Intraday volatility modelling

Sarahnour Ghaith & Thomas Roiseux

ENSIIE

September 2023 – January 2024

Table of contents I

Table of contents II

Introduction

Context

Context

Use of high frequency data to provide volatility analysis.

Project goals

1. Estimating realized volatility with variable time intervals.
2. Comparizon with a long-range volatility estimation.
3. Estimating the size of the mcirostructure noise.

Introduction

Context

Context

Use of high frequency data to provide volatility analysis.

Project goals

1. Estimating realized volatility with variable time intervals.
2. Comparizon with a long-range volatility estimation.
3. Estimating the size of the mcirostructure noise.

Introduction

Context

Context

Use of high frequency data to provide volatility analysis.

Project goals

1. Estimating realized volatility with variable time intervals.
2. Comparizon with a long-range volatility estimation.
3. Estimating the size of the mcirostructure noise.

Introduction

Context

Context

Use of high frequency data to provide volatility analysis.

Project goals

1. Estimating realized volatility with variable time intervals.
2. Comparizon with a long-range volatility estimation.
3. Estimating the size of the mcirostructure noise.

Data collection

Data collection

Chosen dataset

IShare S&P values (IVE) tick data, from September 2009 until today.

Data cleaning and filtering

Outliers

Among the whole data set: 9 prices were lower than 10 $.

Dropping outliers

9 values among more than 11 millions: we can easily drop them without taking care of the consequences.

Data cleaning and filtering

Outliers

Among the whole data set: 9 prices were lower than 10 $.

Dropping outliers

9 values among more than 11 millions: we can easily drop them without taking care of the consequences.

Data cleaning and filtering

Outliers

Among the whole data set: 9 prices were lower than 10 $.

Dropping outliers

9 values among more than 11 millions: we can easily drop them without taking care of the consequences.

Prices plot

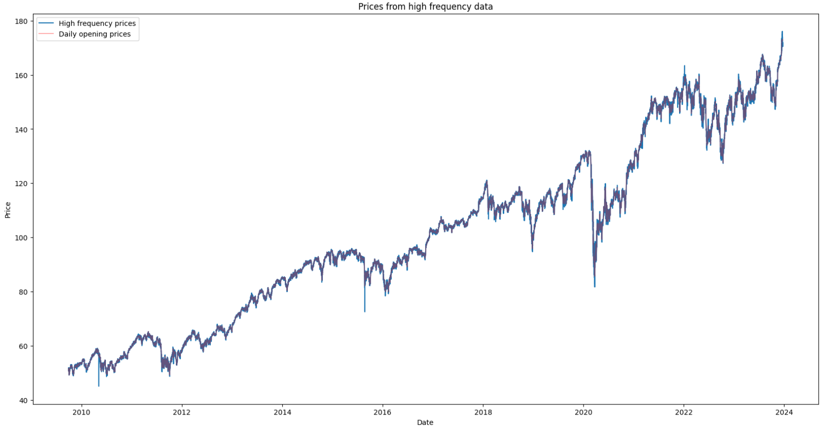

Figure: Prices with respect to time

Introduction to realized volatility

Mathematical definition

$$RV = \sum_{i=1}^{n} \left(X_{t_{i+1}} - X_{t_i}\right)^2$$

with n the number of time samples.

Building the time intervals

How to properly select data with respect to time intervals when available data don't have a fixed δ?

Solution

- If available, we choose the point that fits the interval.
- If not, we take the closest neet one.

Introduction to realized volatility

Mathematical definition

$$RV = \sum_{i=1}^{n} \left(X_{t_{i+1}} - X_{t_i}\right)^2$$

with n the number of time samples.

Building the time intervals

How to properly select data with respect to time intervals when available data don't have a fixed δ?

Solution

- If available, we choose the point that fits the interval.
- If not, we take the closest neet one.

Introduction to realized volatility

Mathematical definition

$$RV = \sum_{i=1}^{n} \left(X_{t_{i+1}} - X_{t_i} \right)^2$$

with n the number of time samples.

Building the time intervals

How to properly select data with respect to time intervals when available data don't have a fixed δ?

Solution

- If available, we choose the point that fits the interval.
- If not, we take the closest neet one.

Paris | Évry-Courcouronnes

Introduction to realized volatility

Mathematical definition

$$RV = \sum_{i=1}^{n} \left(X_{t_{i+1}} - X_{t_i}\right)^2$$

with n the number of time samples.

Building the time intervals

How to properly select data with respect to time intervals when available data don't have a fixed δ?

Solution

- If available, we choose the point that fits the interval.
- If not, we take the closest neet one.

Introduction to realized volatility

Mathematical definition

$$RV = \sum_{i=1}^{n} \left(X_{t_{i+1}} - X_{t_i} \right)^2$$

with n the number of time samples.

Building the time intervals

How to properly select data with respect to time intervals when available data don't have a fixed δ?

Solution

- If available, we choose the point that fits the interval.
- If not, we take the closest neet one.

Paris | Évry-Courcouronnes

Introduction to realized volatility

Mathematical definition

$$RV = \sum_{i=1}^{n} \left(X_{t_{i+1}} - X_{t_i}\right)^2$$

with n the number of time samples.

Building the time intervals

How to properly select data with respect to time intervals when available data don't have a fixed δ?

Solution

- If available, we choose the point that fits the interval.
- If not, we take the closest neet one.

Paris | Évry-Courcouronnes

Estimated realized volatility plot

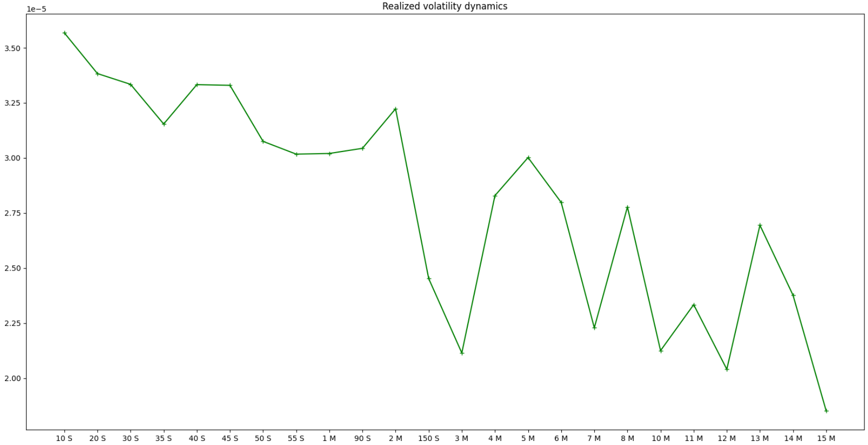

Figure: Estimated realized volatility with several time intervals

Estimated realized volatility plot

Some remarks on the previous plot

1. Volatility decreases when time interval grows.
2. Explained by less selected price variations.
3. The larger the time interval is, the fewer selected prices are.

Estimated realized volatility plot

Some remarks on the previous plot

1. Volatility decreases when time interval grows.
2. Explained by less selected price variations.
3. The larger the time interval is, the fewer selected prices are.

Estimated realized volatility plot

Some remarks on the previous plot

1. Volatility decreases when time interval grows.
2. Explained by less selected price variations.
3. The larger the time interval is, the fewer selected prices are.

Estimated realized volatility plot

Some remarks on the previous plot

1. Volatility decreases when time interval grows.
2. Explained by less selected price variations.
3. The larger the time interval is, the fewer selected prices are.

Selected data

Chosen time line

To compute the daily realized volatility: selection of the same month as before.

Reminder of goal

Comparizon between estimated relazied volatility and long range daily volatility.

Selected data

Chosen time line

To compute the daily realized volatility: selection of the same month as before.

Reminder of goal

Comparizon between estimated relazied volatility and long range daily volatility.

Prices overview

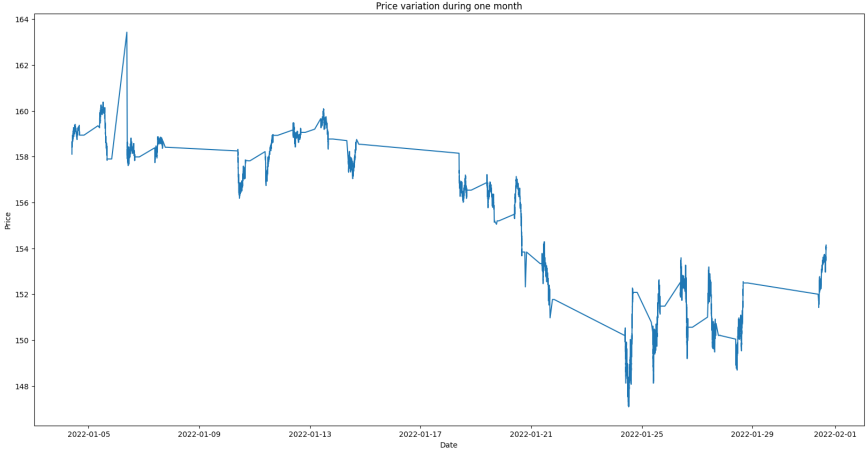

Figure: Prices used for comparing. Flat curves correspond to closed market times.

Comparizon results

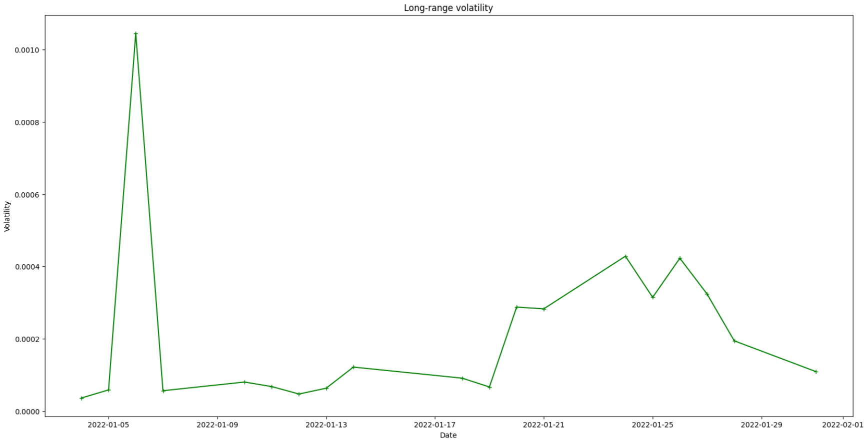

Figure: Long range volatility over the month

Some remarks

1. Dynamics are different:
   1.1 Decresing for intraday.
   1.2 No one for long range.
2. For high freuency data: long range volatility is not useful, as transactions are concluded within a second.
3. Compared to prices graph: the long range volatility tends to show no remarkable variation, which contrasts with all the small ones on that graph.

Some remarks

1. Dynamics are different:
   1.1 Decresing for intraday.
   1.2 No one for long range.
2. For high freuency data: long range volatility is not useful, as transactions are concluded within a second.
3. Compared to prices graph: the long range volatility tends to show no remarkable variation, which contrasts with all the small ones on that graph.

Some remarks

1. Dynamics are different:
   1.1 Decresing for intraday.
   1.2 No one for long range.
2. For high freuency data: long range volatility is not useful, as transactions are concluded within a second.
3. Compared to prices graph: the long range volatility tends to show no remarkable variation, which contrasts with all the small ones on that graph.

Some remarks

1. Dynamics are different:
   1.1 Decresing for intraday.
   1.2 No one for long range.
2. For high freuency data: long range volatility is not useful, as transactions are concluded within a second.
3. Compared to prices graph: the long range volatility tends to show no remarkable variation, which contrasts with all the small ones on that graph.

Some remarks

1. Dynamics are different:
    1.1 Decresing for intraday.
    1.2 No one for long range.
2. For high freuency data: long range volatility is not useful, as transactions are concluded within a second.
3. Compared to prices graph: the long range volatility tends to show no remarkable variation, which contrasts with all the small ones on that graph.

Some remarks

1. Dynamics are different:
   1.1 Decresing for intraday.
   1.2 No one for long range.
2. For high freuency data: long range volatility is not useful, as transactions are concluded within a second.
3. Compared to prices graph: the long range volatility tends to show no remarkable variation, which contrasts with all the small ones on that graph.

Preliminaries

Definition

Can be defined as the result of the ask & bid mechanism. Provides interferences in prices.

Why estimating it?

Noise has impact on prices: estimating it leads to better models.

Estimation

Very easy to estimate.

Preliminaries

Definition

Can be defined as the result of the ask & bid mechanism. Provides interferences in prices.

Why estimating it?

Noise has impact on prices: estimating it leads to better models.

Estimation

Very easy to estimate.

Preliminaries

Definition

Can be defined as the result of the ask & bid mechanism. Provides interferences in prices.

Why estimating it?

Noise has impact on prices: estimating it leads to better models.

Estimation

Very easy to estimate.

Preliminaries

Definition

Can be defined as the result of the ask & bid mechanism. Provides interferences in prices.

Why estimating it?

Noise has impact on prices: estimating it leads to better models.

Estimation

Very easy to estimate.

Methodology

Constructing the estimator

We know that, with η such estimator:

$$\frac{RV}{n} \xrightarrow[n\to\infty]{\mathrm{P}} 2\eta^2 \iff \sqrt{\frac{RV}{2n}} \xrightarrow[n\to\infty]{\mathrm{P}} \eta$$

Computing the estimator

We need to use a large sample of size $n \to \infty$ to have best results for η, the micro-structure noise size estimator.

Methodology

Constructing the estimator

We know that, with η such estimator:

$$\frac{RV}{n} \xrightarrow[n\to\infty]{\mathrm{p}} 2\eta^2 \iff \sqrt{\frac{RV}{2n}} \xrightarrow[n\to\infty]{\mathrm{p}} \eta$$

Computing the estimator

We need to use a large sample of size $n \to \infty$ to have best results for η, the micro-structure noise size estimator.

Methodology

Constructing the estimator

We know that, with η such estimator:

$$\frac{RV}{n} \xrightarrow[n\to\infty]{\mathrm{p}} 2\eta^2 \iff \sqrt{\frac{RV}{2n}} \xrightarrow[n\to\infty]{\mathrm{p}} \eta$$

Computing the estimator

We need to use a large sample of size $n \to \infty$ to have best results for η, the micro-structure noise size estimator.

Daily microstructure noise

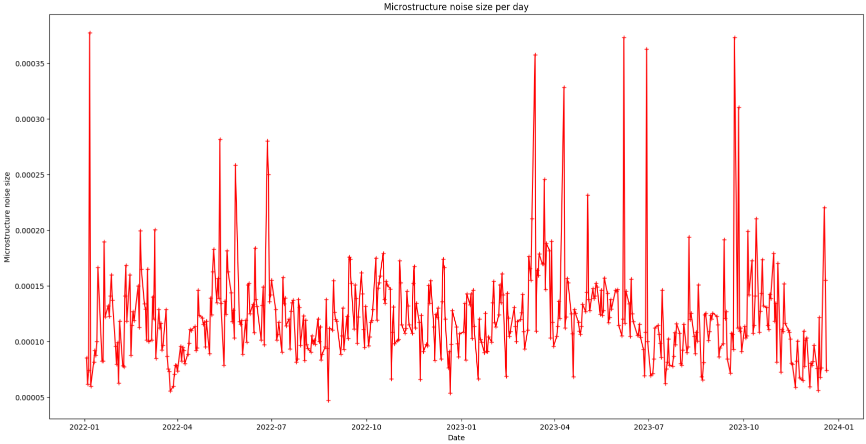

Figure: Daily microstructure noise

Autocorrelation method

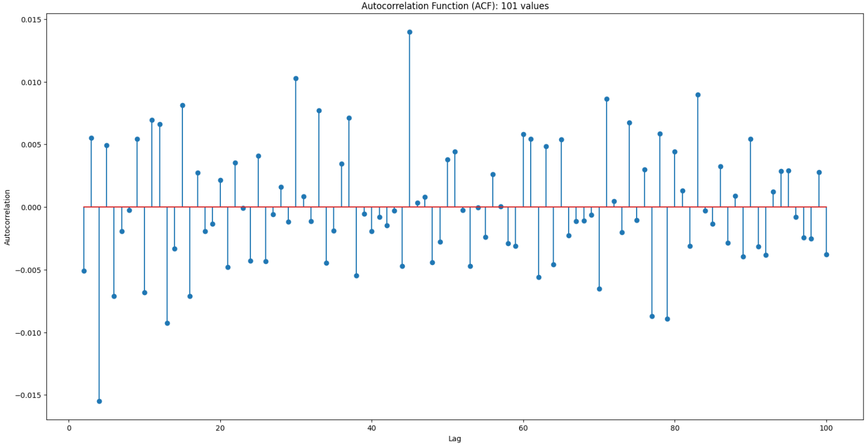

Figure: Micro structure noise after autocorrelation

Some remarks

- Both previous graphs shows that microstructure noise exists among our data set.
- The autocorrelation allows us to see how it impacts the prices (positively or negatively).

Some remarks

- Both previous graphs shows that microstructure noise exists among our data set.
- The autocorrelation allows us to see how it impacts the prices (positively or negatively).

Some remarks

- Both previous graphs shows that microstructure noise exists among our data set.
- The autocorrelation allows us to see how it impacts the prices (positively or negatively).

Computed estimated daily volatility

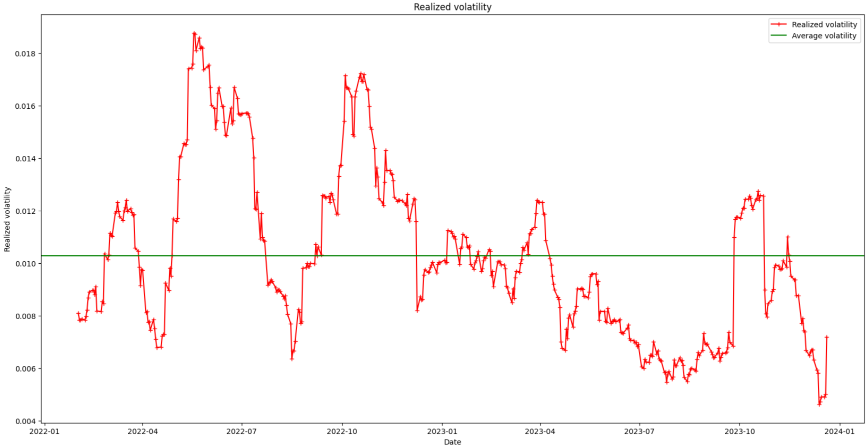

Figure: Estimated daily volatility over the last year (green is the average)

A few words to conclude

1. Analysis of realized volatility is quite easy to implement.
2. Difficulties when time interval between returns decreases.
3. Overnight returns: neglecting them leads to underestimate volatility. Our data wasn't accurate enough.
4. Various other methods developped to mitigate the microstructure noise.

A few words to conclude

1. Analysis of realized volatility is quite easy to implement.
2. Difficulties when time interval between returns decreases.
3. Overnight returns: neglecting them leads to underestimate volatility. Our data wasn't accurate enough.
4. Various other methods developped to mitigate the microstructure noise.

A few words to conclude

1. Analysis of realized volatility is quite easy to implement.
2. Difficulties when time interval between returns decreases.
3. Overnight returns: neglecting them leads to underestimate volatility. Our data wasn't accurate enough.
4. Various other methods developped to mitigate the microstructure noise.

A few words to conclude

1. Analysis of realized volatility is quite easy to implement.
2. Difficulties when time interval between returns decreases.
3. Overnight returns: neglecting them leads to underestimate volatility. Our data wasn't accurate enough.
4. Various other methods developped to mitigate the microstructure noise.

A few words to conclude

1. Analysis of realized volatility is quite easy to implement.
2. Difficulties when time interval between returns decreases.
3. Overnight returns: neglecting them leads to underestimate volatility. Our data wasn't accurate enough.
4. Various other methods developped to mitigate the microstructure noise.