Threshold estimation of jump-diffusion models and interest rate modeling

Roberto Renò

reno@unisi.it

Dipartimento di Economia Politica

Università di Siena

Joint work with Cecilia Mancini, Università di Firenze.

Overlay of the presentation

- The econometric problem

Overlay of the presentation

- The econometric problem
- Review of the literature

Overlay of the presentation

- The econometric problem
- Review of the literature
- The proposed solution

Overlay of the presentation

- The econometric problem
- Review of the literature
- The proposed solution
- Monte Carlo simulations

Overlay of the presentation

- The econometric problem
- Review of the literature
- The proposed solution
- Monte Carlo simulations
- The problem of proxying the short rate

Overlay of the presentation

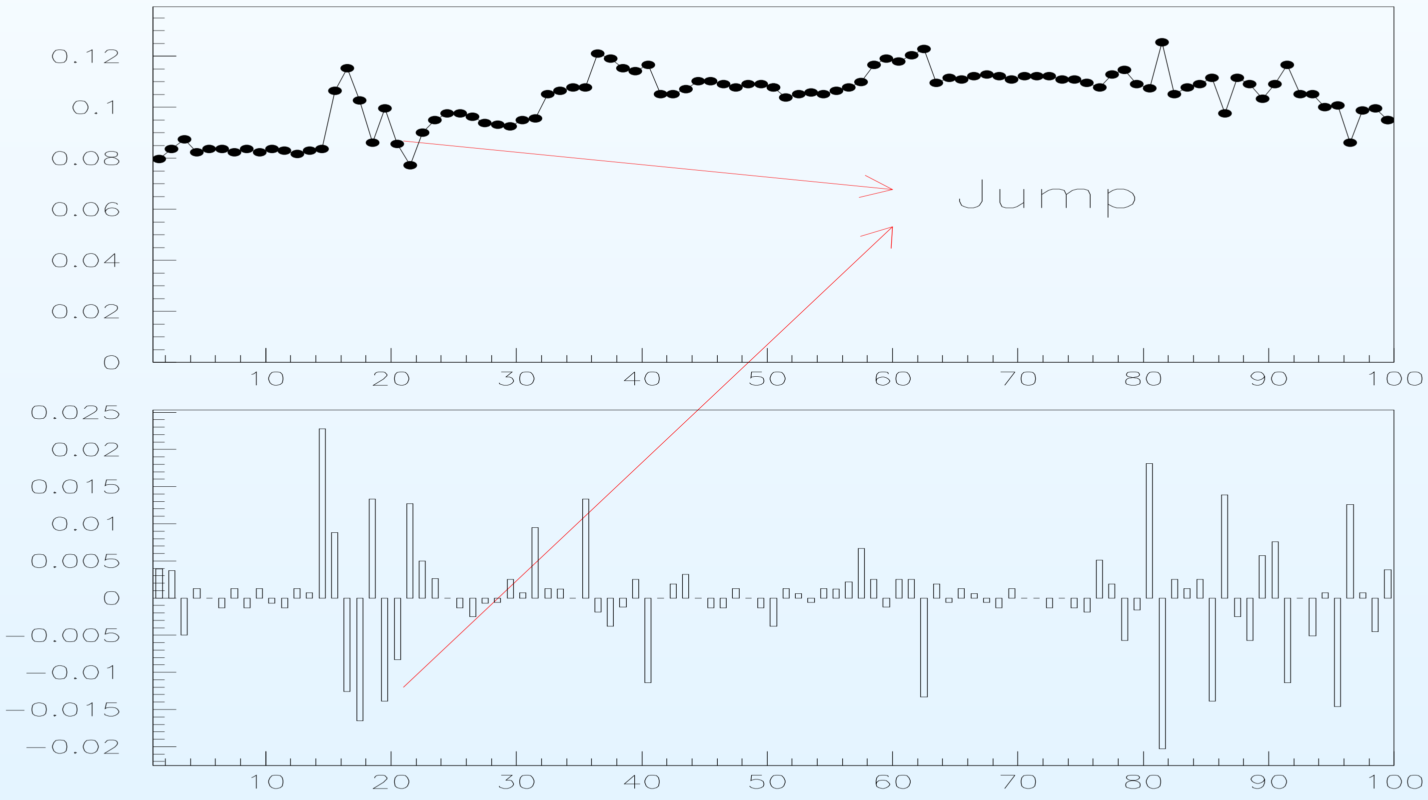
- The econometric problem
- Review of the literature
- The proposed solution
- Monte Carlo simulations
- The problem of proxying the short rate
- A final comparison

Overlay of the presentation

- The econometric problem
- Review of the literature
- The proposed solution
- Monte Carlo simulations
- The problem of proxying the short rate
- A final comparison
- Conclusions

The econometric problem

When we observe variations in the level of any state variable, are these variations continouous or discontinuous?



The econometric problem

When we observe variations in the level of any state variable, are these variations continouous or discontinuous?

In the financial econometrics terminology, can we disentangle diffusive variations from jumps?

The econometric problem

When we observe variations in the level of any state variable, are these variations continouous or discontinuous?

In the financial econometrics terminology, can we disentangle diffusive variations from jumps?

At a first sight, it looks like an hopeless problem: observations are discrete.

The econometric problem

When we observe variations in the level of any state variable, are these variations continouous or discontinuous?

In the financial econometrics terminology, can we disentangle diffusive variations from jumps?

At a first sight, it looks like an hopeless problem: observations are discrete.

Disentangling continuity from abrut changes needs using *infill* asymptotics.

The importance of jumps

The need for abrupt variations of the state variable comes, among others, from the following facts:

The importance of jumps

The need for abrupt variations of the state variable comes, among others, from the following facts:

- Interest rates, expecially short term, are influenced by macroeconomic news and announcement effects

The importance of jumps

The need for abrupt variations of the state variable comes, among others, from the following facts:

- Interest rates, expecially short term, are influenced by macroeconomic news and announcement effects
- Stock prices are influenced by endogenous crashes (e.g. Black Friday, 1987) or exogenous crashes (e.g. September 11)

The importance of jumps

The need for abrupt variations of the state variable comes, among others, from the following facts:

- Interest rates, expecially short term, are influenced by macroeconomic news and announcement effects
- Stock prices are influenced by endogenous crashes (e.g. Black Friday, 1987) or exogenous crashes (e.g. September 11)
- The term structure of smiles in option prices needs a jump component.

The importance of jumps

The need for abrupt variations of the state variable comes, among others, from the following facts:

- Interest rates, expecially short term, are influenced by macroeconomic news and announcement effects
- Stock prices are influenced by endogenous crashes (e.g. Black Friday, 1987) or exogenous crashes (e.g. September 11)
- The term structure of smiles in option prices needs a jump component.
- Commodity prices suffer abrupt spikes due to shortages.

The importance of jumps

The need for abrupt variations of the state variable comes, among others, from the following facts:

- Interest rates, expecially short term, are influenced by macroeconomic news and announcement effects
- Stock prices are influenced by endogenous crashes (e.g. Black Friday, 1987) or exogenous crashes (e.g. September 11)
- The term structure of smiles in option prices needs a jump component.
- Commodity prices suffer abrupt spikes due to shortages.
- Electricity prices are highly inelastic to demand, thus spikes are very frequent.

The literature: parametric models

To disentangle jumps from variation we can use:

- **Parametric models**

Examples are Das (2002), Pan (2002), Eraker, Johannes and Polson (2003), Andersen, Benzoni and Lund (2003), Chernov et al. (2003), Piazzesi (2005).

The literature: parametric models

To disentangle jumps from variation we can use:

- **Parametric models**

Examples are Das (2002), Pan (2002), Eraker, Johannes and Polson (2003), Andersen, Benzoni and Lund (2003), Chernov et al. (2003), Piazzesi (2005).

Test the significance of dJ_t, for example in:

$$\begin{aligned} dX_t &= \mu\, dt + \sqrt{v_t} dW_{x,t} + dJ_t, \\ d \log v_t &= (\alpha - \beta \log v_t)\, dt + \eta dW_{v,t}, \end{aligned} \tag{2}$$

The literature: parametric models

To disentangle jumps from variation we can use:

- **Parametric models**

Examples are Das (2002), Pan (2002), Eraker, Johannes and Polson (2003), Andersen, Benzoni and Lund (2003), Chernov et al. (2003), Piazzesi (2005).

Test the significance of dJ_t, for example in:

$$\begin{aligned} dX_t &= \mu\, dt + \sqrt{v_t} dW_{x,t} + dJ_t, \\ d\log v_t &= (\alpha - \beta \log v_t)\, dt + \eta dW_{v,t}, \end{aligned} \quad (3)$$

Ait-Sahalia (2002,2004) directly exploits the properties of likelihood function.

Power variation

Comes from the recent contributions of Barndorff-Nielsen and Shephard, (2004) and Woerner (2005).

- The realized power variation of order γ is defined as:

$$PV_\delta(X)_t^\gamma = \delta^{1-\gamma/2} \sum_{j=1}^{[t/\delta]} |\Delta_j X|^\gamma$$

Power variation

Comes from the recent contributions of Barndorff-Nielsen and Shephard, (2004) and Woerner (2005).

- The realized power variation of order γ is defined as:

$$PV_\delta(X)_t^\gamma = \delta^{1-\gamma/2} \sum_{j=1}^{[t/\delta]} |\Delta_j X|^\gamma$$

- the normalization term $\delta^{1-\gamma/2}$ disappears when $\gamma = 2$, goes to infinity when $\gamma > 2$ and goes to zero when $\gamma < 2$.

Power variation

Comes from the recent contributions of Barndorff-Nielsen and Shephard, (2004) and Woerner (2005).

- The realized power variation of order γ is defined as:

$$PV_\delta(X)_t^\gamma = \delta^{1-\gamma/2} \sum_{j=1}^{[t/\delta]} |\Delta_j X|^\gamma$$

- the normalization term $\delta^{1-\gamma/2}$ disappears when $\gamma = 2$, goes to infinity when $\gamma > 2$ and goes to zero when $\gamma < 2$.
-

$$p - \lim_{\delta \to 0} \mu_\gamma^{-1} PV_\delta(X)_t^\gamma = \begin{cases} \int_0^t \sigma_s^\gamma ds & if\ \gamma < 2 \\ [X]_t & if\ \gamma = 2 \\ +\infty & if\ \gamma > 2 \end{cases}$$

Bipower variation

We define *realized bipower variation* of order $[\gamma_1, \gamma_2]$ as:

$$BPV_\delta(X)_t^{[\gamma_1,\gamma_2]} = \delta^{1-\frac{1}{2}(\gamma_1+\gamma_2)} \sum_{j=2}^{[t/\delta]} |\Delta_{j-1}X|^{\gamma_1} \cdot |\Delta_j X|^{\gamma_2}$$

Bipower variation

We define *realized bipower variation* of order $[\gamma_1, \gamma_2]$ as:

$$BPV_\delta(X)_t^{[\gamma_1,\gamma_2]} = \delta^{1-\frac{1}{2}(\gamma_1+\gamma_2)} \sum_{j=2}^{[t/\delta]} |\Delta_{j-1}X|^{\gamma_1} \cdot |\Delta_j X|^{\gamma_2}$$

$$p - \lim_{\delta\to 0} \mu_{\gamma_1}^{-1} \mu_{\gamma_2}^{-1} BPV_\delta(X)_t^{[\gamma_1,\gamma_2]} =$$

$$= \begin{cases} \int_0^t \sigma_s^{\gamma_1+\gamma_2} ds & if\ max(\gamma_1,\gamma_2) < 2 \\ [X]_t & if\ max(\gamma_1,\gamma_2) = 2 \\ +\infty & if\ max(\gamma_1,\gamma_2) > 2 \end{cases}$$

Standard bipower variation

The $[1,1]$-order bipower variation, when it exists, is defined as:

$$BPV(X)_t^{[1,1]} = p-\lim_{\delta\to 0}\sum_{j=2}^{[t/\delta]} |\Delta_{j-1}X|\cdot|\Delta_j X|$$

Standard bipower variation

The $[1,1]$-order bipower variation, when it exists, is defined as:

$$BPV(X)_t^{[1,1]} = p - \lim_{\delta \to 0} \sum_{j=2}^{[t/\delta]} |\Delta_{j-1}X| \cdot |\Delta_j X|$$

In the case $\mu = 0$ and σ independent from W_t, we have:

$$BPV(X)_t^{[1,1]} = \mu_1 \int_0^t \sigma_s^2 ds = \mu_1^2 [X^c]_t$$

where

$$\mu_1 = \frac{\sqrt{2}}{\sqrt{\pi}}$$

This result can be extended to the case $\mu \neq 0$

Estimation based on sample moments

- An alternative estimator can be devised using nonparametric sample moment estimation (Bandi and Nguyen, 2003, Johannes, 2004)

Estimation based on sample moments

- An alternative estimator can be devised using nonparametric sample moment estimation (Bandi and Nguyen, 2003, Johannes, 2004)
- Suppose $\lambda = \lambda(X_t)$ and the jump sizes $c_i = C$ are all identically distributed. Then:

$$\lim_{\delta \to 0} \frac{1}{\delta} \mathbb{E}_t \left[X_{t+\delta} - X_t \right] = \mu(X_t)$$

$$\lim_{\delta \to 0} \frac{1}{\delta} \mathbb{E}_t \left[(X_{t+\delta} - X_t)^2 \right] = \sigma(X_t) + \lambda(X_t) \mathbb{E} \left[C^2 \right]$$

$$\lim_{\delta \to 0} \frac{1}{\delta} \mathbb{E}_t \left[(X_{t+\delta} - X_t)^k \right] = \lambda(X_t) \mathbb{E} \left[C^k \right]$$

Threshold estimation

Mancini (2003,2007) disentangles diffusion from jumps using the *modulus of continuity* of the Brownian motion:

$$r(\delta) = \sqrt{2\delta \log \frac{1}{\delta}}$$

which has the following property, as established by Lévy:

$$\mathcal{P}\left[\limsup_{\delta\to 0} \frac{\max_{|t-s|\leq\delta} |W(t) - W(s)|}{r(\delta)} = 1\right] = 1$$

It measures the *speed* at which the Brownian motion shrinks to zero.

The intuition

When $\delta \to 0$, diffusive variations go to zero, while jumps do not.

Moreover, we know the rate at which the diffusive variations shrink to zero: the modulus of continuity.

The intuition

When $\delta \longrightarrow 0$, diffusive variations go to zero, while jumps do not.

Moreover, we know the rate at which the diffusive variations shrink to zero: the modulus of continuity.

Thus, we can identify the jumps as those variations which are larger than a suitable *threshold* $\vartheta(\delta)$ which goes to zero, as $\delta \longrightarrow 0$, slower than $r(\delta)$.

The theorem (Mancini, 2007)

Suppose

$$X = Y + J$$

Where Y is a Brownian martingale plus drift and J is a jump process with counting process N with $E[N_T] < \infty$ and time horizon $T < \infty$.

If $\vartheta(\delta)$ is a real deterministic function such that

$$\lim_{\delta \to 0} \vartheta(\delta) = 0 \quad \text{and} \quad \lim_{\delta \to 0} \frac{\delta \log \frac{1}{\delta}}{\vartheta(\delta)} = 0$$

then for P-almost all ω, $\exists \bar{\delta}(\omega)$ such that $\forall \delta < \bar{\delta}(\omega)$ we have

$$\forall i = 1, \ldots, n, \quad I_{\{\Delta N = 0\}}(\omega) = I_{\{(\Delta X)^2 \leq \vartheta(\delta)\}}(\omega).$$

Threshold power and multipower variation

Power variation can be generalized with the following definition of threshold power variation, for $\gamma \geq 1$:

$$tPV_\delta(X)_t^{[\gamma]} = \delta^{1-\frac{1}{2}\gamma} \sum_{j=1}^{[t/\delta]} |\Delta_j X|^\gamma I_{\{|\Delta_j X|^2 \leq \vartheta_j(\delta)\}} \longrightarrow \mu_\gamma \int_0^t \sigma_s^\gamma ds \qquad (4)$$

Threshold power and multipower variation

Power variation can be generalized with the following definition of threshold power variation, for $\gamma \geq 1$:

$$tPV_{\delta}(X)_t^{[\gamma]} = \delta^{1-\frac{1}{2}\gamma} \sum_{j=1}^{[t/\delta]} |\Delta_j X|^{\gamma} I_{\{|\Delta_j X|^2 \leq \vartheta_j(\delta)\}} \longrightarrow \mu_{\gamma} \int_0^t \sigma_s^{\gamma} ds \qquad (6)$$

Limit theorems in law for tPV in the case $\gamma = 2$ are established in Mancini (2007).

Threshold power and multipower variation

Power variation can be generalized with the following definition of threshold power variation, for $\gamma \geq 1$:

$$tPV_\delta(X)_t^{[\gamma]} = \delta^{1-\frac{1}{2}\gamma} \sum_{j=1}^{[t/\delta]} |\Delta_j X|^\gamma I_{\{|\Delta_j X|^2 \leq \vartheta_j(\delta)\}} \longrightarrow \mu_\gamma \int_0^t \sigma_s^\gamma ds \qquad (8)$$

Limit theorems in law for tPV in the case $\gamma = 2$ are established in Mancini (2007).

We define the (realized) threshold multipower variation as:

$$TMPV_\delta(X)_t^{[\gamma_1,\ldots,\gamma_M]} = \delta^{1-\frac{1}{2}(\gamma_1+\ldots+\gamma_M)} \sum_{j=M}^{[t/\delta]} \prod_{k=1}^{M} |\Delta_{j-k+1} X|^{\gamma_k} I_{\{|\Delta_{j-k+1} X|^2 \leq \vartheta_{j-k+1}(\delta)\}} \qquad (9)$$

Threshold multipower variation asymptotics

As $\delta \longrightarrow 0$ we have (Pirino and Renò, 2007)

1. if $\max(\gamma_1, \ldots, \gamma_M) < 2$ we have:

$$TMPV_\delta(X)_t^{[\gamma_1,\ldots,\gamma_M]} \longrightarrow \left(\prod_{k=1}^{M} \mu_{\gamma_k}\right) \int_0^t \sigma_s^{\gamma_1+\ldots+\gamma_M} ds \qquad (10)$$

   where the above convergence is in probability.

Threshold multipower variation asymptotics

As $\delta \longrightarrow 0$ we have (Pirino and Renò, 2007)

1. if $\max(\gamma_1, \ldots, \gamma_M) < 2$ we have:

$$TMPV_\delta(X)_t^{[\gamma_1,\ldots,\gamma_M]} \longrightarrow \left(\prod_{k=1}^{M} \mu_{\gamma_k}\right) \int_0^t \sigma_s^{\gamma_1+\ldots+\gamma_M} ds \qquad (12)$$

   where the above convergence is in probability.

2. if $\max(\gamma_1, \ldots, \gamma_M) < 1$ we have:

$$\delta^{-\frac{1}{2}} \left(TMPV_\delta(X)_t - \int_0^t \sigma_s^{\gamma_1+\ldots+\gamma_M} ds \right) \longrightarrow c_\gamma \int_0^t \sigma_s^{\gamma_1+\ldots+\gamma_M} dW_s' \qquad (13)$$

   where the above convergence is in law.

The model

We use the above intuition to estimate univariate models of the kind:

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t + dJ_t, \qquad t \in [0, T],$$

The model

We use the above intuition to estimate univariate models of the kind:

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t + dJ_t, \qquad t \in [0,T],$$

Drift and diffusion are level-dependent. J_t can be a finite activity doubly stochastic Poisson process with level-dependent intensity $\lambda(X_t)$, or a Lévy process with finite or infinite activity.

The model

We use the above intuition to estimate univariate models of the kind:

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t + dJ_t, \qquad t \in [0,T],$$

Drift and diffusion are level-dependent. J_t can be a finite activity doubly stochastic Poisson process with level-dependent intensity $\lambda(X_t)$, or a Lévy process with finite or infinite activity.

This is the first estimator of the diffusion coefficient under an infinite activity jump component.

A primer to nonparametric estimation

The advantage of nonparametric estimation is clearly its flexibility.

A primer to nonparametric estimation

The advantage of nonparametric estimation is clearly its flexibility.

We introduce the setup through a well-known example.

A primer to nonparametric estimation

The advantage of nonparametric estimation is clearly its flexibility.

We introduce the setup through a well-known example.

Suppose that n observations $X_1, \ldots, X_n$ are drawn from an unknown density $f(x)$, and we want to estimate the density.

Histograms

One popular way to estimate the density is the histogram.
The estimate through an histogram can be written:

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^{n} I_{\{|X_i - x| < h/2\}}$$

where h is a bandwith parameter (the histogram bin).

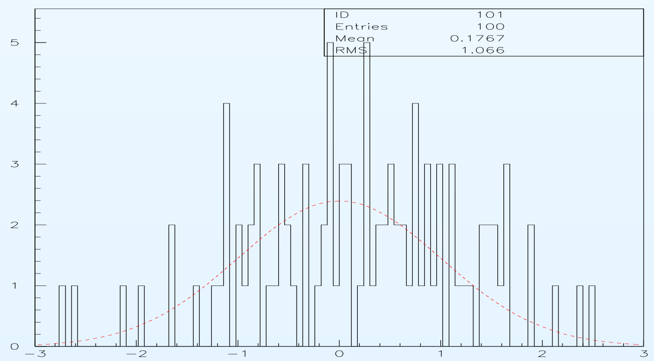
Histograms

One popular way to estimate the density is the histogram.
The estimate through an histogram can be written:

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^{n} I_{\{|X_i - x| < h/2\}}$$

where h is a bandwith parameter (the histogram bin).

Small h



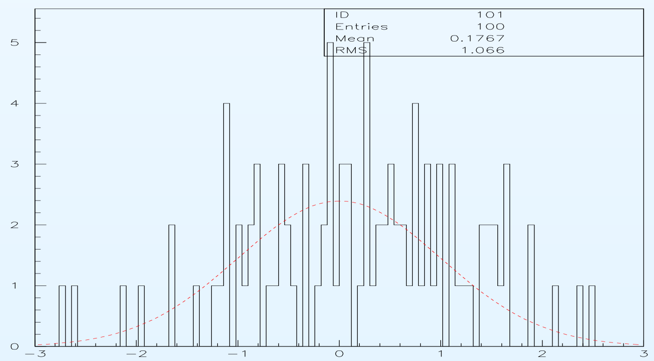
Histograms

One popular way to estimate the density is the histogram.
The estimate through an histogram can be written:

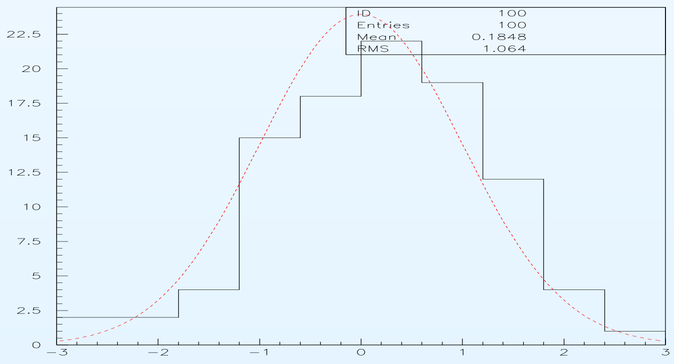
$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} I_{\{|X_i - x| < h/2\}}$$

where h is a bandwith parameter (the histogram bin).

Small h



Large h



Histograms and empirical CDF

- The empirical CDF is

$$F(x) = \frac{1}{n} \sum_{i=1}^{n} I_{\{X_i \leq x\}}$$

Histograms and empirical CDF

- The empirical CDF is

$$F(x) = \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i \leq x\}}$$

- it is a nonsense to differentiate it to get an estimate of $f(x)$

Histograms and empirical CDF

- The empirical CDF is

$$F(x) = \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i \leq x\}}$$

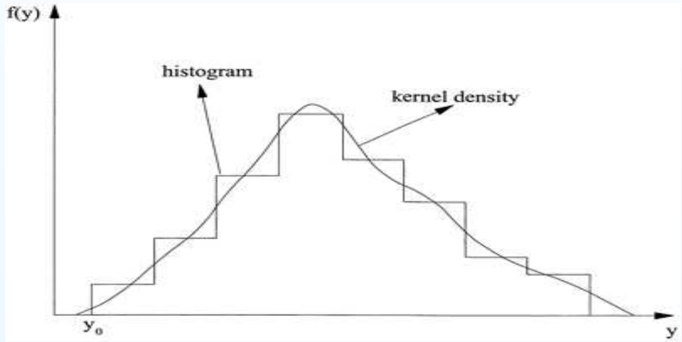
- it is a nonsense to differentiate it to get an estimate of $f(x)$
- but

$$\hat{f}(x) = \frac{F(x+\frac{h}{2}) - F(x-\frac{h}{2})}{h} = \frac{1}{nh}\sum_{i=1}^{n} I_{\{|X_i - x| < h/2\}}$$

Kernel smoothing

Histograms are not smooth by definition. We can obtain a smoother estimate replacing the indicator function with a continuous function, the *kernel* $K(\cdot)$:

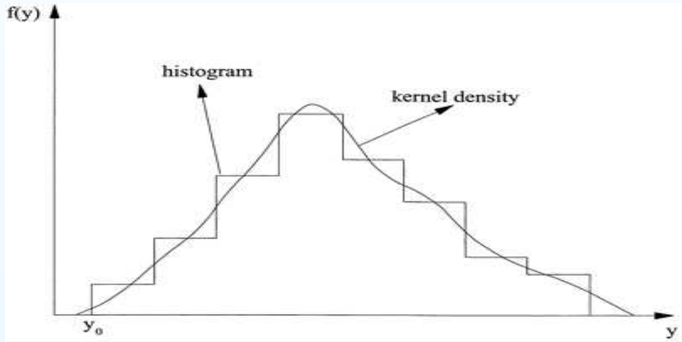
$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{X_i - x}{h}\right)$$



Kernel smoothing

Histograms are not smooth by definition. We can obtain a smoother estimate replacing the indicator function with a continuous function, the *kernel* $K(\cdot)$:

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{X_i - x}{h}\right)$$



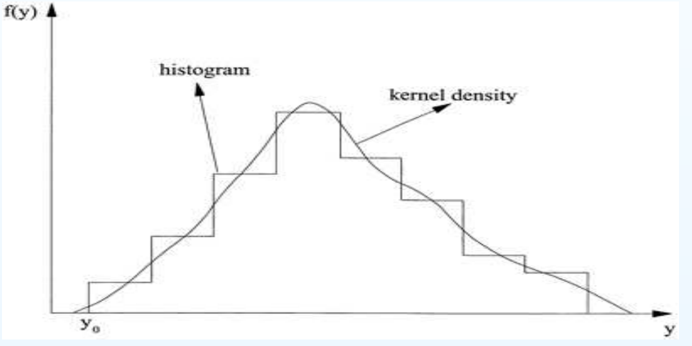
A kernel is a function $K : \mathbb{R} \in \mathbb{R}^{+}$ such that

$$\int_{\mathbb{R}} K(s)ds = 1, \quad \int_{\mathbb{R}} sK(s)ds = 0, \quad \int_{\mathbb{R}} s^2K(s)ds = \sigma_K^2 < +\infty$$

Kernel smoothing

Histograms are not smooth by definition. We can obtain a smoother estimate replacing the indicator function with a continuous function, the *kernel* $K(\cdot)$:

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{X_i - x}{h}\right)$$



A kernel is a function $K : \mathbb{R} \in \mathbb{R}^+$ such that

$$\int_{\mathbb{R}} K(s)ds = 1, \quad \int_{\mathbb{R}} sK(s)ds = 0, \quad \int_{\mathbb{R}} s^2K(s)ds = \sigma_K^2 < +\infty$$

A bandwidth h is a sequence of positive real number such that

$$\lim_{n\to\infty} h = 0, \quad \lim_{n\to\infty} nh = +\infty$$

Smooth histograms

One example with $K(s) = \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}}$, the Gaussian kernel.

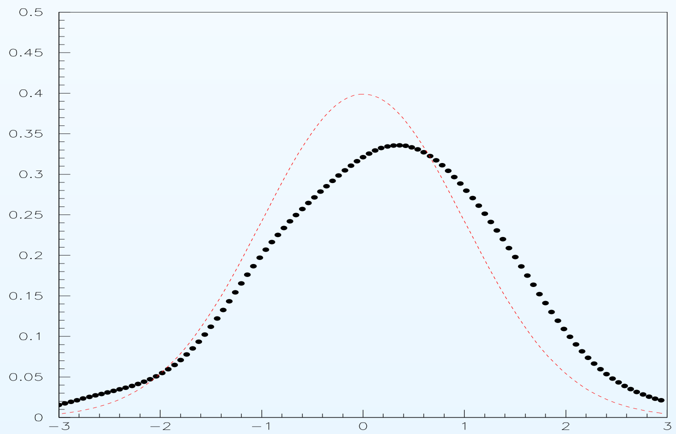
The choice of h has an impact on the estimate:

Smooth histograms

One example with $K(s) = \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}}$, the Gaussian kernel.

The choice of h has an impact on the estimate:

optimal h

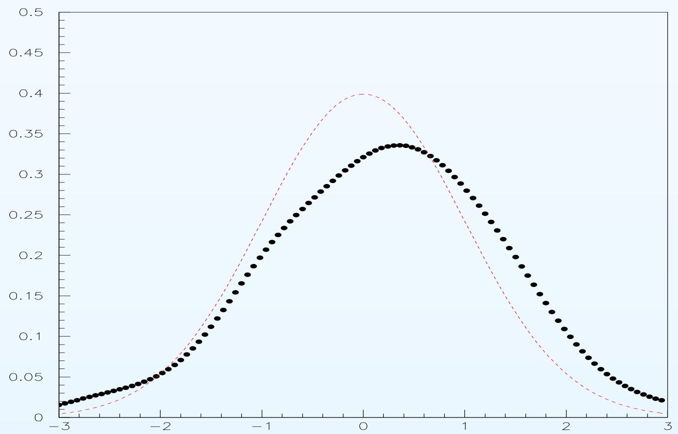


Smooth histograms

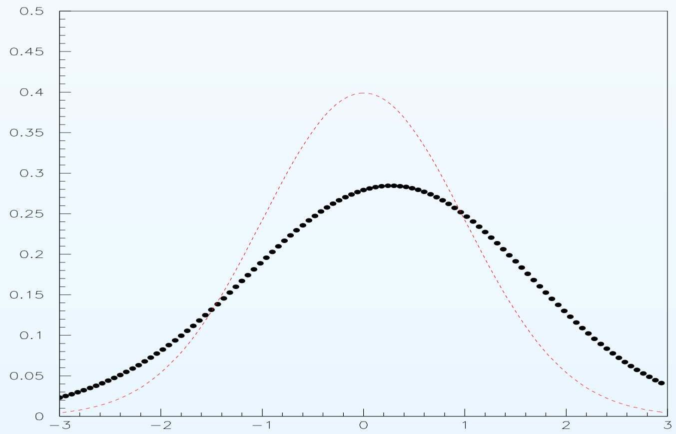
One example with $K(s) = \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}}$, the Gaussian kernel.

The choice of h has an impact on the estimate:

optimal h

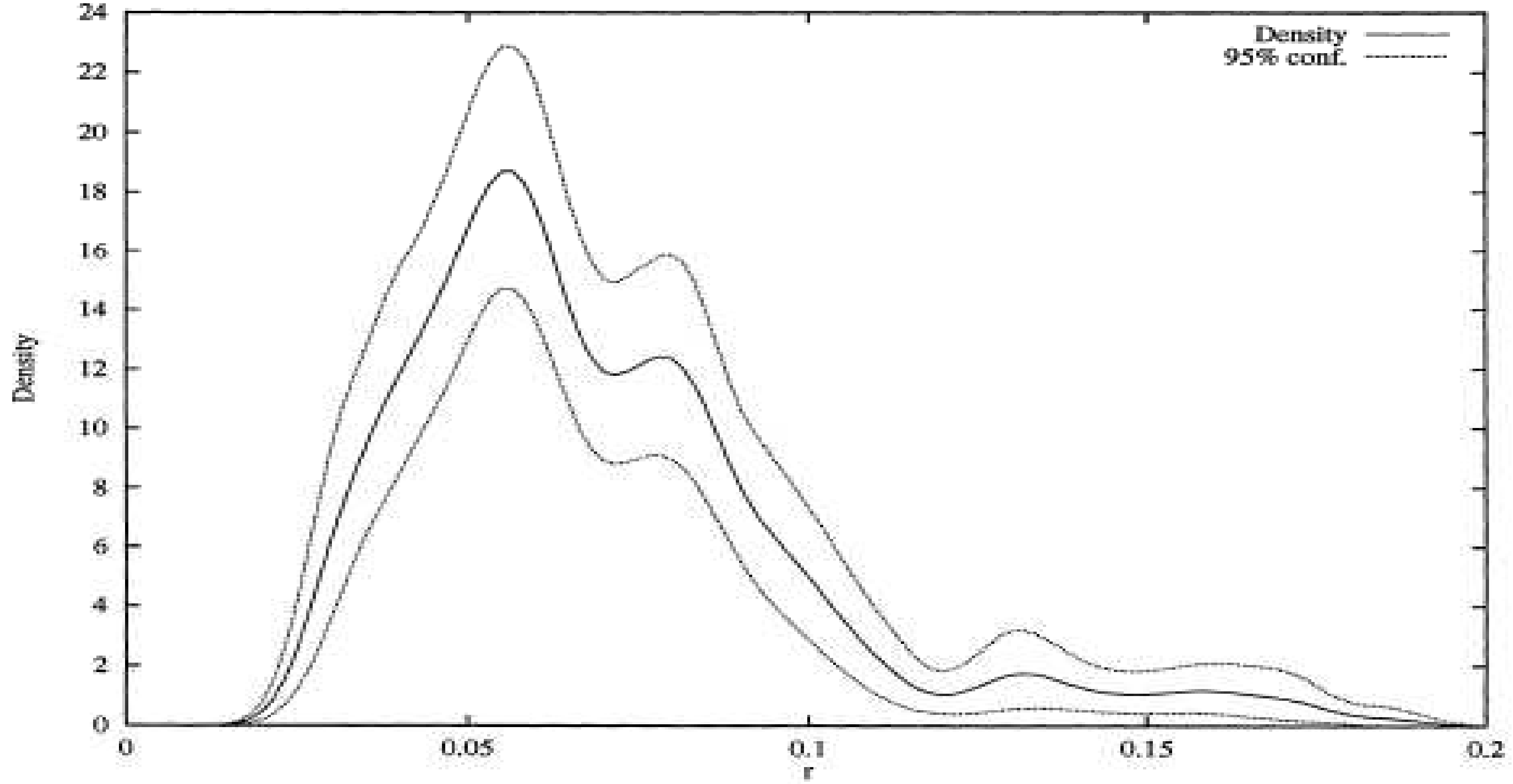


oversmoothing



Interest rate density

Font: Stanton (1997)



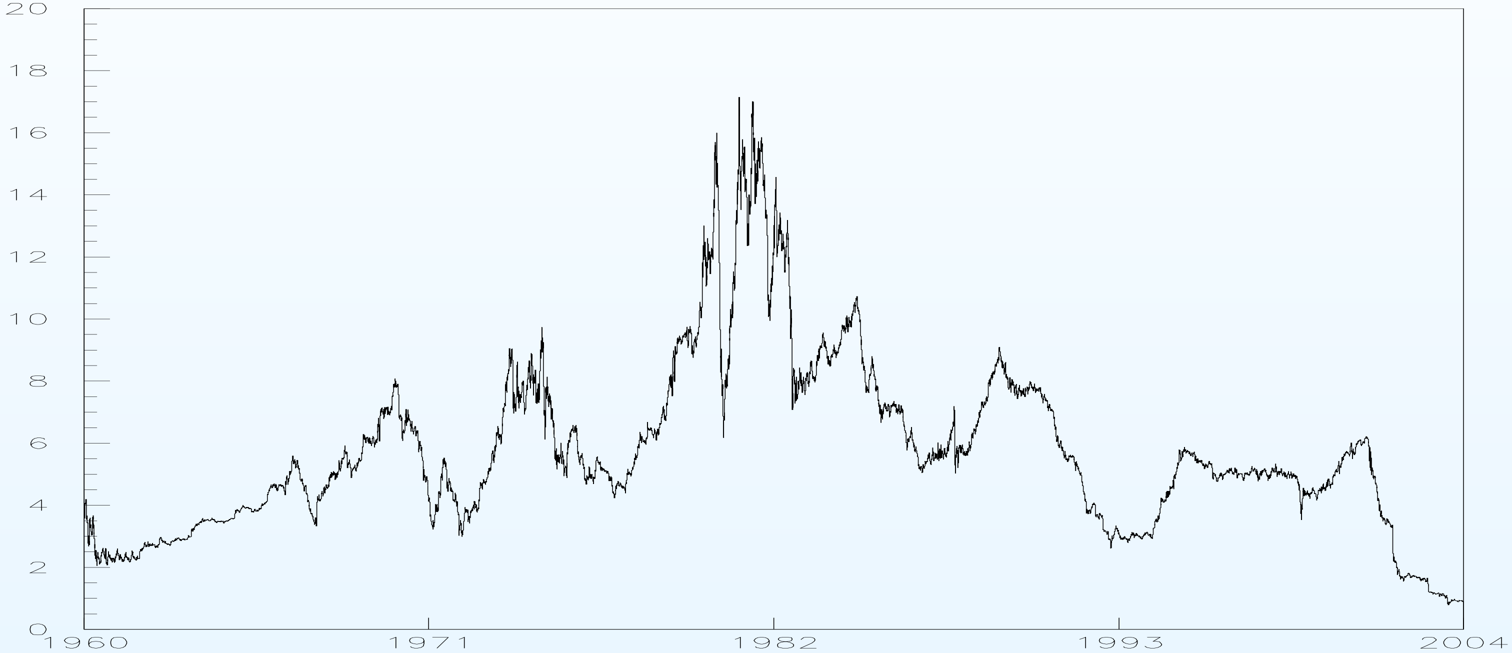
Nonparametric estimation: the continuous case

A popular estimator is that proposed by Florens-Zmirou (1993), and subsequently by Stanton (1997) and Jiang and Knight (1998).
The estimator is very intuitive:

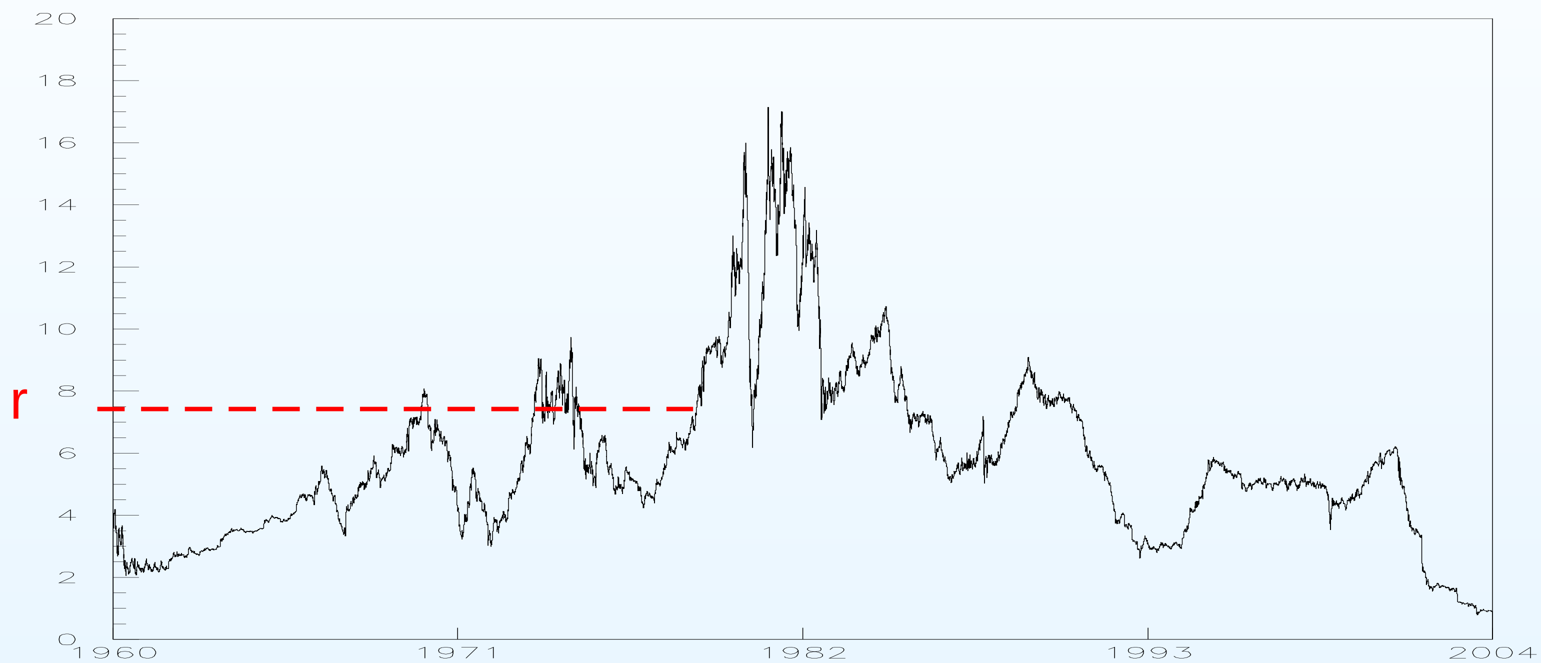
$$\hat{\sigma}^2(r) = \frac{n \sum_{t=1}^{n-1} K\left(\frac{r-\hat{r}_t}{h}\right) (\hat{r}_{t+1} - \hat{r}_t)^2}{T \sum_{t=1}^{n} K\left(\frac{r-\hat{r}_t}{h}\right)}.$$

Asymptotic properties are fully assessed in Florens-Zmirou (1993).
A generalization is given by Bandi and Phillips (2003).

The intuition behind non-parametric estimation



The intuition behind non-parametric estimation

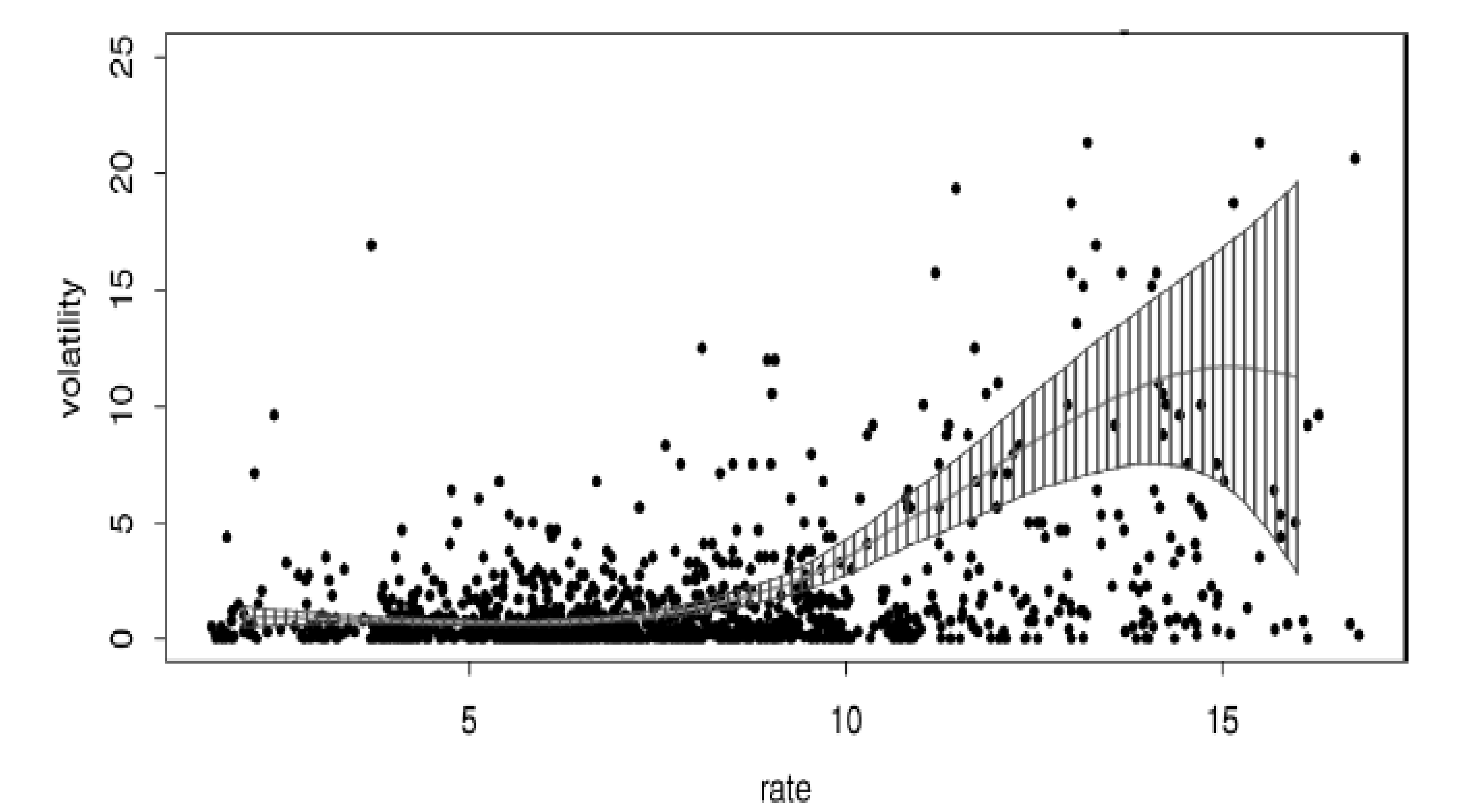


The intuition behind non-parametric estimation

Estimate the quadratic variation here!

The diffusion coefficient estimator

Font: Fan (2003)



Bandi-Phillips estimators

Bandi and Phillips (2003) propose an alternative estimator which is very close to that of Florens-Zmirou (1993):

$$\hat{\sigma}^2(r) = \frac{n \sum_{i=1}^{n} K\left(\frac{r-\hat{r}_i}{h}\right)\left(\frac{1}{m_i}\sum_{j=0}^{m_i}\left[\hat{r}_{t_{i,j}+1} - \hat{r}_{t_{i,j}}\right]^2\right)}{T \sum_{i=1}^{n} K\left(\frac{r-\hat{r}_i}{h}\right)}$$

where $t_{i,j}$ is a subset of indexes such that

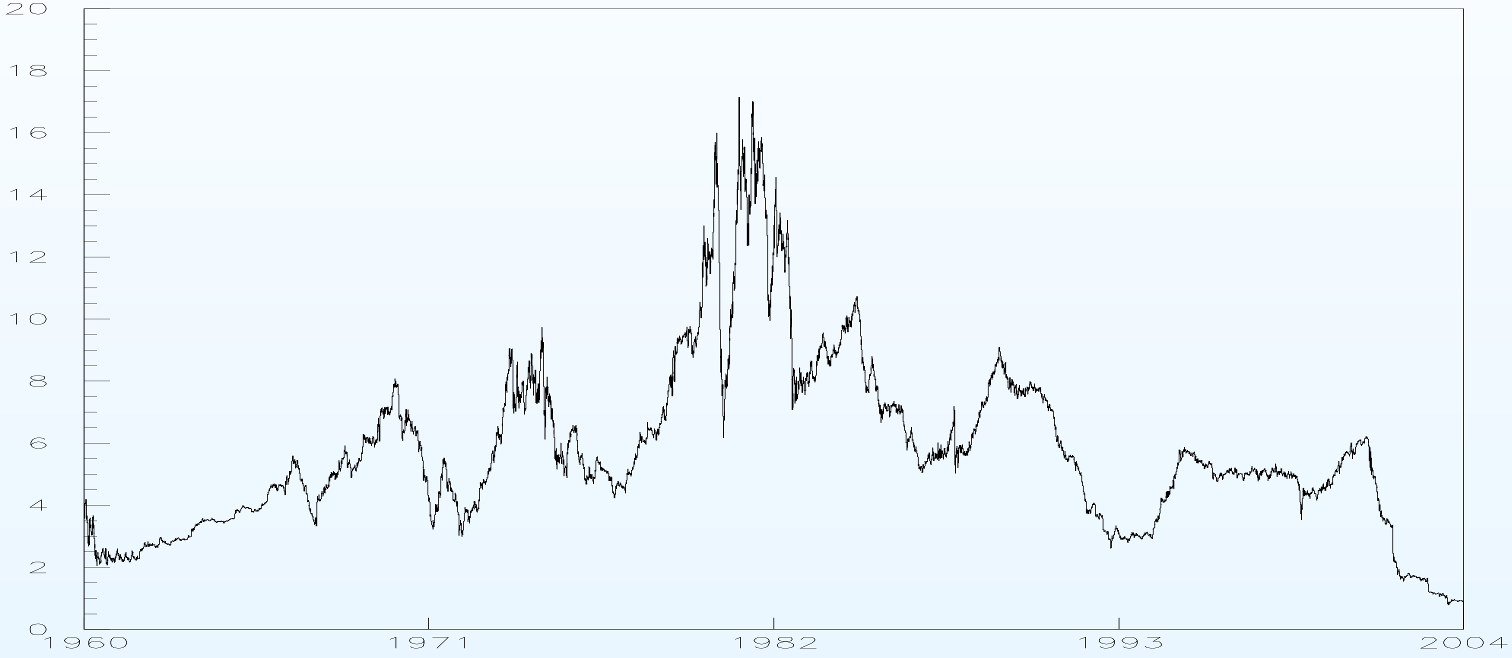
$$t_{i,0} = \inf\{t \geq 0 : |\hat{r}_t - \hat{r}_i| \leq \varepsilon_s\},$$

and

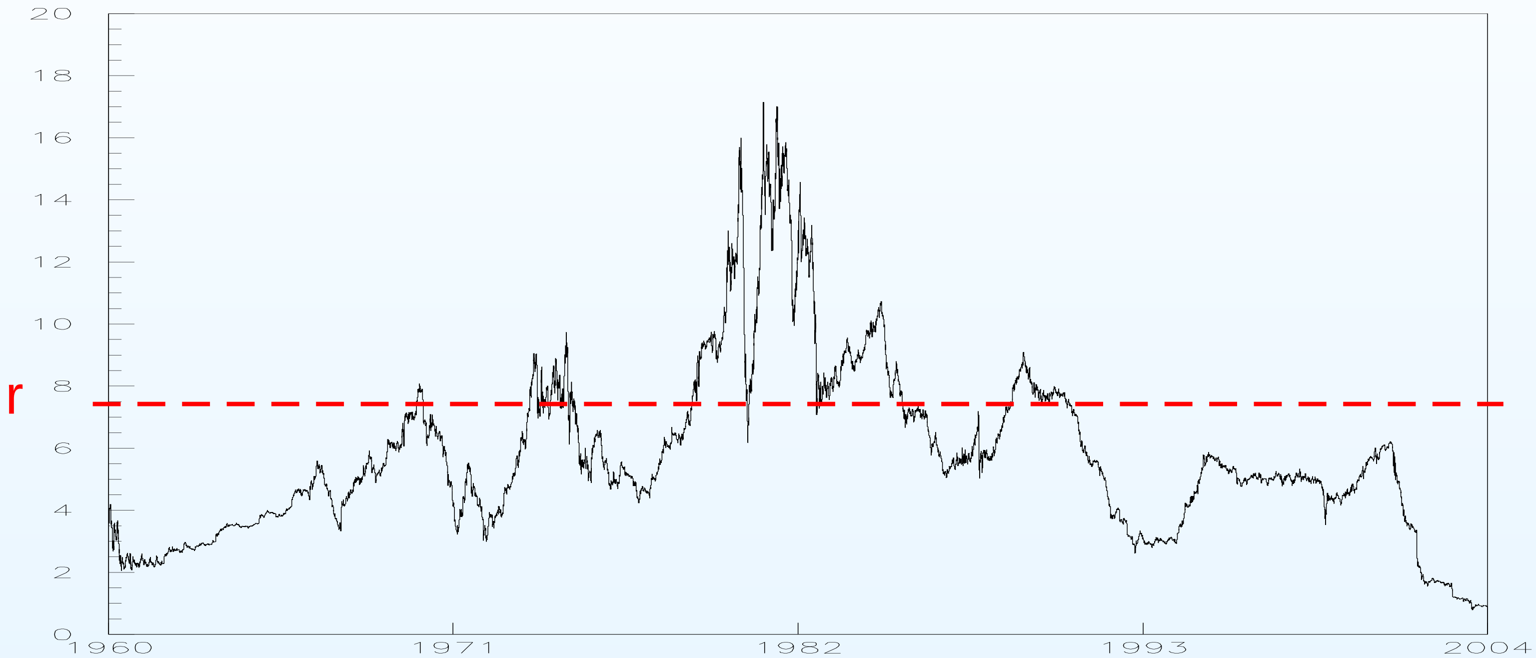
$$t_{i,j+1} = \inf\left\{t \geq t_{i,j} + 1 : |\hat{r}_t - \hat{r}_i| \leq \varepsilon_s\right\},$$

m_i is the number of times that $|\hat{r}_t - \hat{r}_i| \leq \varepsilon_s$ and ε_s is a parameter.

The intuition of Bandi-Phillips

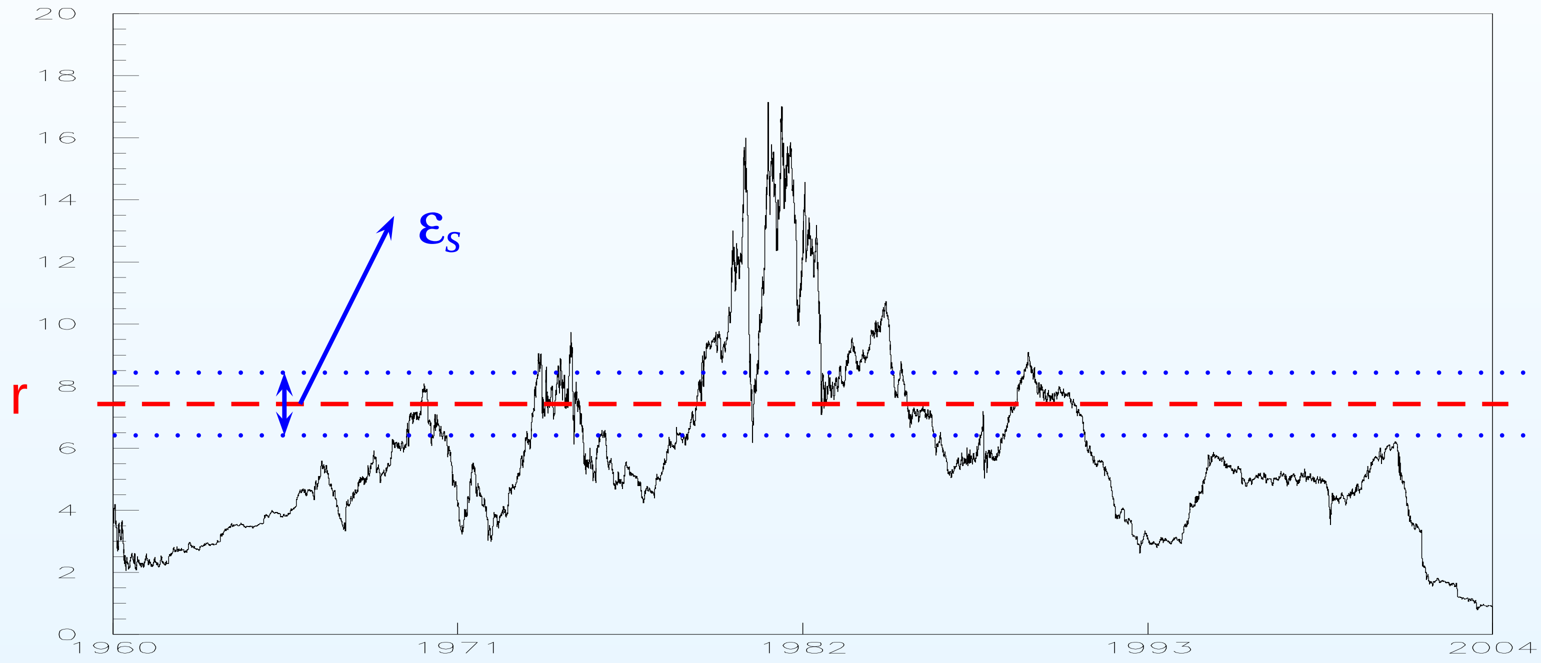


The intuition of Bandi-Phillips



The intuition of Bandi-Phillips

The intuition of Bandi-Phillips



Estimate the quadratic variation in the strip

Estimation of the jump process

We first get (in the case of finite activity) an estimate of the whole jump process using:

$$\hat{J}_{1,t} = \sum_{\{i: t_i \le t\}} \hat{\gamma}_{\tau^{(i)}},$$

where

$$\hat{\gamma}_{\tau^{(i)}} \doteq (X_{t_{i+1}} - X_{t_i}) \cdot I_{\{(X_{t_{i+1}} - X_{t_i})^2 > \vartheta(\delta)\}}$$

It is remarkable that we estimate contemporaneously jump times and sizes.

The proposed estimator

Let X be a jump-diffusion process with finite activity, and assume that:

- $P\{\gamma_\ell = 0\} = 0$;
- $t_i = i\delta$ (equally spaced observations)
- as $\delta \to 0$ both the threshold function $\vartheta(\delta)$ and $\frac{\delta \ln \frac{1}{\delta}}{\vartheta(\delta)}$ tend to zero;
- $nh^4 \to 0$ as $n \to \infty$ and $\exists \beta > 1 : nh^\beta \to \infty$.

Then for all x visited by X

$$\hat{\sigma}^2_{n,h}(x) = \frac{n \sum_{i=1}^{n} K\left(\frac{X_{t_i} - \hat{J}_{1,t_i} - x}{h}\right) (X_{t_{i+1}} - X_{t_i})^2 I_{\{(X_{t_{i+1}} - X_{t_i})^2 \leq \vartheta(\delta)\}}}{T \sum_{i=1}^{n} K\left(\frac{X_{t_i} - \hat{J}_{1,t_i} - x}{h}\right)} \to_P \sigma^2(x)$$

Threshold setting

How do we set the threshold $\vartheta(\delta)$?

Threshold setting

How do we set the threshold $\vartheta(\delta)$?

If ϑ is too small $\longrightarrow$ many variations will be detected as jumps.

Threshold setting

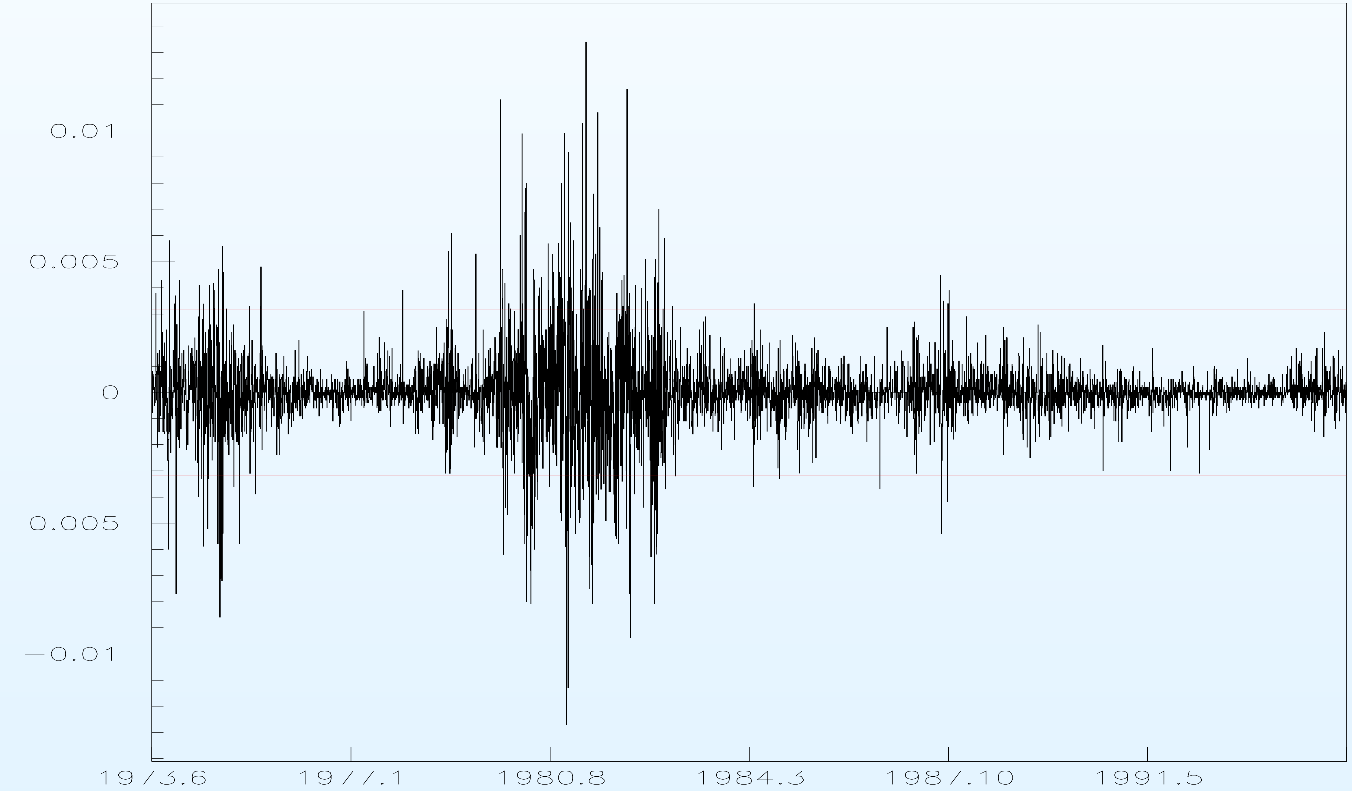
How do we set the threshold $\vartheta(\delta)$?

If ϑ is too small $\longrightarrow$ many variations will be detected as jumps.

If ϑ is too large $\longrightarrow$ many jumps will not be detected.

An auxiliary model for the threshold

Moreover, we need a non-constant threshold to take into account volatility persistence.



An auxiliary model for the threshold

Moreover, we need a non-constant threshold to take into account volatility persistence.

We use a GARCH(1,1) model. Even if misspecified, it asymptotically provides the optimal volatility forecast (Nelson and Foster, 1994).

We then fit:

$$r_t - r_{t-1} = \bar{r} + \sqrt{h_t} \cdot \varepsilon_t$$
$$h_t = \omega + \alpha(r_{t-1} - r_{t-2})^2 + \beta h_{t-1}$$

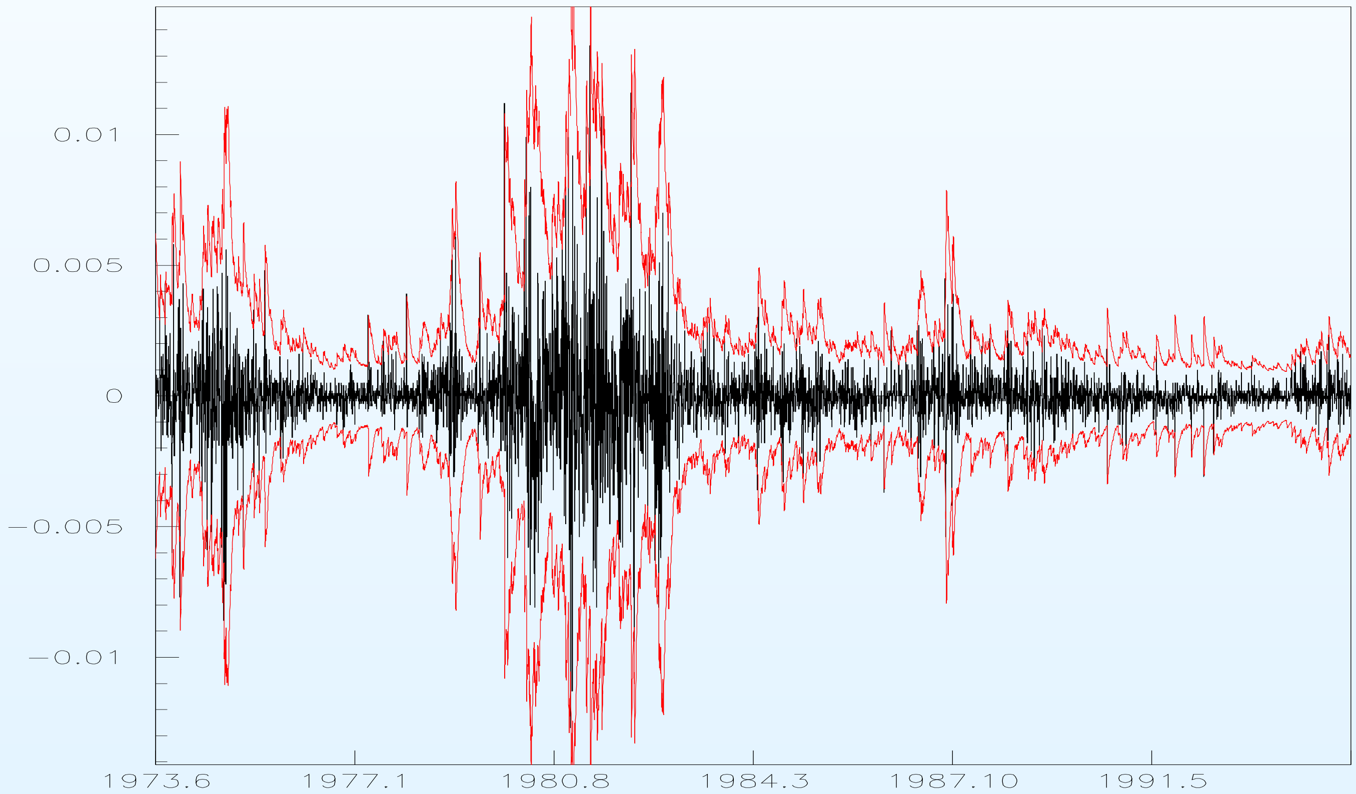
then we set:

$$\vartheta(t) = c \cdot h_t$$

with $c = 9$ (three standard deviations).

An auxiliary model for the threshold

Moreover, we need a non-constant threshold to take into account volatility persistence.



The simulated model

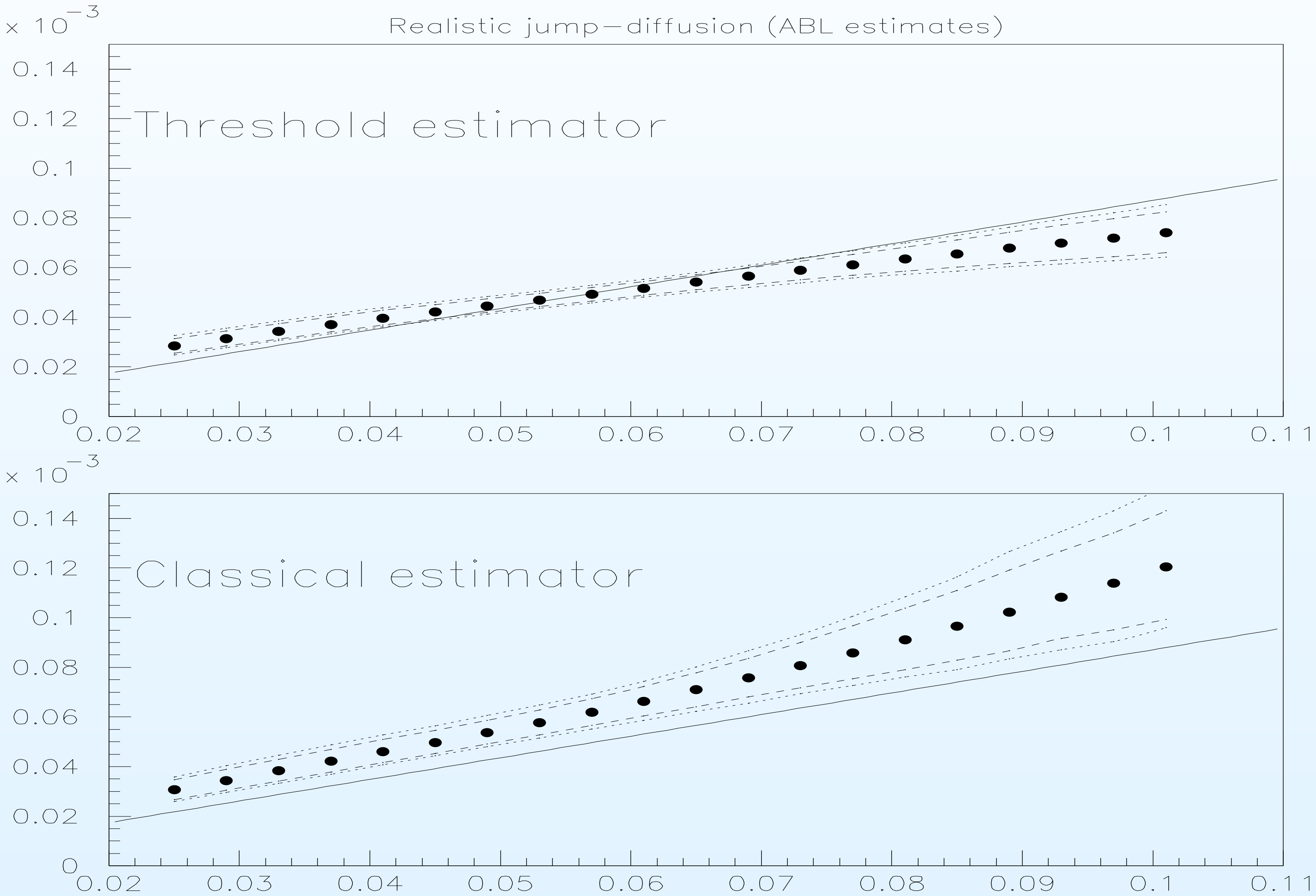
$$
\begin{aligned}
dr_t &= \kappa_1(b_t - r_t)dt - \bar{\kappa}\lambda r_t dt + \gamma\sqrt{r_t}dW_{1,t} + (e^{Z_t} - 1)r_t dN_t \\
db_t &= \kappa_3(\theta - b_t)dt + \eta_2\sqrt{b_t}dW_{2,t}, \\
Z_t &\simeq \mathcal{N}(\mu_J, \sigma_J^2), \\
\bar{\kappa} &= \mathbb{E}\left[e^{Z_t} - 1\right] = e^{\left(\mu_J + \frac{1}{2}\sigma_J^2\right)},
\end{aligned}
$$

This is a modified version of the models estimated by Andersen, Benzoni and Lund (2004) on a time series of 3-months Treasury Bills annualized rates.

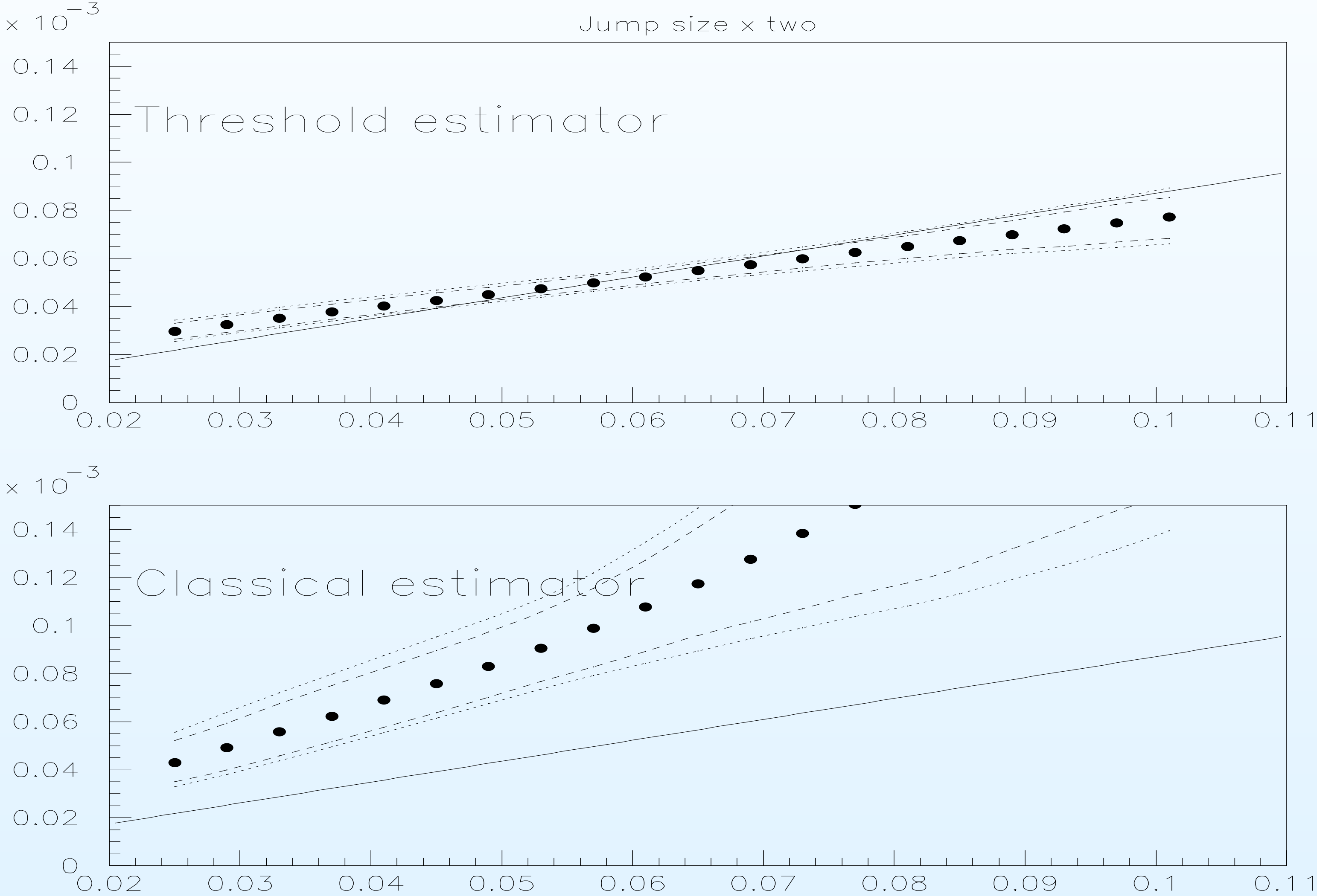
We use their estimates of the $SV_1J - SD$ model.

The starting values r_0 and b_0 are sampled from the unconditional distribution.

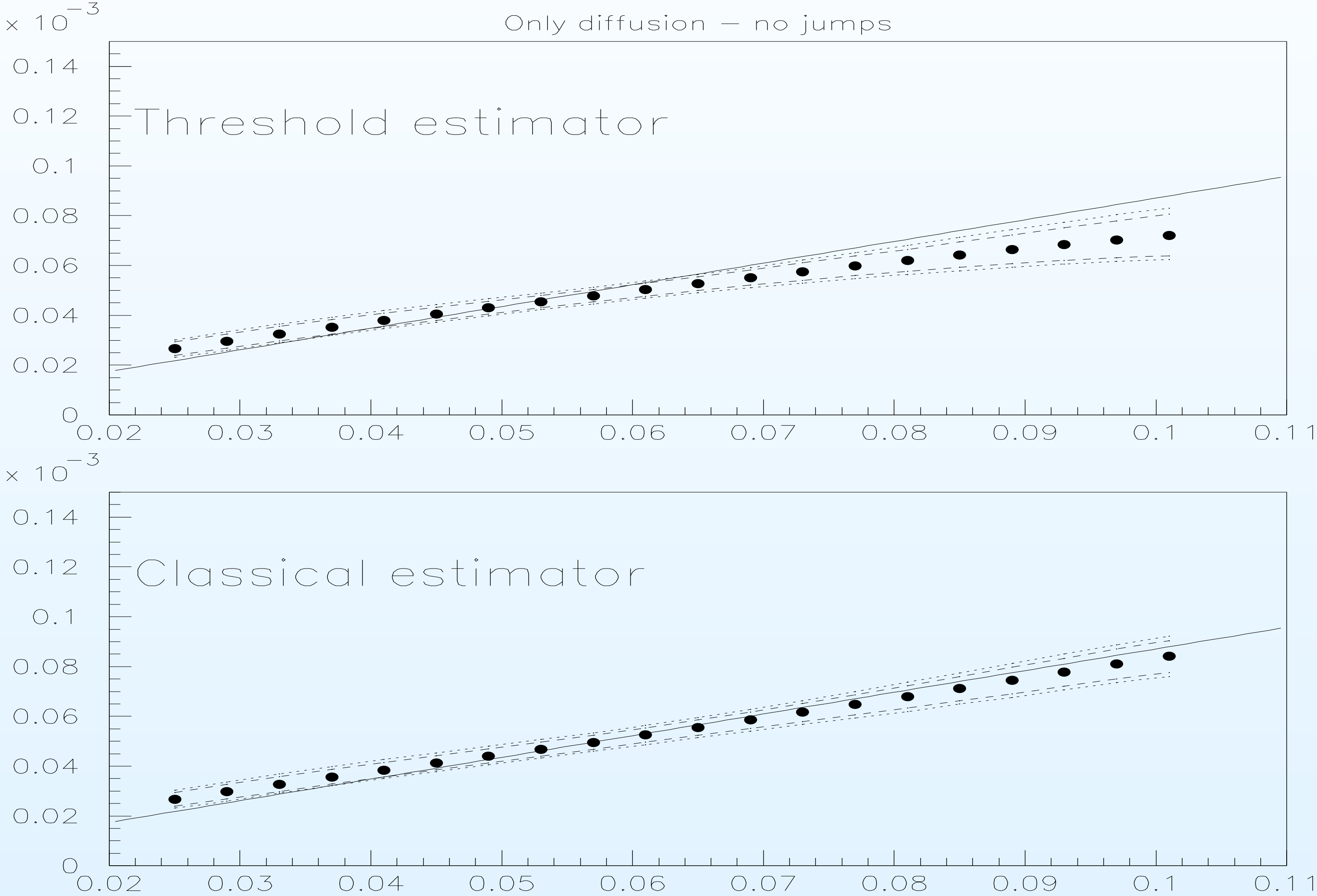
Simulations - realistic jump sizes



Simulations - doubled jump sizes



Simulations - no jumps



Estimating the drift function

When J is a doubly stochastic Poisson process it is possible to estimate the drift and the jump intensity functions by letting $T \longrightarrow \infty$ and $T/n \longrightarrow 0$.

Estimating the drift function

When J is a doubly stochastic Poisson process it is possible to estimate the drift and the jump intensity functions by letting $T \longrightarrow \infty$ and $T/n \longrightarrow 0$.

The estimator for the drift is:

$$\hat{\mu}_n(x) = \frac{n \sum_{i=1}^{n} K\left(\frac{X_{t_{i-1}} - \hat{J}_{1,t_{i-1}} - x}{h}\right) (X_{t_{i+1}} - X_{t_i}) \cdot I_{\{(X_{t_{i+1}} - X_{t_i})^2 \leq \vartheta(\frac{T}{n})\}}}{T \sum_{i=1}^{n} K\left(\frac{X_{t_{i-1}} - \hat{J}_{1,t_{i-1}} - x}{h}\right)}$$

Then for each x visited by X we have

$$\hat{\mu}_n(x) \longrightarrow_P \mu(x).$$

Estimation of the intensity function

The estimator for $\lambda(x)$ is:

$$\hat{\lambda}_n(x) = \frac{n \sum_{i=1}^{n} K\left(\frac{X_{t_{i-1}} - x}{h}\right) c_{i,n} I_{\{(X_{t_{i+1}} - X_{t_i})^2 > \vartheta(\delta)\}}}{T \sum_{i=1}^{n} K\left(\frac{X_{t_{i-1}} - x}{h}\right)}$$

where $\sup_i |1 - c_{i,n}| \to 0$ when $n \to \infty$.

Then, for each x visited by X,

$$\hat{\lambda}_n(x) \to_P \lambda(x).$$

A small sample correction

The coefficients c_i can help in recovering an unbiased estimated intensity by making assumptions on the distribution of jumps.

A small sample correction

The coefficients c_i can help in recovering an unbiased estimated intensity by making assumptions on the distribution of jumps.

Under normality, we choose:

$$c_{i,n} = \frac{1}{2\mathcal{N}(-\sqrt{\vartheta}/\sigma_J)}$$

A small sample correction

The coefficients c_i can help in recovering an unbiased estimated intensity by making assumptions on the distribution of jumps.

Under normality, we choose:

$$c_{i,n} = \frac{1}{2\mathcal{N}(-\sqrt{\vartheta}/\sigma_J)}$$

We can estimate σ_J using, for example, a simple method of moments estimator, fitting:

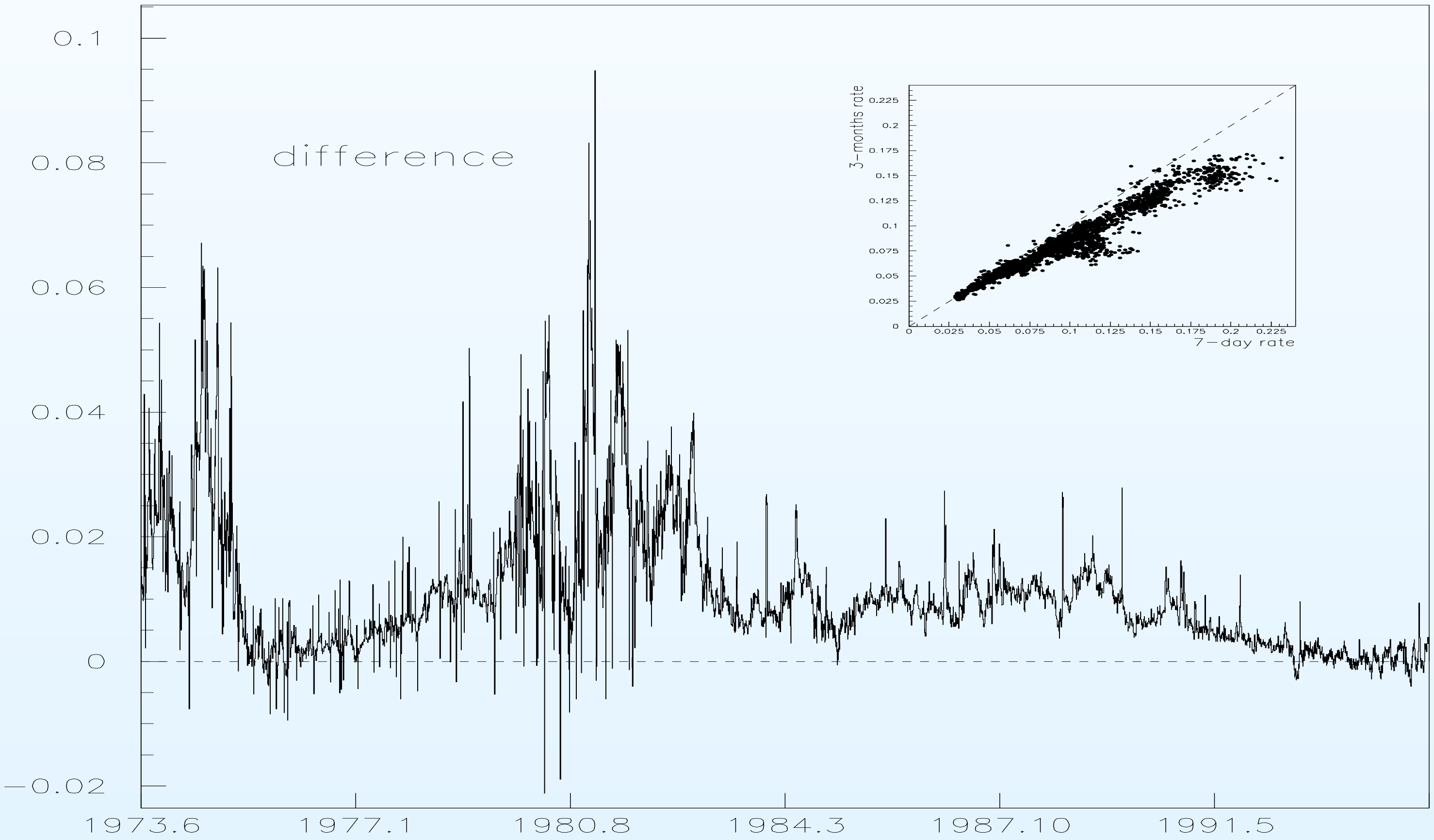
$$m_2(\vartheta) = \sigma_J^2 + \frac{\sigma_J\sqrt{\vartheta}\exp\left(\frac{-\vartheta}{2\sigma_J^2}\right)}{\mathcal{N}(-\sqrt{\vartheta}/\sigma_J)\sqrt{2\pi}} \tag{16}$$

Short rate modelling

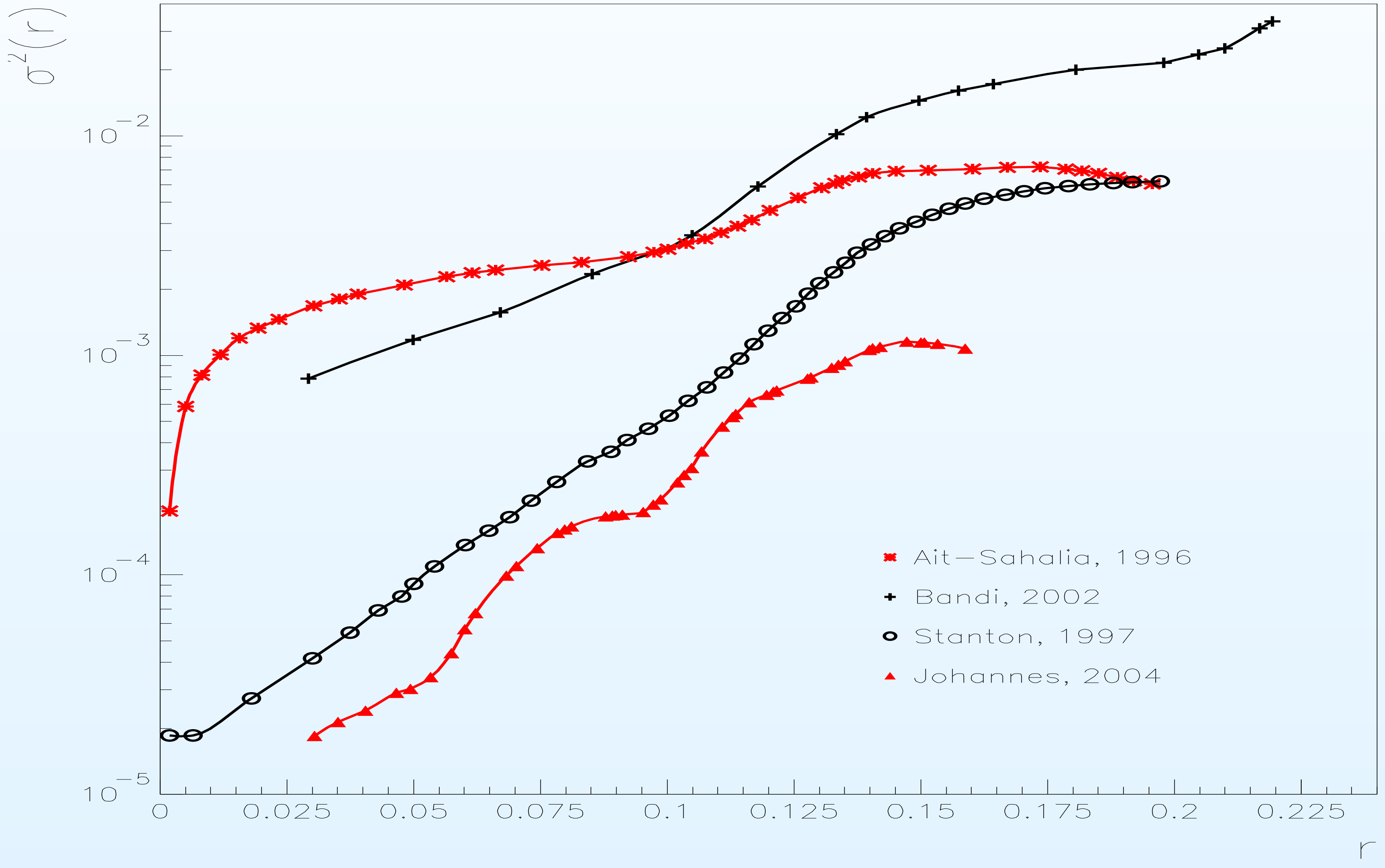
Which is the short rate?

Short rate modelling

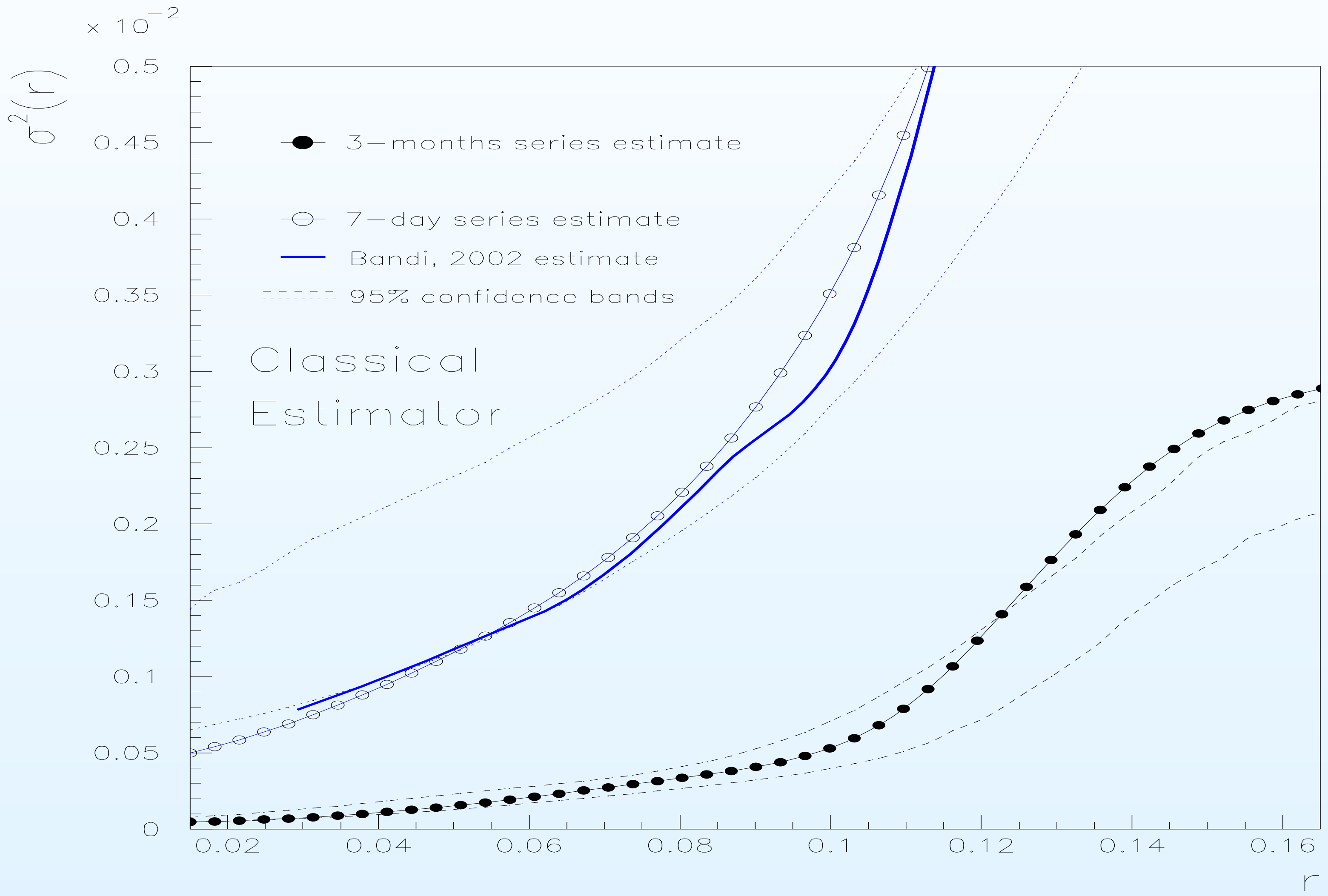
Which is the difference?



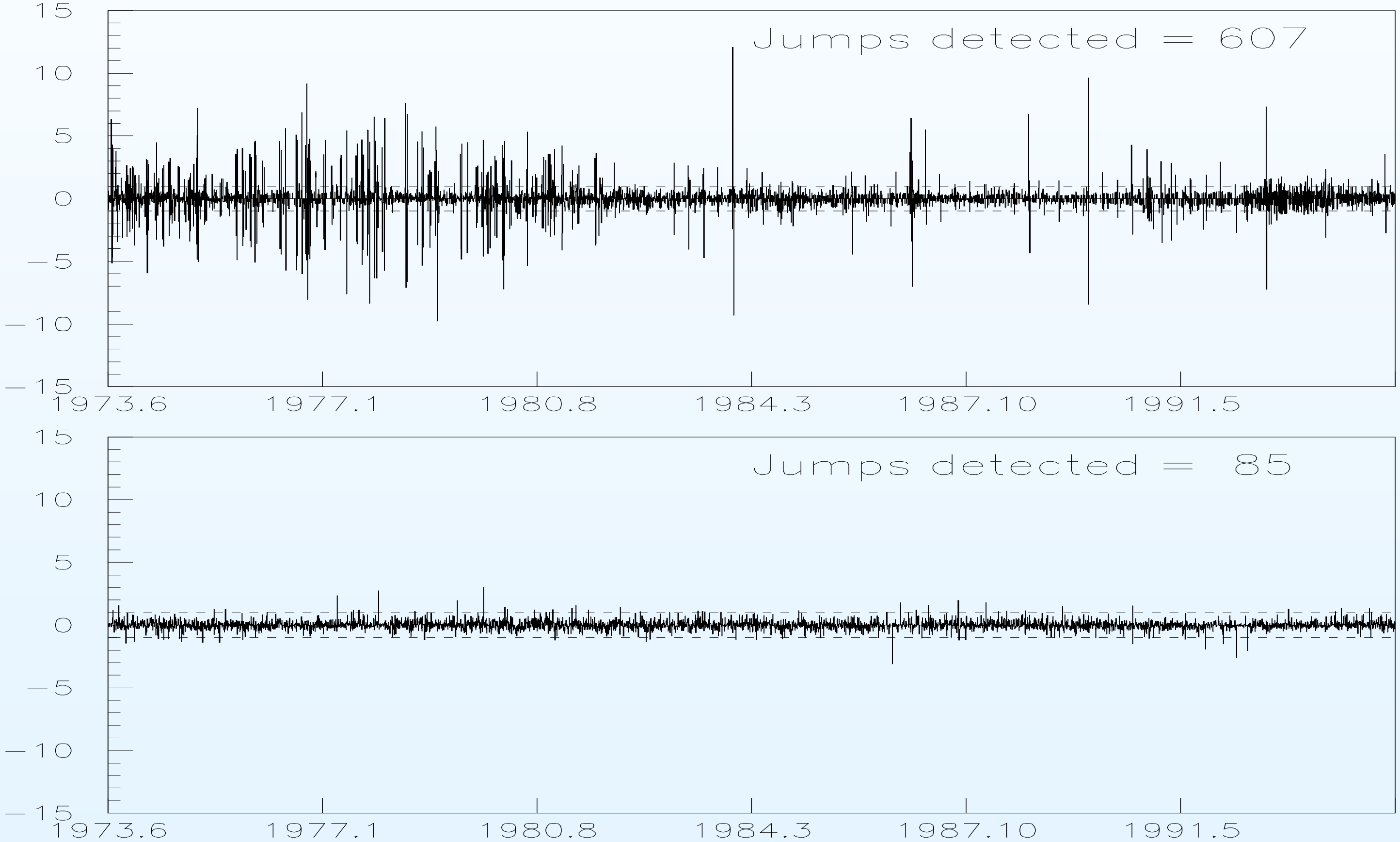
The state of the art



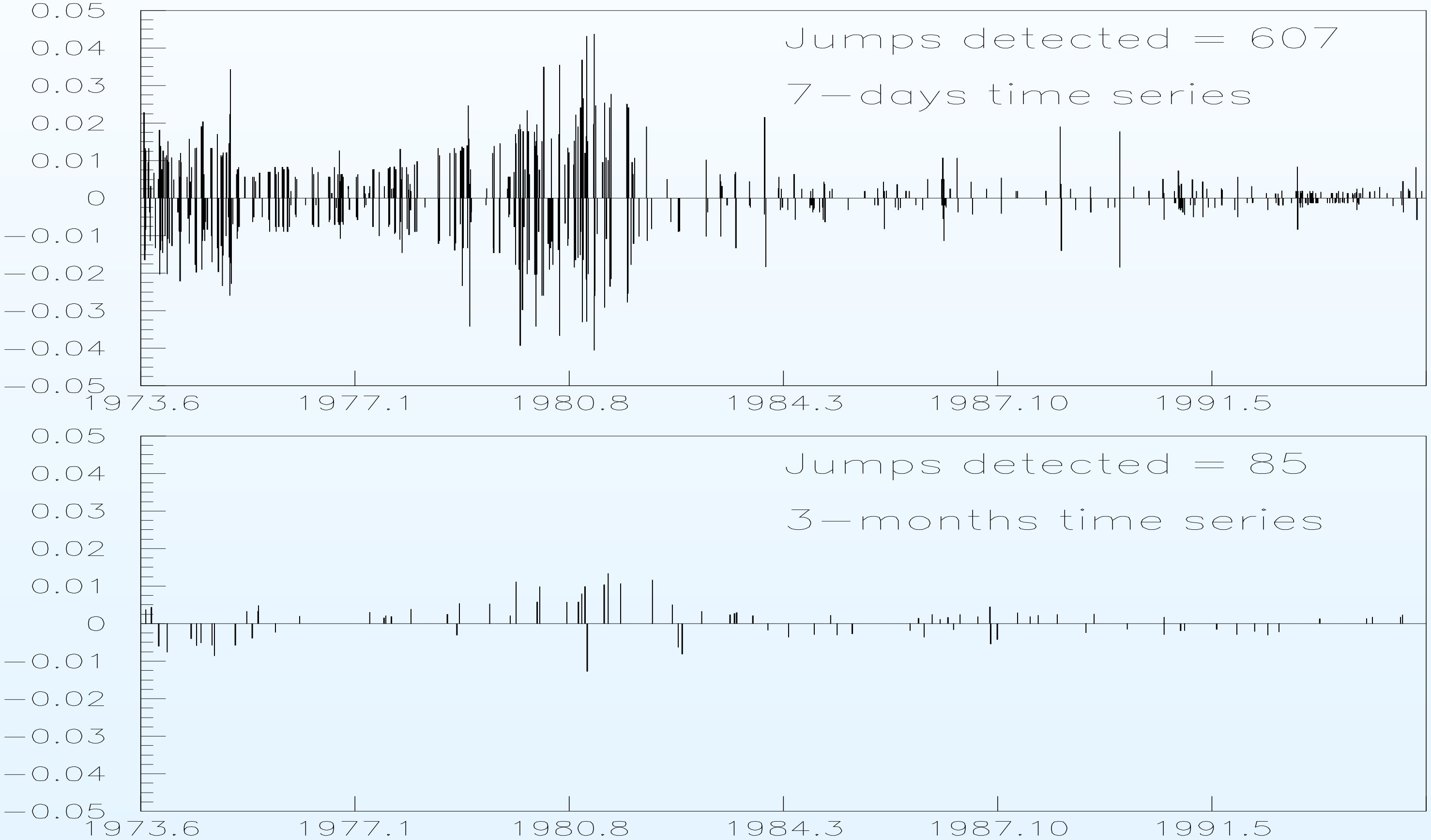
Estimation with no jumps



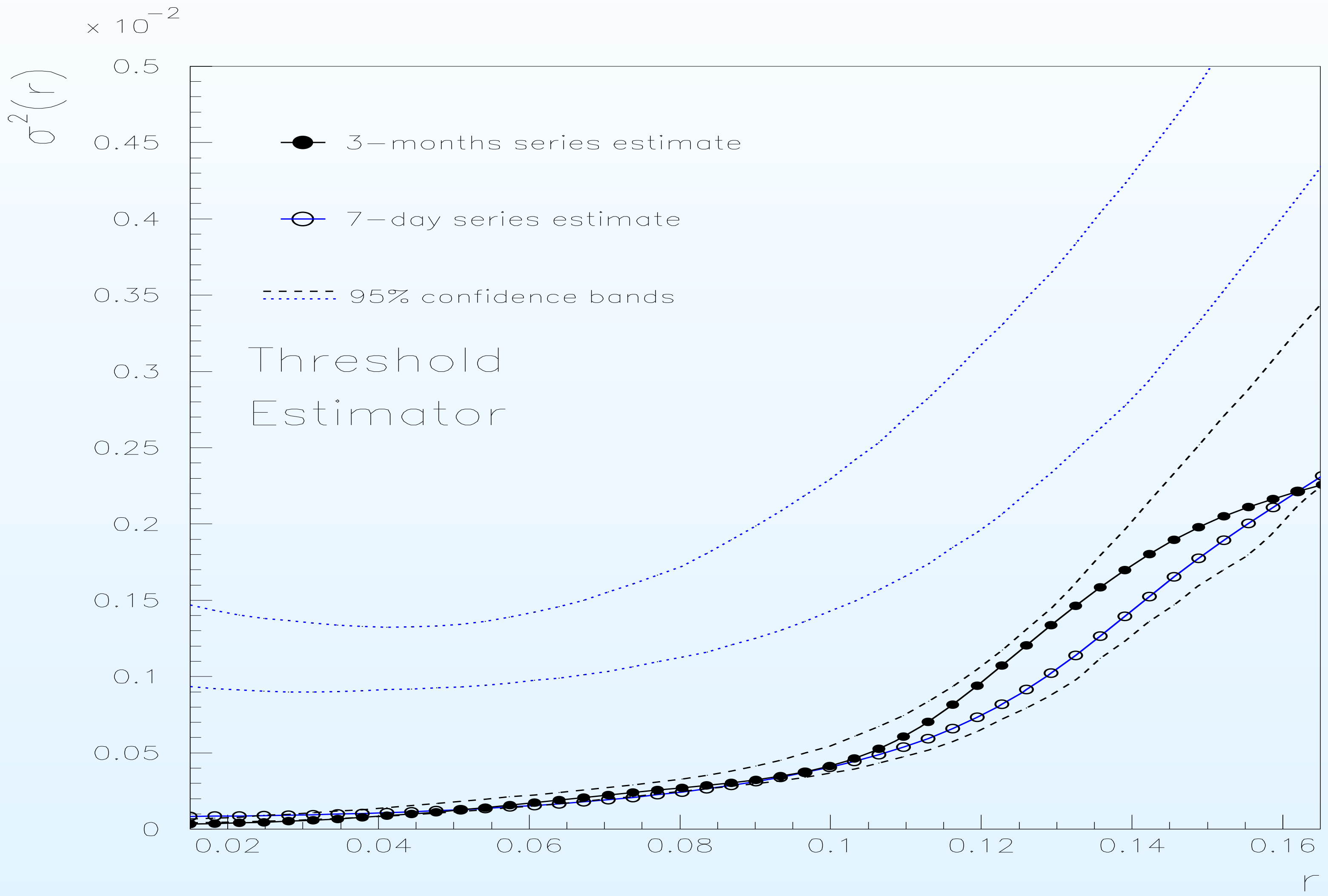
Jumps detected



Jumps detected



Estimation with the threshold estimator



Comparison with Bandi-Nguyen-Johannes

- An alternative estimation (under finite activity jumps) is developed by Bandi and Nguyen (2003) and Johannes (2004).

Comparison with Bandi-Nguyen-Johannes

- An alternative estimation (under finite activity jumps) is developed by Bandi and Nguyen (2003) and Johannes (2004).
- Using the fourth and the sixth moment, and assuming normality for C, with $\mathbb{E}[C] = 0$ and denoting by $\sigma_J^2 = \mathbb{E}[C^2]$, we have:

$$\lim_{\delta \to 0} \frac{1}{\delta} \mathbb{E}_t \left[(X_{t+\delta} - X_t)^4 \right] = 3\lambda(X_t)\sigma_J^4$$

$$\lim_{\delta \to 0} \frac{1}{\delta} \mathbb{E}_t \left[(X_{t+\delta} - X_t)^6 \right] = 15\lambda(X_t)\sigma_J^6,$$

Comparison with Bandi-Nguyen-Johannes

- An alternative estimation (under finite activity jumps) is developed by Bandi and Nguyen (2003) and Johannes (2004).
- Using the fourth and the sixth moment, and assuming normality for C, with $\mathbb{E}[C] = 0$ and denoting by $\sigma_J^2 = \mathbb{E}[C^2]$, we have:

$$\lim_{\delta \to 0} \frac{1}{\delta} \mathbb{E}_t\left[(X_{t+\delta} - X_t)^4\right] = 3\lambda(X_t)\sigma_J^4$$

$$\lim_{\delta \to 0} \frac{1}{\delta} \mathbb{E}_t\left[(X_{t+\delta} - X_t)^6\right] = 15\lambda(X_t)\sigma_J^6,$$

- After estimating $\lambda(X_t), \sigma_J$ we have:

$$\lim_{\delta \to 0} \frac{1}{\delta} \mathbb{E}_t\left[(X_{t+\delta} - X_t)^2\right] = \sigma(X_t) + \lambda(X_t)\sigma_J^2$$

Comparison with Bandi-Nguyen-Johannes

- Moment estimators can be obtained using:

$$\frac{1}{\delta}\mathbb{E}_t\left[\left(X_{t+\delta}-X_t\right)^k\right]=\frac{n\sum_{i=1}^{n-1}K\left(\frac{X_i-x}{h_n}\right)\left(X_{i+1}-X_i\right)^k}{T\sum_{i=1}^{n-1}K\left(\frac{X_i-x}{h_n}\right)}$$

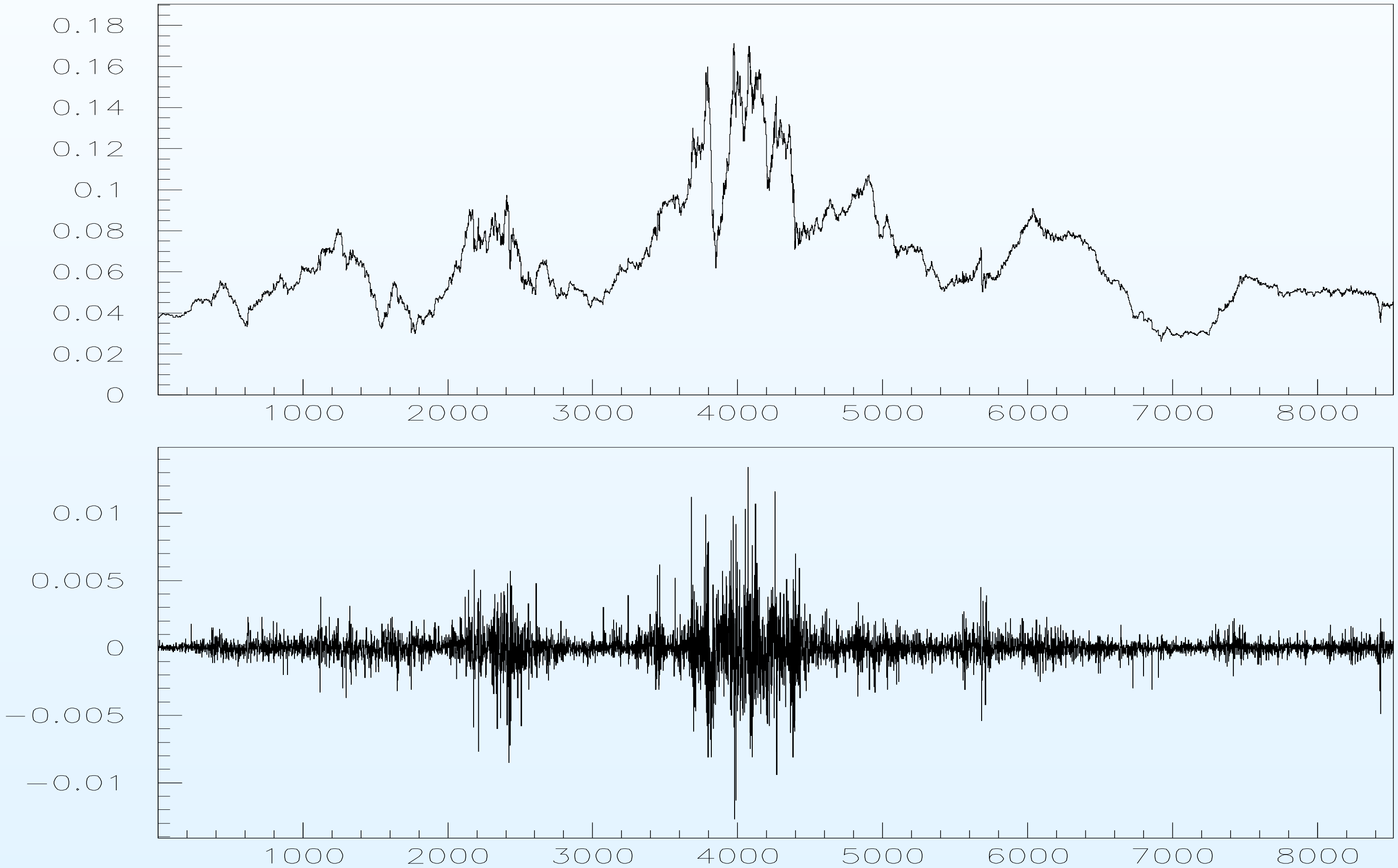
Comparison with Bandi-Nguyen-Johannes

- Moment estimators can be obtained using:

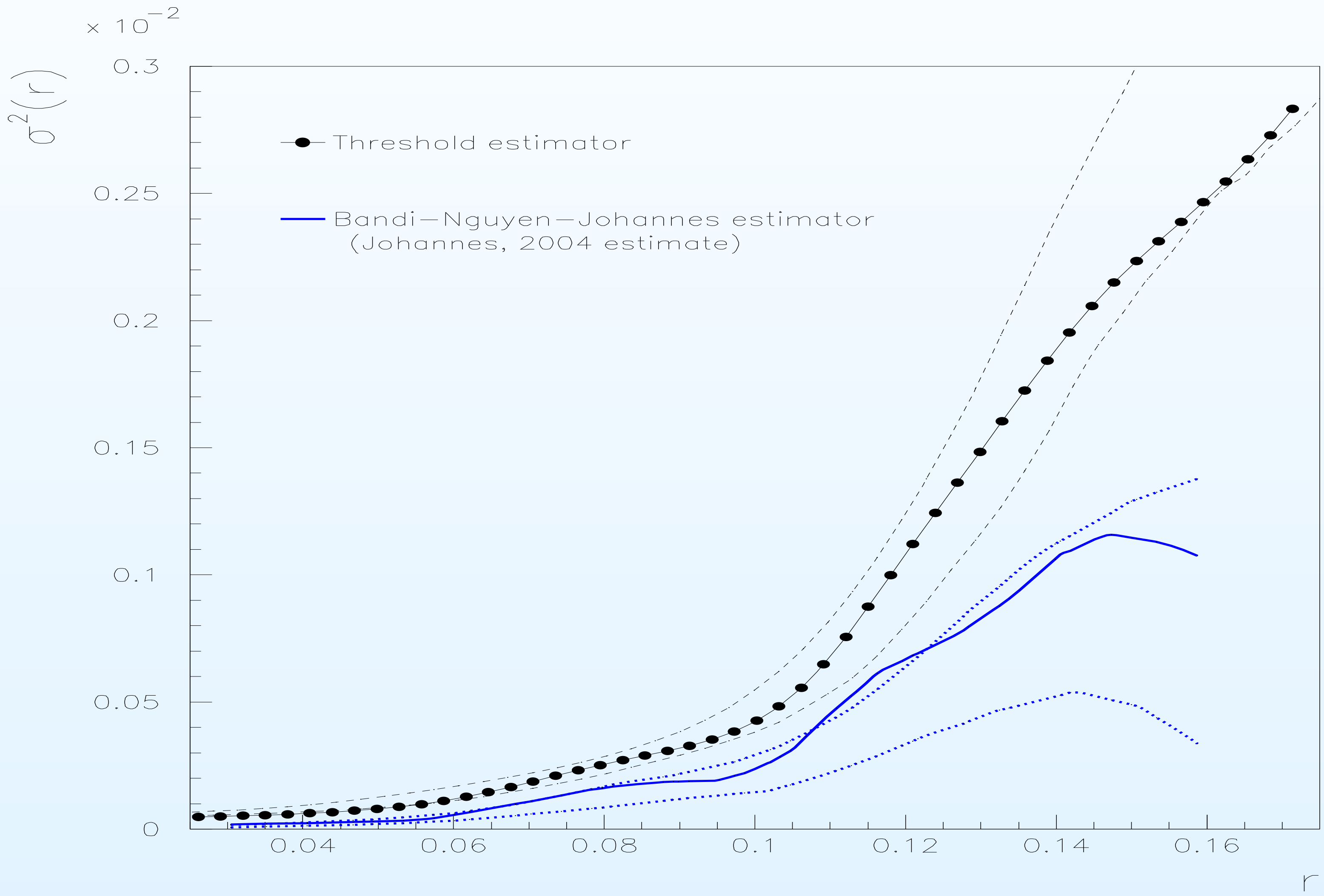
$$\frac{1}{\delta}\mathbb{E}_t\left[\left(X_{t+\delta}-X_t\right)^k\right]=\frac{n\sum_{i=1}^{n-1}K\left(\frac{X_i-x}{h_n}\right)\left(X_{i+1}-X_i\right)^k}{T\sum_{i=1}^{n-1}K\left(\frac{X_i-x}{h_n}\right)}$$

- Full limit theorems are developed in Bandi and Nguyen (2003), estimation on interest rate data is accomplished in Johannes (2004).

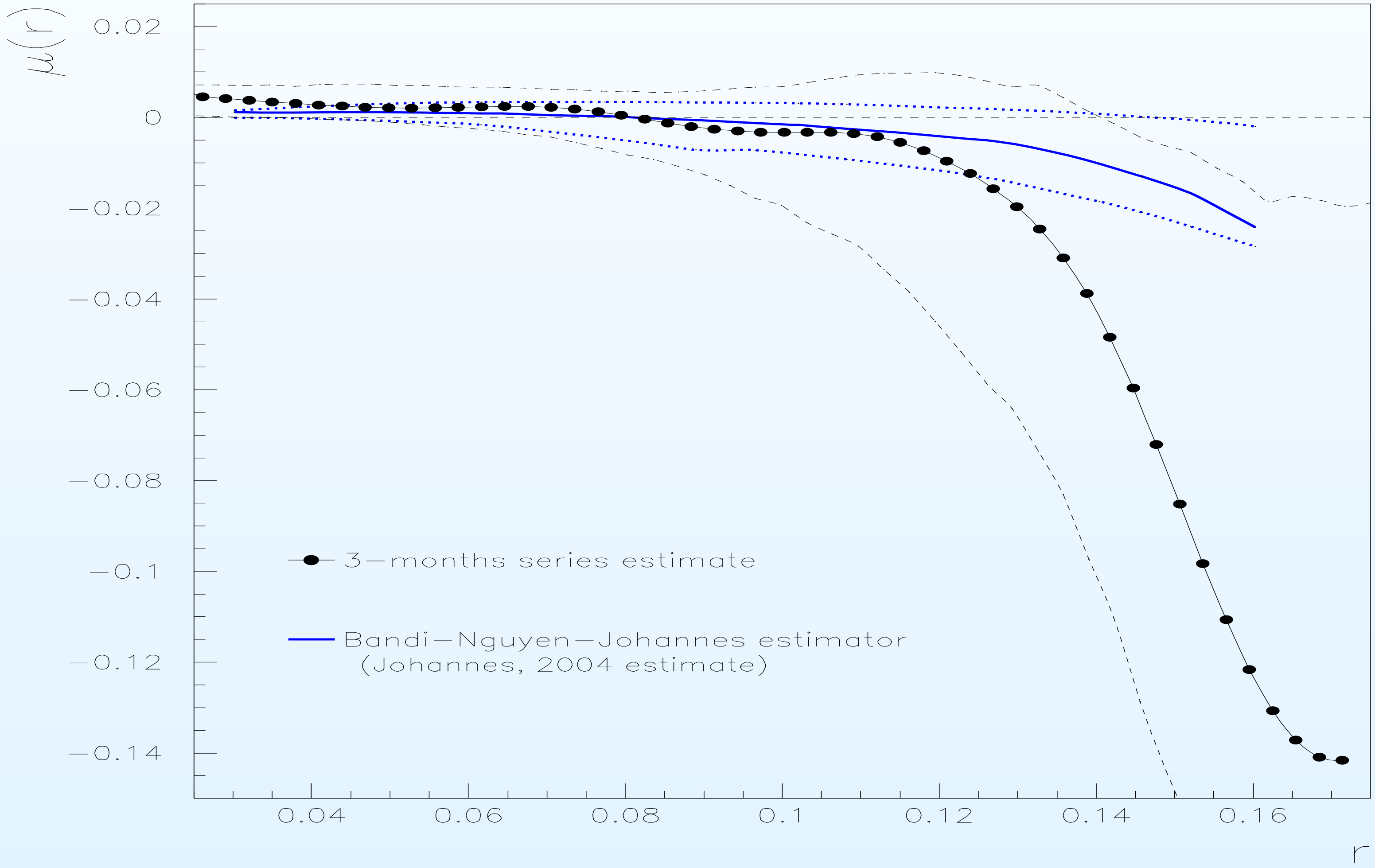
Johannes (2004) dataset



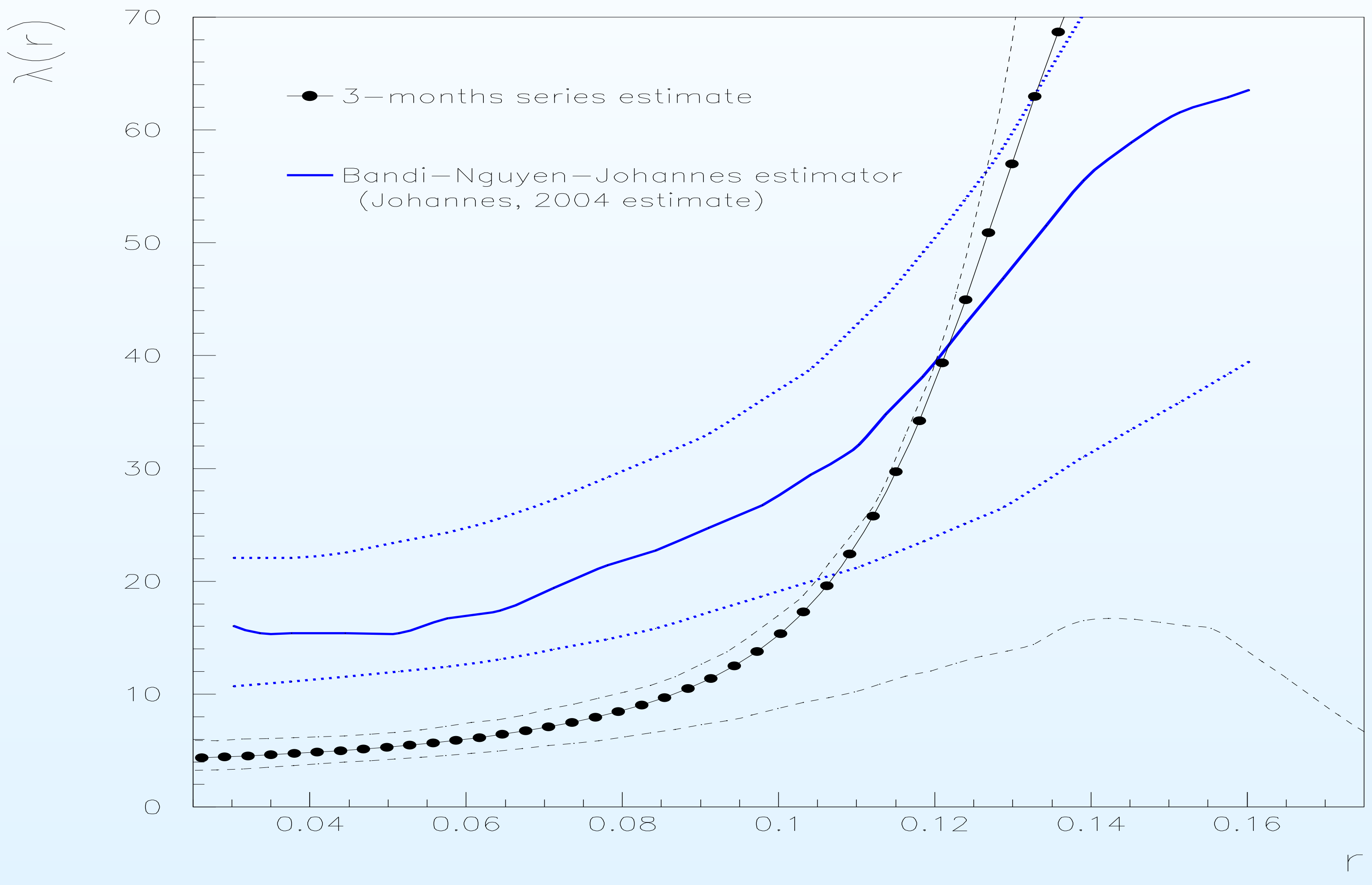
Comparison with Bandi-Nguyen-Johannes



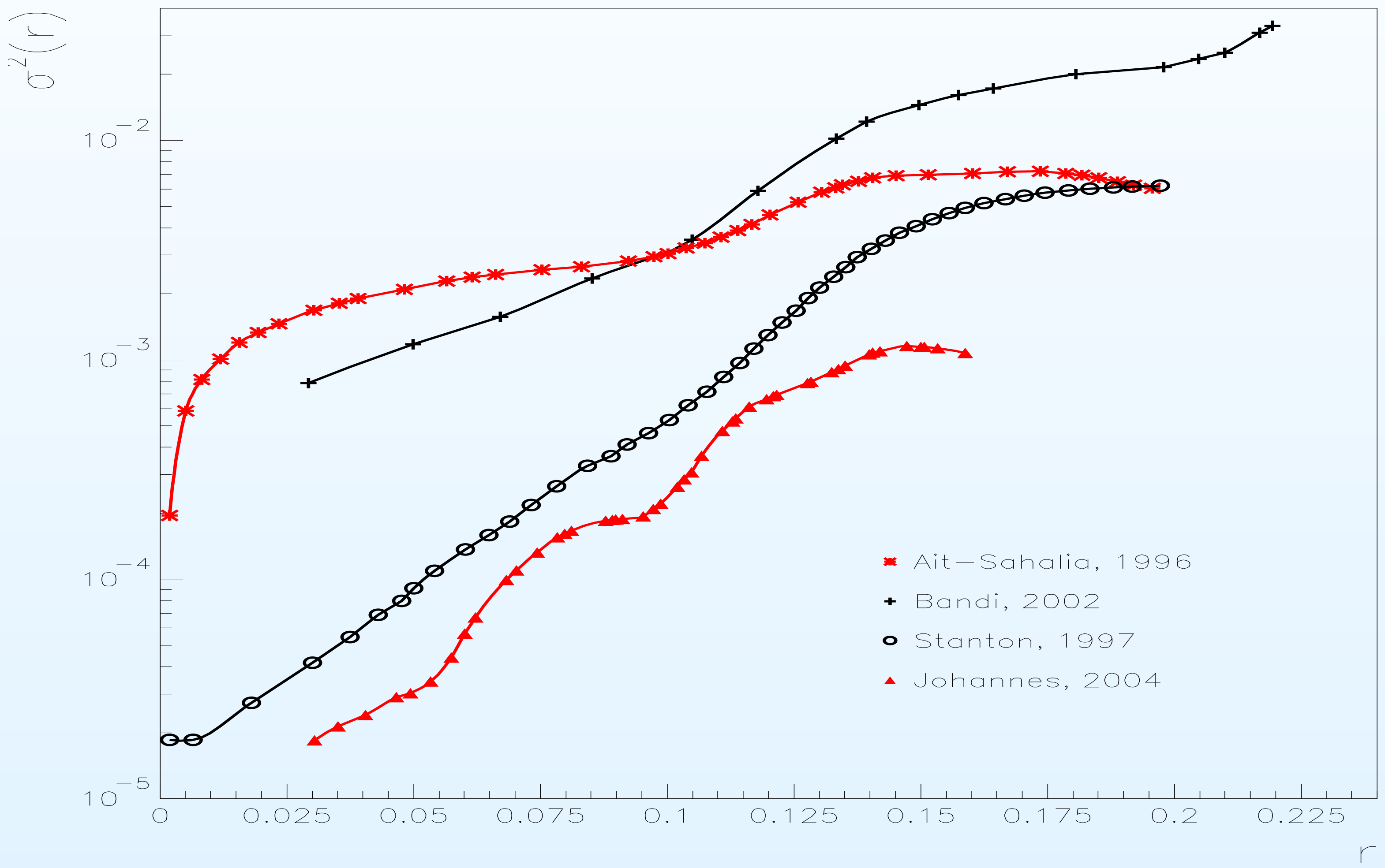
Comparison with Bandi-Nguyen-Johannes



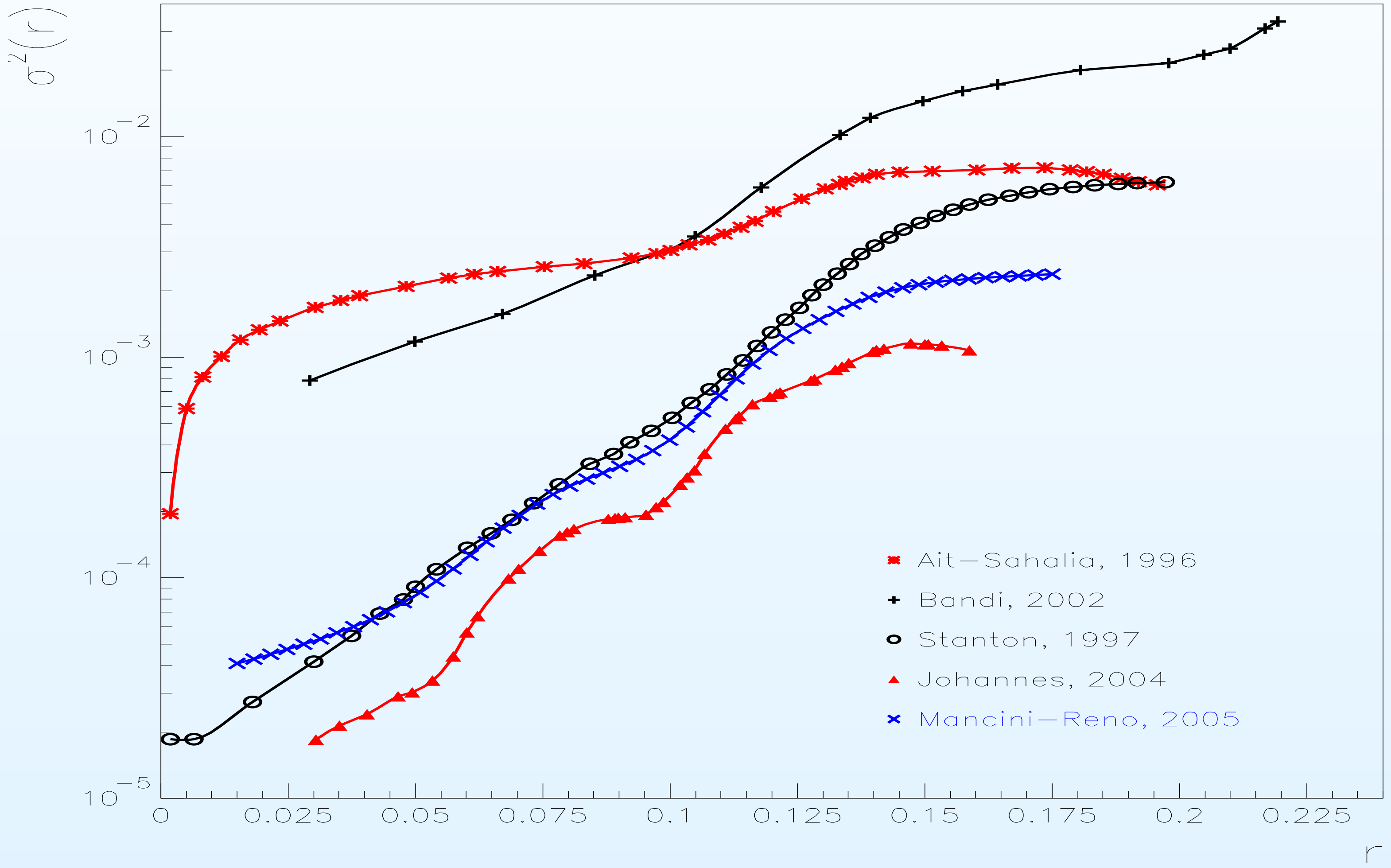
Comparison with Bandi-Nguyen-Johannes



Comparison with the literature



Comparison with the literature



Conclusions

- We propose a nonparametric estimator for univariate jump-diffusion models based on threshold estimation.

Conclusions

- We propose a nonparametric estimator for univariate jump-diffusion models based on threshold estimation.
- We show that the threshold can be set using standard econometric techniques (GARCH modelling)

Conclusions

- We propose a nonparametric estimator for univariate jump-diffusion models based on threshold estimation.
- We show that the threshold can be set using standard econometric techniques (GARCH modelling)
- Monte Carlo simulations show that the proposed estimator of the diffusion coefficient is superior to alternatives which do not consider jumps

Conclusions

- We propose a nonparametric estimator for univariate jump-diffusion models based on threshold estimation.
- We show that the threshold can be set using standard econometric techniques (GARCH modelling)
- Monte Carlo simulations show that the proposed estimator of the diffusion coefficient is superior to alternatives which do not consider jumps
- We show that the 7-day time series is characterized by an inherent jump process which is absent on the 3-months time series. This is due to liquidity reasons.

Conclusions

- We propose a nonparametric estimator for univariate jump-diffusion models based on threshold estimation.
- We show that the threshold can be set using standard econometric techniques (GARCH modelling)
- Monte Carlo simulations show that the proposed estimator of the diffusion coefficient is superior to alternatives which do not consider jumps
- We show that the 7-day time series is characterized by an inherent jump process which is absent on the 3-months time series. This is due to liquidity reasons.
- Our estimators provide different results than those obtained so far on jump-diffusion models: more diffusive variance and less jumps.

Work in progress

1. Explain the difference with Bandi-Nguyen-Johannes estimator

Work in progress

1. Explain the difference with Bandi-Nguyen-Johannes estimator
2. Use threshold estimators to estimate $\int \sigma_s^\gamma ds$

Work in progress

1. Explain the difference with Bandi-Nguyen-Johannes estimator
2. Use threshold estimators to estimate $\int \sigma_s^\gamma ds$
   SV1FJ model - One day one jump (Pirino and Renò, 2007)

	Mean(%)	Standard Dev.(%)
BPV	27.72	37.77
TBPV	-7.91	18.75
C-TBPV	3.85	18.92
QV	53.71	113.41
TQV	-13.36	40.41
C-TQV	7.72	46.17
TriV	105.62	243.84
TTriV	-12.71	38.73
C-TTriV	10.36	46.94

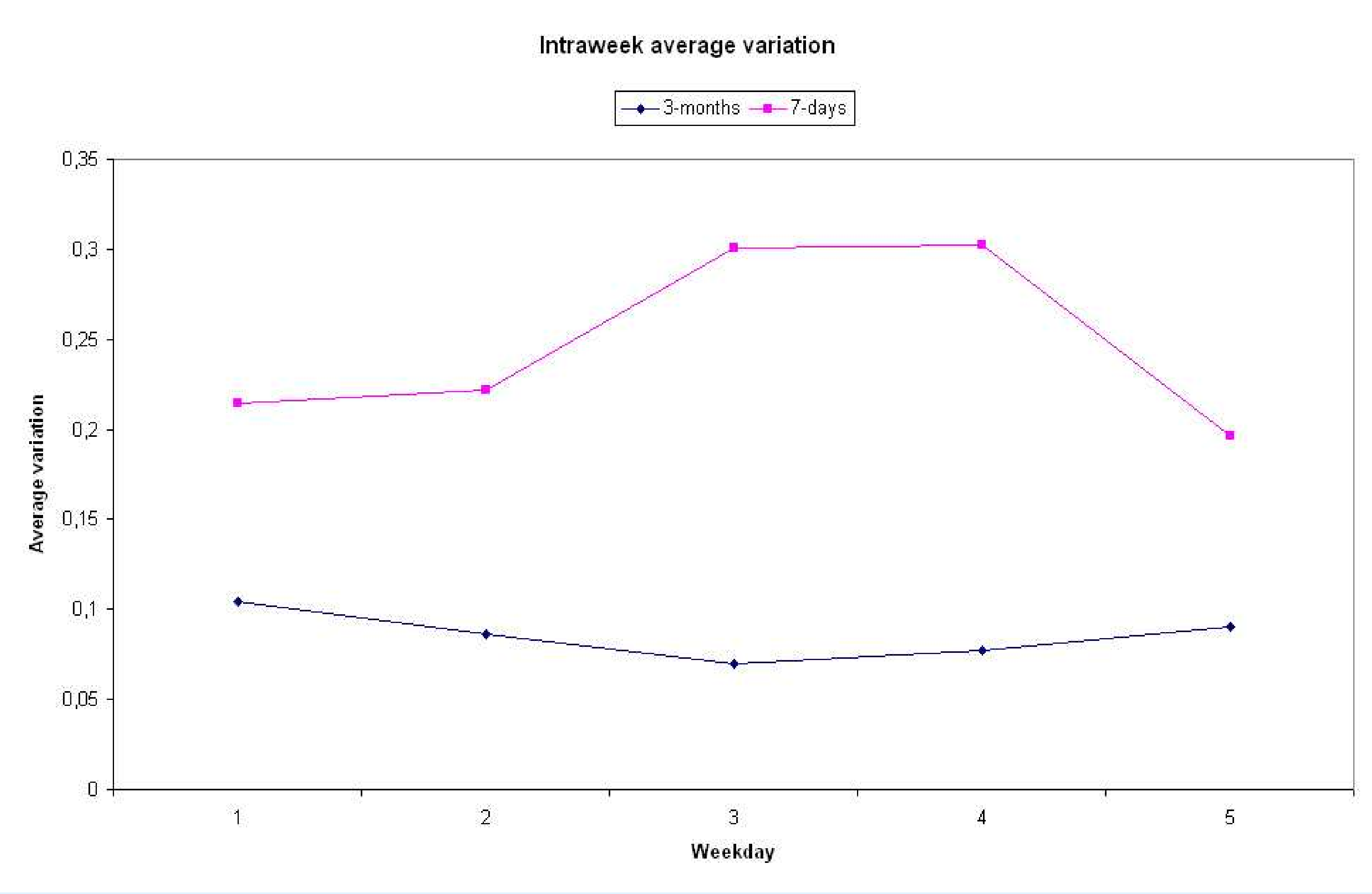
Work in progress

1. Explain the difference with Bandi-Nguyen-Johannes estimator
2. Use threshold estimators to estimate $\int \sigma_s^\gamma ds$
3. Apply above results to volatility forecasting, extending the HAR-RV model of Corsi (2004).

Extra-figures

Extra-figures

Extra-figures



Extra-figures