Estimation of Intraday Realized Variance, Trading Volume, and Correlation

Yiwen Shen, with Prof. Paul Glasserman
Decision, Risk, and Operations, Columbia Business School

Motivation

- Most studies on stock volatility, trading volume, and correlation are based on “low-frequency” daily data.
- However, it is also important to examine and understand the intraday patterns of volatility, trading volume, and correlation.
- This helps us to gain insight of the trading behaviors of market participants, and has huge potentials in various applications.
- We use a large-scale data set that has not been studied in previous literature: second-level trade data for all S&P 500 components from 2003 to 2018.
- We construct efficient and robust estimators to study the general intraday patterns for US large-cap & liquid stocks.

Methodology

- Traditional high-frequency estimates for realized variance (RV) and covariance (RCV) suffer from market microstructure noise:
  - The bias increases linearly with the number of sampling points.
- But if we reduce the sampling frequency, we will lose most price information, as only last observation is kept.
- It is hard to choose the sampling frequency that balances the market microstructure noise versus information loss. Most previous study chose this in an ad hoc way.
- We use the Two-Scale RV estimator in Ait-Sahalia et al (JASA 2005), which combines observations from a fast and a slow time scale.
- It is unbiased under microstructure noise and uses all observations.
- The TSRV estimator is given by:
  \[ \text{TSRV} = \frac{1}{n} \sum_{i=1}^{n} (X_i - X_j)^2 \]
where sum of squared returns is used.
- Similarly, we use a Two-Scale RCV estimator in Zhang (JoE 2011) for realized covariance:
  \[ \text{RCV} = \frac{1}{n} \sum_{i=1}^{n} (Y_i - Y_j)^2 \]
  where correlation is computed using squared returns.
- Then we compute realized correlation:
  \[ \text{Corr}(X, Y) = \frac{\text{RCV}}{\sqrt{\text{TSRV} \times \text{TSRCV}}} \]
- Moreover, the computation of TSRV and TSRCV can be vectorized, which facilitates its implementation for our large-scale data set.
- In our study, fast scale samples every 10 seconds, and slow scale samples every minute.
- We compute the TSRV and TSRCV for every 30 minutes with a five-minute moving window from 9:30AM to 16PM.

Realized Variance & Trading Volume

For stock \( i \), day \( t \), and time \( t' \), we define normalized RV as:
\[ \text{NormRV}_{t,t'} = \frac{RV_{t,t'}}{\max[RV_{t',t}]}. \]
- Intraday RV generally decreases with trading time, with a small spike near market close.
- We define volume share as:
  \[ \text{VolShare}_{t,t'} = \frac{Vol_{t,t'}}{\text{Vol}_{t,t'}}. \]
- The intraday volume changes from U-shape to lopsided smile, i.e., trading volume near market close.
- This may be due to increased passive trading in recent years.

Realized Correlation of Stock Pair Buckets

We show by numerical experiments that the TSRV estimator is robust to different choices of time scales. Besides, the estimates match the results from parametric approach by MLE (Zhang et al, RFS 2005).

Realized Correlation Between Sectors

We can compute the portfolio-implied average correlation between sectors:
\[ \rho_{AB} = \frac{1}{2} \sum_{k=1}^{K} \rho_{ij} = \frac{1}{2} \sum_{k=1}^{K} \frac{RCV_{ij}}{\text{Vol}_{ij}}, \]
- For example, “0-1” bucket includes the stock pairs that have lowest daily correlation, below the first percentile of all stock pairs (total 100,000+ pairs).
- Before crisis, intraday correlation exhibits a U-shape in the middle of day, and is significantly lower in the morning and near market close.
- Starting from crisis, the intraday correlation increases near market close.

In recent years, we see increasing variation in the intraday correlation patterns:
- The buckets with highest daily correlation have similar patterns as before.
- While the intraday correlation of lowest buckets starts from negative regime, and monotonically increases during the trading hours.

Interday Dispersion of Trading Volume

We measure the dispersion of trading volume across trading days at each time \( t \) as:
\[ \text{VolStd}_{t} = \frac{\text{Std}(\text{Vol}_{t})}{\text{Mean}(\text{Vol}_{t})}. \]
- Interday dispersion generally decreases with time: more active trading in the morning, more passive trading around market close.
- The decrease in dispersion is more significant in recent years.

References
