observed AVAR is

$$\widehat{\mathrm{AVAR}}(\widehat{\Theta}_n) \; = \; MSQV + \frac{1}{2}\widehat{AMEE}.$$

Theorem

(FEASIBLE ESTIMATION.) *Assume the conditions of precious theorem, and also that $(L_{\mathcal{T}}, R_0, \tilde{R}_{\mathcal{T}})$ is conditionally Gaussian given $\mathcal{F}$. Then*

$$\frac{\widehat{\Theta}_n - \Theta}{\widehat{AVAR}_n^{1/2}} \xrightarrow{\mathcal{L}} N(0,1) \text{ stably in law.} \tag{3}$$

Two Implementations

- Using exactly three scales $(m = 3)$:
  In this case, the three $\gamma_{n,l}$ are determined by the three linear equations (1) and (2), the solution being

$$\gamma_{n,1} = -\frac{1}{v_n} K_{n,1}(K_{n,3}^3 - K_{n,2}^3),$$

$$\gamma_{n,2} = \frac{1}{v_n} K_{n,2}(K_{n,3}^3 - K_{n,1}^3), \text{ and}$$

$$\gamma_{n,3} = -\frac{1}{v_n} K_{n,3}(K_{n,2}^3 - K_{n,1}^3), \text{ where}$$

$$v_n = 2(K_{n,1} + K_{n,2} + K_{n,3})(K_{n,2} - K_{n,1})(K_{n,3} - K_{n,1})(K_{n,3} - K_{n,2}).$$

- Using several scales, to minimize the variance (later slide)

Estimating the Volatility of θ

- We can similarly use the Multiscale Theorem to estimate the volatility $[\theta,\theta]_{\mathcal{T}-} - [\theta,\theta]_0$
- The side conditions now become (instead of (2))

$$\sum_{l=1}^{m} \gamma_{n,l} = 0 \text{ and } \sum_{l=1}^{m} \gamma_{n,l} K_{n,l}^2 = \frac{3}{2}(\Delta T)^{-2}. \tag{4}$$

- There is here no need to worry about Meta edge effect.
- Produces consistent estimator $\widehat{[\theta,\theta]} \xrightarrow{p} [\theta,\theta]_{\mathcal{T}-}$

Optimized Estimators of AVAR and of Volatility

We now consider the case of general number m of scales. Set

$$\mathbb{A}_n = \begin{pmatrix} K_{n,1}^{-1} & K_{n,2}^{-1} & \cdots & K_{n,m}^{-1} \\ 1 & 1 & \cdots & 1 \\ K_{n,1}^{2} & K_{n,2}^{2} & \cdots & K_{n,m}^{2} \end{pmatrix} \text{ and } \underline{\gamma}_n = \begin{pmatrix} \gamma_{n,1} \\ \cdots \\ \gamma_{n,m} \end{pmatrix}$$

and $\mathbb{C}_n = \text{diag}(K_{n,1}^{-2}(B_n - 2K_{n,1} + 1), \cdots, K_{n,m}^{-2}(B_n - 2K_{n,m} + 1))$. We note that $\mathcal{E}_n = \underline{\gamma}_n^* \mathbb{C}_n \underline{\gamma}_n$, where "$*$" denotes transpose. Our two optimization problems thus become

$$\min \underline{\gamma}_n^* \mathbb{C}_n \underline{\gamma}_n \text{ subject to } \mathbb{A}_n \underline{\gamma}_n = \underline{b}_n \tag{5}$$

with standard solution

$$\underline{\gamma}_n = \mathbb{A}_n^* (\mathbb{A}_n \mathbb{C}_n^{-1} \mathbb{A}_n^*)^{-1} \underline{b}_n. \tag{6}$$

- To estimate $\text{AVAR}(\widehat{\Theta}_n)$, set $\underline{b}_n = \left(0, \frac{1}{2}, 0\right)^*$
- To estimate $[\theta, \theta]_{\mathcal{T}-} - [\theta, \theta]_0$, set $\underline{b}_n = \left(0, 0, \frac{3}{2}(\Delta T_n)^{-2}\right)^*$

For the optimal solution, $\mathcal{E}_n = \underline{b}_n^* (\mathbb{A}_n \mathbb{C}_n^{-1} \mathbb{A}_n^*)^{-1} \underline{b}_n$.

Conclusions

- Introduces the Observed Asymptotic Variance – an alternative (nonparametric) approach to assessing the estimation error.
- Theory development:
  - Integral-to-Spot Device
  - Observed AVAR under small and big edge effect
- Applied motivation: observed AVAR can be adopted to existing estimators in the literature.
- Methodology estimates AVAR for general estimators (not just volatility) and general processes.
- Asynchronicity and irregular sampling are treated as part of the edge effect.
- Multiscale construction also provides a new consistent estimator for $[\theta, \theta]_{\mathcal{T}}$.
- Straightforward generalization to multivariate dimensions.