



Modelling ultra high frequency financial data: Model with uncertainty zones and associated statistical procedures for volatility, correlation and lead-lag

Mathieu Rosenbaum

Ecole Polytechnique

Fall 2020



- 1 Modelling ultra high frequency data
- 2 Model with uncertainty zones
- 3 Volatility estimation
- 4 Covariation estimation
- 5 Lead-Lag estimation



- 1 Modelling ultra high frequency data
- 2 Model with uncertainty zones
- 3 Volatility estimation
- 4 Covariation estimation
- 5 Lead-Lag estimation



Properties of the microstructure model

We want the following properties for our model :

- Model for prices **and** durations → no hesitation about the sampling frequency.
- Discrete prices.
- Bid-Ask bounce.
- Stylized facts of returns, durations and volatility. In particular, inverse relation between durations and volatility.
- A diffusive behavior at large sampling scales.



Properties of the microstructure model (2)

- No hesitation about the price.
- Finite quadratic variation for the microstructure noise.
- An interpretation of the model.
- **A useful model.**

A proposed answer :

- The model with uncertainty zones.
- In this model, prices and durations are functionals of some hitting times of an underlying continuous semi-martingale.



- 1 Modelling ultra high frequency data
- 2 Model with uncertainty zones
- 3 Volatility estimation
- 4 Covariation estimation
- 5 Lead-Lag estimation



Aversion for price changes

- In an idealistic framework, transactions would occur when the efficient price crosses the tick grid.
- In practice, uncertainty about the efficient price and aversion for price changes of market participants.
- The price changes only when market participants are convinced that the efficient price is far from the last traded price.
- We introduce a parameter η quantifying this aversion for price changes.



Model with uncertainty zones : notation

- Efficient price : X_t .
- α : tick size.
- t_i : time of the i -th transaction with price change.
- P_{t_i} : transaction price at time t_i .
- $L_i = |P_{t_{i+1}} - P_{t_i}|/\alpha$: size of the i -th price jump.
- Explanatory variables process for the size of the jumps : χ_t .
- Uncertainty zones : $U_k = [0, \infty) \times (d_k, u_k)$ with

$$d_k = (k + 1/2 - \eta)\alpha \text{ and } u_k = (k + 1/2 + \eta)\alpha.$$

- τ_i : i -th exit time of an uncertainty zone.

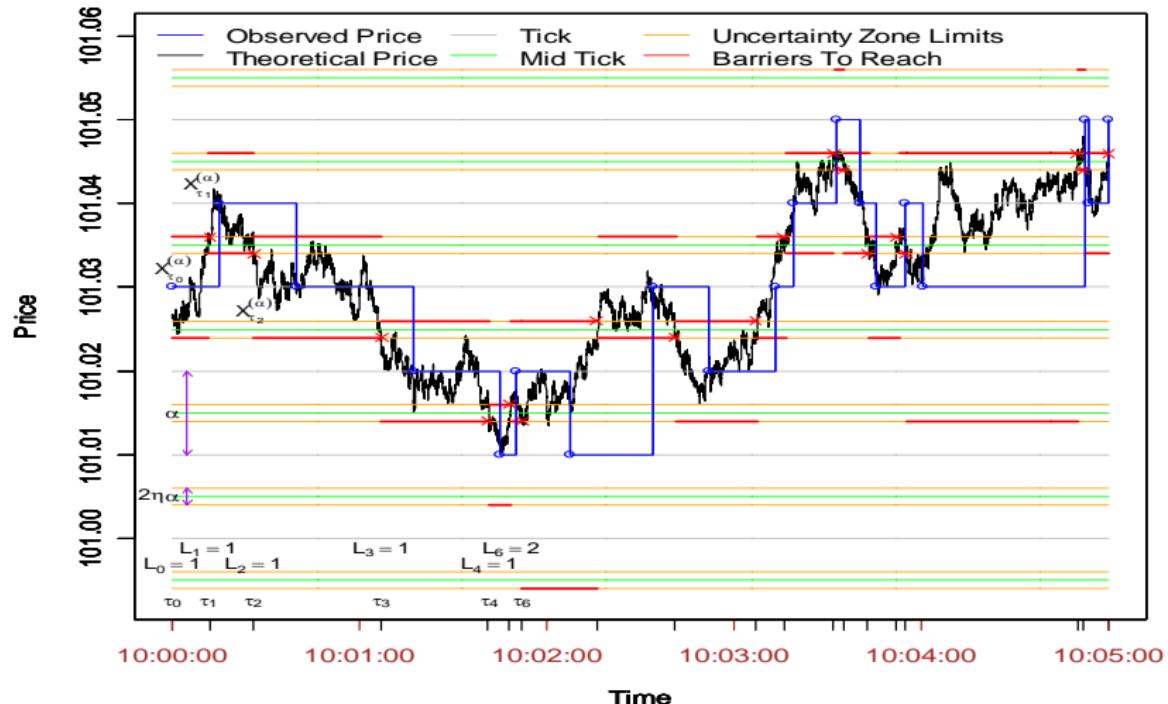
Model with uncertainty zones : dynamics

- $d \log X_u = a_u du + \sigma_u dW_u.$
- $\mathbb{P}[L_i = s | \mathcal{F}_{\tau_i}] = \phi_s(\chi_{\tau_i}).$
- $\tau_{i+1} = \inf \left\{ t > \tau_i, X_t = X_{\tau_i}^{(\alpha)} \pm \alpha(L_i - \frac{1}{2} + \eta) \right\},$
with $X_{\tau_i}^{(\alpha)}$ the value of X_{τ_i} rounded to the nearest multiple of α .
- $\tau_i \leq t_i < \tau_{i+1}$ and $P_{t_i} = X_{\tau_i}^{(\alpha)}.$

Model with uncertainty zones

© Théo Jalabert







Comments on the model

- Price/Durations model for the last traded price.
- Latent semi-martingale efficient price.
- Random times for the transactions. Inverse relation between durations and volatility.
- Bid-Ask bounce.
- Stylized facts of auto and cross correlogramms reproduced, both in tick time and calendar time.
- Jumps of several ticks. Size of the jumps determined by explanatory variables, involving for example the order book.
- Structural model for the microstructure noise.



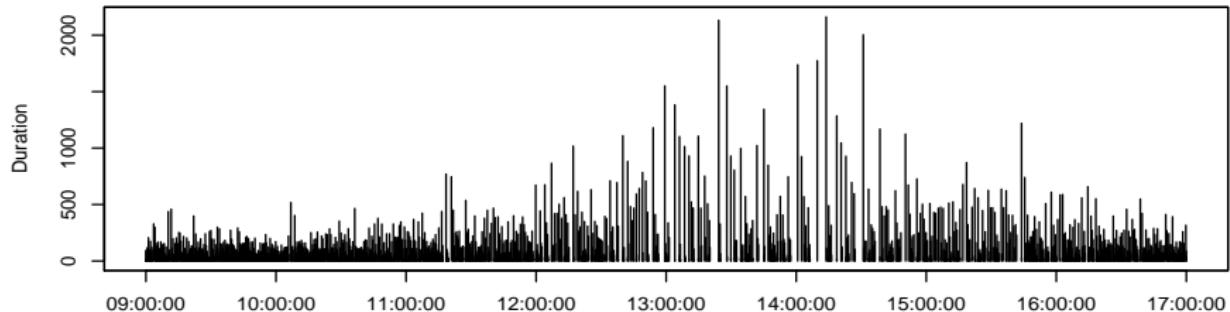
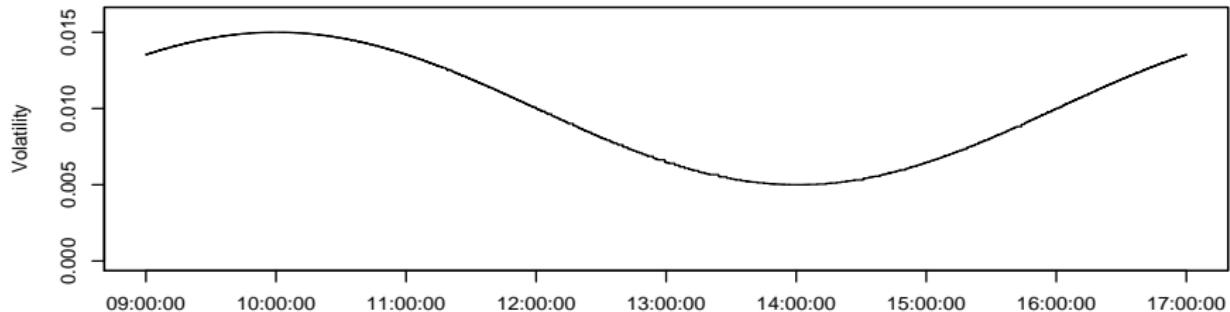
Interpretation of η

- Quantifies the aversion for price changes (with respect to the tick size) of market participants.
- In the UHF, the order book can not “follow” the efficient price and is reluctant to price changes. Reluctancy measured by η .
- $2\eta\alpha$ represents the **implicit spread** of a large tick asset (see later).
- A small $\eta (< 1/2)$ means that for market participants, the tick size is too big and conversely.

Some properties : Durations

© Théo Jalabert

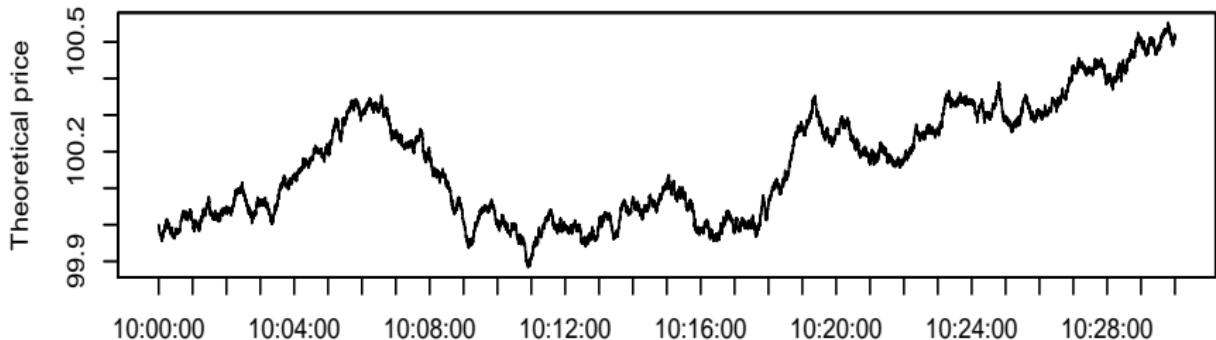
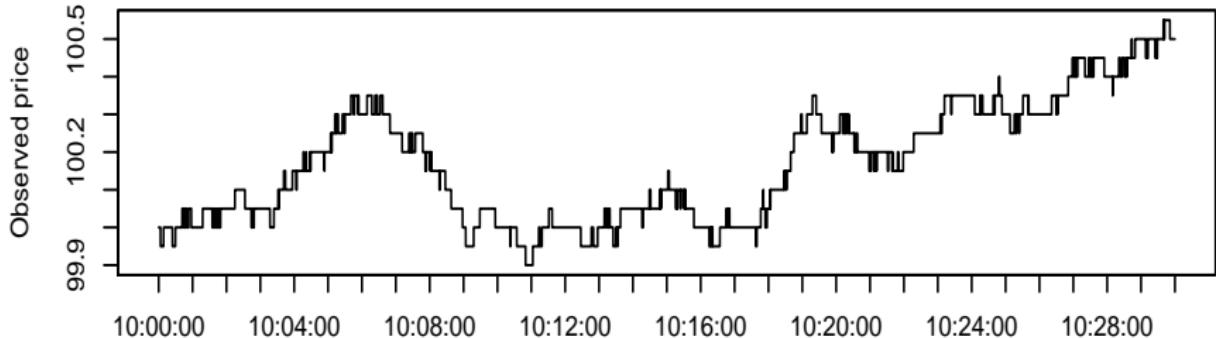




Some properties : The price

© Théo Jalabert

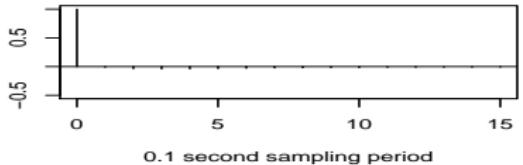




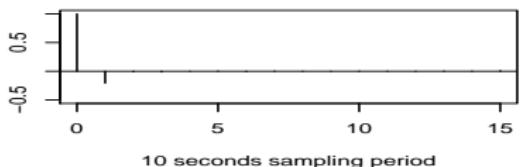
Some properties : ACF log returns ($n < 1/2$)

© Théo Jérôme

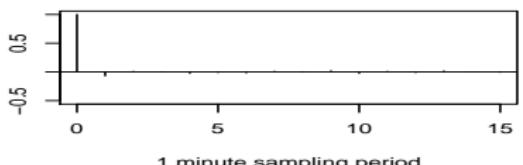
Théo Jérôme



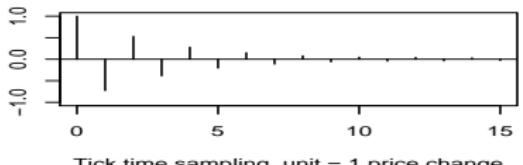
0.1 second sampling period



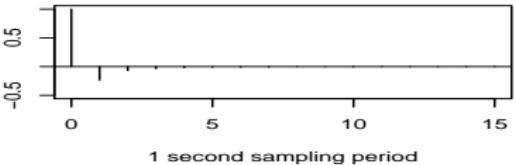
10 seconds sampling period



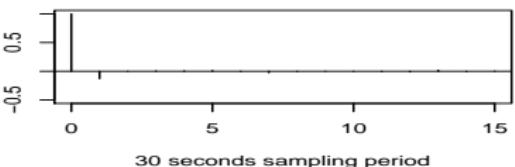
1 minute sampling period



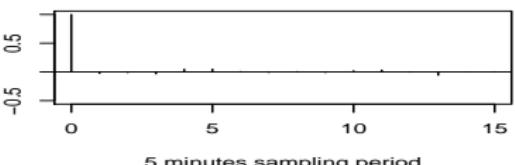
Tick time sampling, unit = 1 price change



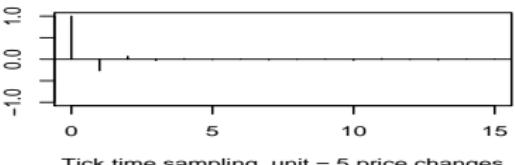
1 second sampling period



30 seconds sampling period



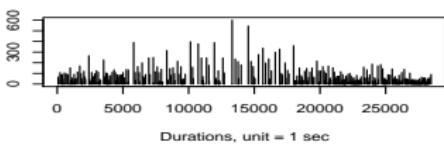
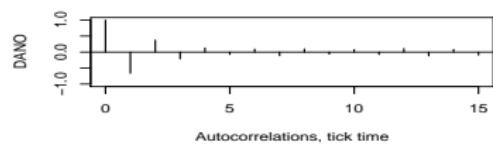
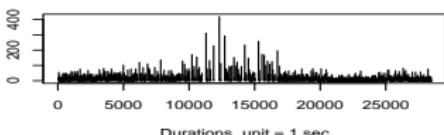
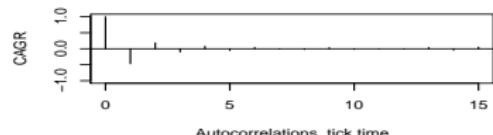
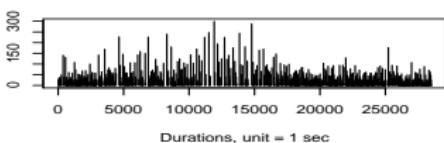
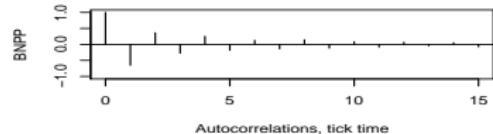
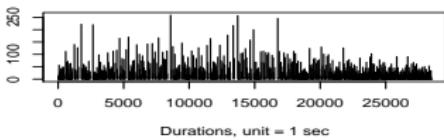
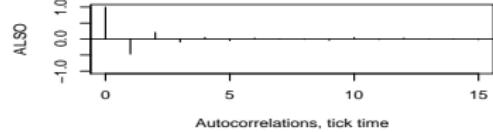
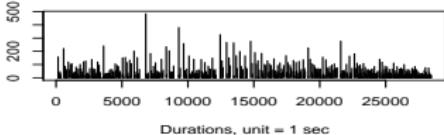
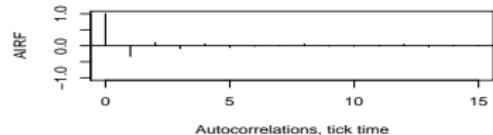
5 minutes sampling period



Tick time sampling, unit = 5 price changes

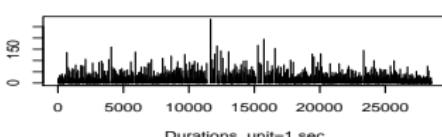
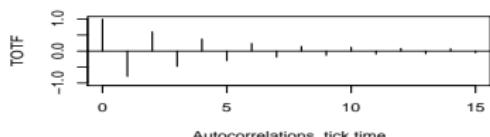
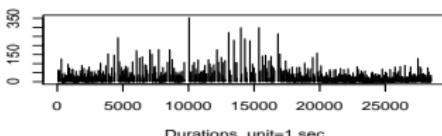
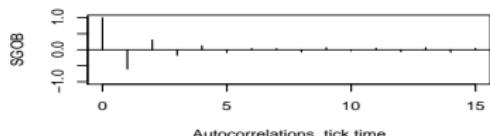
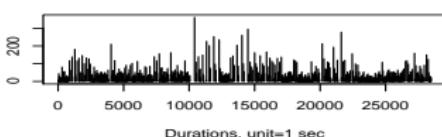
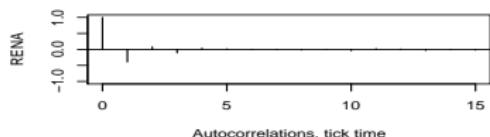
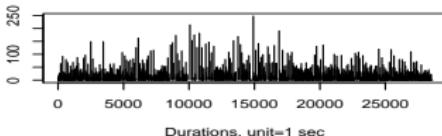
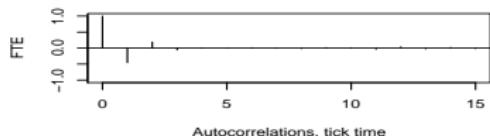
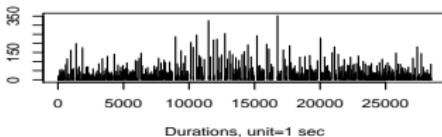
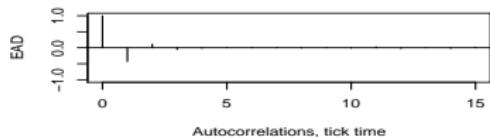
Real data : 2007/01/16

© Théo Jalabert



Real data : 2007/01/16

© Théo Jalabert





- 1 Modelling ultra high frequency data
- 2 Model with uncertainty zones
- 3 Volatility estimation
- 4 Covariation estimation
- 5 Lead-Lag estimation



Efficient price

- $X_{\tau_i} = P_{t_i} - \alpha\left(\frac{1}{2} - \eta\right)sign(P_{t_i} - P_{t_{i-1}}).$
- $\hat{X}_{\tau_i} = P_{t_i} - \alpha\left(\frac{1}{2} - \hat{\eta}\right)sign(P_{t_i} - P_{t_{i-1}}).$

t_i : observation time, τ_i : exit time,
 P_t : observed price, X_t : efficient price.

Estimation of η

© Théo Jalabert



Let $N_{\alpha,t} = \text{card}\{t_i, t_i \leq t\}$. For $k = 1, \dots, m$,

$$N_{\alpha,t,k}^{(c)} = \sum_{i=1}^{N_{\alpha,t}} \mathbb{I}_{\{|X_{\tau_i} - X_{\tau_{i-1}}| = \alpha k\}},$$

$$N_{\alpha,t,k}^{(a)} = \sum_{i=1}^{N_{\alpha,t}} \mathbb{I}_{\{|X_{\tau_i} - X_{\tau_{i-1}}| = \alpha(k-1+2\eta)\}}.$$

We define

$$\hat{\eta}_t = \sum_{k=1}^m \lambda_{\alpha,t,k} u_{\alpha,t,k},$$

with

$$\lambda_{\alpha,t,k} = \frac{N_{\alpha,t,k}^{(a)} + N_{\alpha,t,k}^{(c)}}{\sum_{j=1}^m [N_{\alpha,t,j}^{(a)} + N_{\alpha,t,j}^{(c)}]} \text{ and } u_{\alpha,t,k} = \frac{1}{2} \left(k \left(\frac{N_{\alpha,t,k}^{(c)}}{N_{\alpha,t,k}^{(a)}} - 1 \right) + 1 \right).$$



Theorem

Let

$$\widehat{RV}_t = \sum_{i=1}^{N_{\alpha,t}} (\log(\hat{X}_{\tau_i}) - \log(\hat{X}_{\tau_{i-1}}))^2.$$

We have

$$\alpha^{-1}(\widehat{RV}_t - RV_t) \xrightarrow{\mathcal{L}_s} \gamma_t \int_0^t v_u dW_{\theta_u},$$

where W is a Brownian motion independent of B and θ_u , γ_u and v_u depend on X_u , σ_u and explanatory variables, involving for example the order book.



Comments on the theorem

- a_u only needs to be progressively measurable and locally bounded. σ_u only needs to be adapted càdlàg. In particular, σ_u is not necessarily an Ito semi-martingale.
- Rounding happens on the original scale. Thus, estimating the quadratic variation of the log price is more intricate than estimating those of the price.
- The observation times are random, endogenous. So, usual theorems for deterministic or exogenous sampling can not be applied.
- The key ideas are to work in a modified time in which the observation times are equidistant and to use stability properties of the convergence in $\mathbb{D}[0, T]$.

Bund and DAX, estimation of η , October 2010

© Theo Jalabert

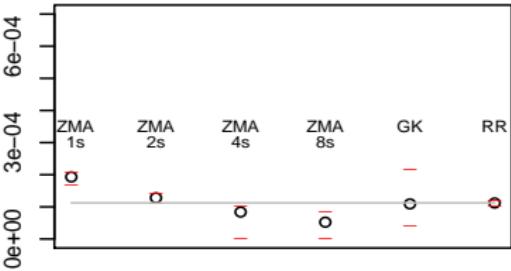
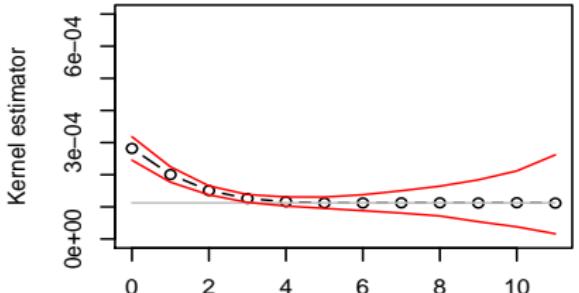
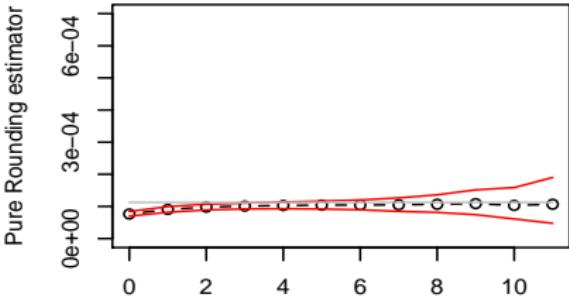
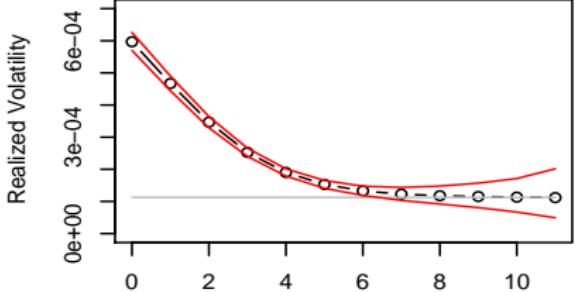


Day	η (Bund)	η (FDAX)	Day	η (Bund)	η (FDAX)
1 Oct.	0.18	0.41	18 Oct.	0.16	0.33
5 Oct.	0.15	0.37	19 Oct.	0.13	0.37
6 Oct.	0.15	0.37	20 Oct.	0.13	0.33
7 Oct.	0.15	0.38	21 Oct.	0.15	0.33
8 Oct.	0.15	0.41	22 Oct.	0.11	0.33
11 Oct.	0.14	0.36	25 Oct.	0.12	0.31
12 Oct.	0.14	0.36	26 Oct.	0.14	0.31
13 Oct.	0.14	0.32	27 Oct.	0.14	0.32
14 Oct.	0.16	0.35	28 Oct.	0.14	0.32
15 Oct.	0.16	0.35	29 Oct.	0.14	0.34

1000 MC simulations, 90% confidence interval

© Theo Jalabert





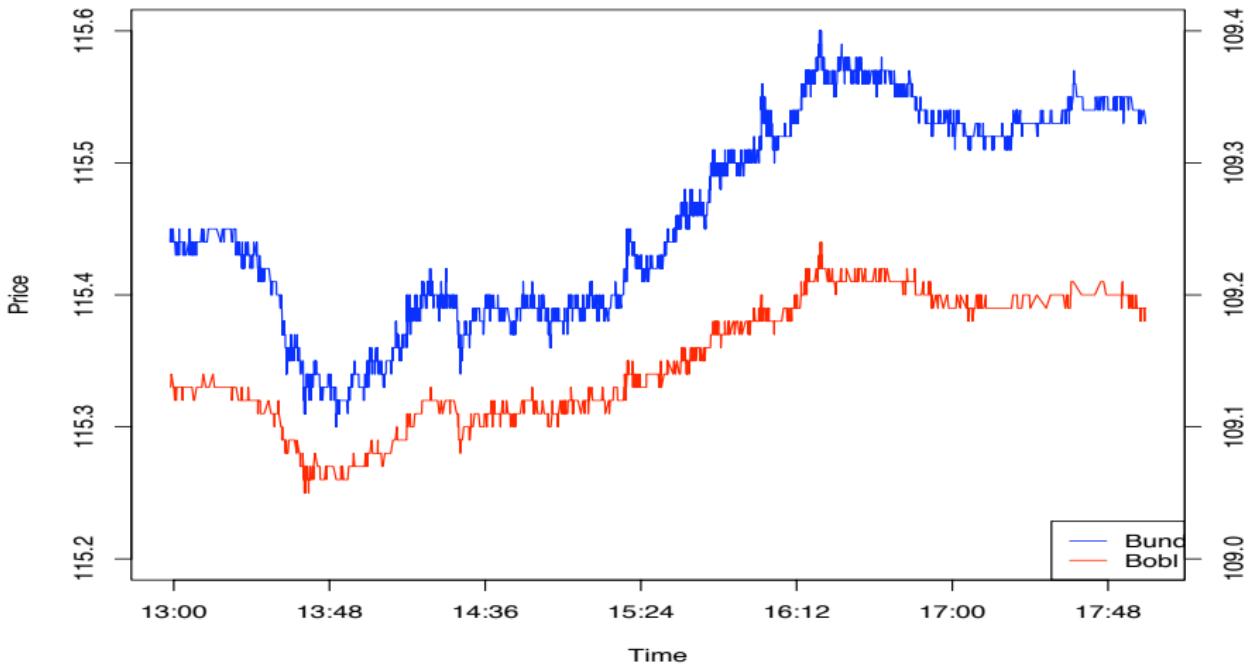


- 1 Modelling ultra high frequency data
- 2 Model with uncertainty zones
- 3 Volatility estimation
- 4 Covariation estimation
- 5 Lead-Lag estimation

Correlated assets : Bund-Bobl

© Théo Jalabert

Bund and Bobl





Dynamics of the efficient log prices

We now consider two assets.

$$d \log X_t = \mu_t^X dt + \sigma_{t-}^X dW_t,$$

$$d \log Y_t = \mu_t^Y dt + \sigma_{t-}^Y dB_t,$$

with

$$d\langle W, B \rangle_t = \rho_t dt.$$



Integrated covariance

We want to estimate

$$\int_0^1 \rho_t \sigma_t^X \sigma_t^Y dt.$$



Practical problems

We will face two main difficulties :

- Asynchronicity of the data,
- Microstructure effects.



Realized covariation

We start with the usual case without asynchronicity or microstructure effects. We observe $(X_{i/n}, Y_{i/n})$, $i = 0, \dots, n$. Let

$$\Delta_i^n X = \log X_{i/n} - \log X_{(i-1)/n}.$$

An estimator of

$$\int_0^1 \rho_t \sigma_t^X \sigma_t^Y dt$$

with accuracy $n^{-1/2}$ is given by

$$\hat{c}_n = \sum_{i=1}^n \Delta_i^n X \Delta_i^n Y.$$



Constant volatilities

Very often, traders like to think in term of correlation. When the correlation and volatility parameters are supposed to be constant :

$$\rho_t = \rho, \quad \sigma_t^X = \sigma^X, \quad \sigma_t^Y = \sigma^Y,$$

an estimator of ρ with accuracy $n^{-1/2}$ is given by

$$\frac{\widehat{c}_n}{\sqrt{\sum_{i=1}^n (\Delta_i^n X)^2 \sum_{i=1}^n (\Delta_i^n Y)^2}}.$$



Non constant volatilities

In the case where the volatility parameters are no longer constant, one can consider

$$\widehat{\rho}_n = \frac{2}{\pi} \frac{\widehat{c}_n}{\widehat{a}_n},$$

with

$$\widehat{a}_n = \sum_{i=1}^{n-1} |\Delta_{i+1}^n X \Delta_i^n Y|.$$

Indeed, \widehat{a}_n is an estimator of

$$\frac{2}{\pi} \int_0^1 \sigma_s^X \sigma_s^Y ds.$$

Non constant volatilities : convergence of \hat{a}_n

- Step 1 : we note that

$$\hat{a}_n \approx \sum_{i=1}^{n-1} \sigma_{\frac{i-1}{n}}^X \sigma_{\frac{i-1}{n}}^Y |W_{\frac{i+1}{n}} - W_{\frac{i}{n}}| |B_{\frac{i}{n}} - B_{\frac{i-1}{n}}|,$$

with the Brownian increments in the preceding sum being independent.

- Step 2 : \hat{a}_n has the same limit as

$$\sum_{i=1}^{n-1} \mathbb{E}_{\mathcal{F}_{\frac{i}{n}}} \left[\sigma_{\frac{i-1}{n}}^X \sigma_{\frac{i-1}{n}}^Y |W_{\frac{i+1}{n}} - W_{\frac{i}{n}}| |B_{\frac{i}{n}} - B_{\frac{i-1}{n}}| \right].$$

Non constant volatilities : convergence of \hat{a}_n

We have

$$\begin{aligned}& \sum_{i=1}^{n-1} \mathbb{E}_{\mathcal{F}_{\frac{i}{n}}} \left[\sigma_{\frac{i-1}{n}}^X \sigma_{\frac{i-1}{n}}^Y |W_{\frac{i+1}{n}} - W_{\frac{i}{n}}| |B_{\frac{i}{n}} - B_{\frac{i-1}{n}}| \right] \\&= \sum_{i=1}^{n-1} \sigma_{\frac{i-1}{n}}^X \sigma_{\frac{i-1}{n}}^Y |B_{\frac{i}{n}} - B_{\frac{i-1}{n}}| \mathbb{E}[|W_{\frac{i+1}{n}} - W_{\frac{i}{n}}|] \\&= \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{n}} \sum_{i=1}^{n-1} \sigma_{\frac{i-1}{n}}^X \sigma_{\frac{i-1}{n}}^Y |B_{\frac{i}{n}} - B_{\frac{i-1}{n}}|.\end{aligned}$$

Non constant volatilities : convergence of \hat{a}_n

The term

$$\sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{n}} \sum_{i=1}^{n-1} \sigma_{\frac{i-1}{n}}^X \sigma_{\frac{i-1}{n}}^Y |B_{\frac{i}{n}} - B_{\frac{i-1}{n}}|$$

has the same limit as

$$\sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{n}} \sum_{i=1}^{n-1} \sigma_{\frac{i-1}{n}}^X \sigma_{\frac{i-1}{n}}^Y \mathbb{E}[|B_{\frac{i}{n}} - B_{\frac{i-1}{n}}|] = \frac{2}{\pi} \frac{1}{n} \sum_{i=1}^{n-1} \sigma_{\frac{i-1}{n}}^X \sigma_{\frac{i-1}{n}}^Y,$$

which tends to (Riemann sum)

$$\frac{2}{\pi} \int_0^1 \sigma_s^X \sigma_s^Y ds.$$



Previous tick scheme

- Assume now we observe X at times $(T^{X,i}), i = 1, \dots$ and Y at times $(T^{Y,i}), i = 1, \dots$
- We build

$$\bar{X}_t = X_{T^{X,i}} \text{ for } t \in [T^{X,i}, T^{X,i+1})$$

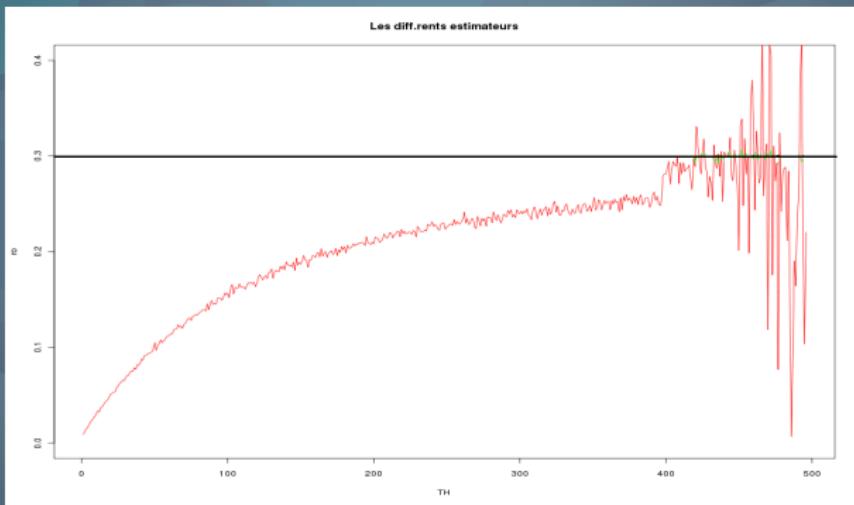
and

$$\bar{Y}_t = Y_{T^{Y,i}} \text{ for } t \in [T^{Y,i}, T^{Y,i+1}).$$

- For given h , the previous tick covariation estimator is

$$V_h = \sum_{i=1}^m (\log \bar{X}_{ih} - \log \bar{X}_{(i-1)h}) (\log \bar{Y}_{ih} - \log \bar{Y}_{(i-1)h}).$$

L'effet Epps





Epps effect

- Systematic bias for this estimator.
- Example : Assume that $\log X$ and $\log Y$ are two Brownian motions with correlation ρ and that the trade times are arrival times of two independent Poisson processes, then one can show that

$$\mathbb{E}[V_h] \rightarrow 0, \text{ as } h \rightarrow 0.$$



Hayashi-Yoshida Estimator

- Let $I_i^X = (T^{X,i}, T^{X,i+1}]$ and $I_j^Y = (T^{Y,i}, T^{Y,i+1}]$
- The Hayashi-Yoshida estimator is

$$U_n = \sum_{i,j} \Delta X(I_i^X) \Delta Y(I_j^Y) \mathbf{1}_{\{I_i^X \cap I_j^Y \neq \emptyset\}}.$$

- This estimator does not need any selection of h and is convergent if the arrival times are independent from the price.
- Nevertheless, it is not robust to microstructure effects.



Theorem

In the model with uncertainty zones, the Hayashi-Yoshida estimator is a consistent estimator of the covariation provided one uses the estimated values of the efficient prices.



- 1000 simulations,
- Black-Scholes model with U-shaped volatility,
- Constant covariation = 4.5×10^{-5} ($\rho = 0.4$).

	Hayashi-Yoshida	New Estimator
• Mean	8.55×10^{-5}	4.48×10^{-5}
Standard Deviation	7.8×10^{-6}	5.1×10^{-6}



- 1 Modelling ultra high frequency data
- 2 Model with uncertainty zones
- 3 Volatility estimation
- 4 Covariation estimation
- 5 Lead-Lag estimation



Observation from practitioners in finance

- Some assets are leading some other assets.
- This means that a “lagger” asset may partially reproduce the behavior of a “leader” asset.
- This common behavior is unlikely to be instantaneous. It is subject to some time delay called “lead-lag”.



Bachelier model

- For $t \in [0, 1]$, and $(B^{(1)}, B^{(2)})$ such that $\langle B^{(1)}, B^{(2)} \rangle_t = \rho t$, set

$$X_t := x_0 + \sigma_1 B_t^{(1)}, \quad \tilde{Y}_t := y_0 + \sigma_2 B_t^{(2)},$$

- Define $Y_t := \tilde{Y}_{t-\theta}$, $t \in [\theta, 1]$. Our lead-lag model is given by the bidimensional process (X_t, Y_t) .
- We have

$$\begin{cases} X_t &= x_0 + \sigma_1 B_t^{(1)} \\ Y_t &= y_0 + \rho \sigma_2 B_{t-\theta}^{(1)} + \sigma_2 (1 - \rho^2)^{1/2} W_{t-\theta} \end{cases}.$$



Estimation idea (1)

- Assume the data arrive at regular and synchronous time stamps in the Bachelier model, i.e. we have data

$$(X_0, Y_0), (X_{\Delta_n}, Y_{\Delta_n}), (X_{2\Delta_n}, Y_{2\Delta_n}), \dots, (X_1, Y_1),$$

and suppose $\theta = k_0 \Delta_n$, $k_0 \in \mathbb{Z}$.

- Let

$$\mathcal{C}_n(k) := \sum_i (X_{i\Delta_n} - X_{(i-1)\Delta_n})(Y_{(i+k)\Delta_n} - Y_{(i+k-1)\Delta_n}).$$



Estimation idea (2)

- Heuristically, we have

$$\mathcal{C}_n(k) \approx \Delta_n^{-1} \mathbb{E}[(X_{\cdot} - X_{\cdot - \Delta_n})(Y_{\cdot + k\Delta_n} - Y_{\cdot + (k-1)\Delta_n})] + \Delta_n^{1/2} \xi^n.$$

- Moreover,

$$\Delta_n^{-1} \mathbb{E}[(X_{\cdot} - X_{\cdot - \Delta_n})(Y_{\cdot + k\Delta_n} - Y_{\cdot + (k-1)\Delta_n})] = \begin{cases} 0 & \text{if } k \neq k_0 \\ \rho \sigma_1 \sigma_2 & \text{if } k = k_0. \end{cases}$$

- Thus we can (asymptotically) detect the value k_0 that defines θ in the very special case $\theta = k_0 \Delta_n$ by maximizing in k the contrast sequence

$$k \rightsquigarrow |\mathcal{C}_n(k)|.$$



Let $\theta > 0$ (for simplicity, extensions are quite straightforward) and set $\mathbb{F}^\theta = (\mathcal{F}_t^\theta)_{t \geq 0}$, with $\mathcal{F}_t^\theta = \mathcal{F}_{t-\theta}$.

Assumptions

- We have

$$X = X^c + A, \quad Y = Y^c + B.$$

- $(X_t^c)_{t \geq 0}$ is a continuous \mathbb{F} -local martingale, and $(Y_t^c)_{t \geq 0}$ is a continuous \mathbb{F}^θ -local martingale.
- $\exists v_n \rightarrow 0$, $v_n^{-1} \max \left\{ \sup\{|I_i^X|\}, \sup\{|I_i^Y|\} \right\} \rightarrow 0$.
- The $T^{X,i}$ are \mathbb{F}^{v_n} -stopping times and the $T^{Y,i}$ are $\mathbb{F}^{\theta+v_n}$ -stopping times.



Estimator

- We set

$$U_n(\theta) = \sum_{i,j} \Delta X(I_i^X) \Delta Y(I_j^Y) 1_{\{I_i^X \cap (I_j^Y)_{-\theta} \neq \emptyset\}},$$

with $(I_j^Y)_{-\theta} = (T^{Y,j} - \theta, T^{Y,j+1} - \theta]$.

- Eventually, $\widehat{\theta}_n$ is defined as a solution of

$$|U_n(\widehat{\theta}_n)| = \max_{\theta \in \mathcal{G}^n} |U_n(\theta)|,$$

where \mathcal{G}^n is a sufficiently fine grid.



Theorem

As $n \rightarrow \infty$,

$$\nu_n^{-1}(\widehat{\theta}_n - \theta) \rightarrow 0,$$

in probability, on the event $\{\langle X^c, \tilde{Y}^c \rangle_T \neq 0\}$.



Idea of proof

We show that

- If we compute the contrast function over points θ_n of the grid \mathcal{G}^n such that $|\theta_n - \theta|$ is bigger than v_n , then the contrast function goes to zero.
- If we compute the contrast function over points θ_n of the grid \mathcal{G}^n such that $|\theta_n - \theta|$ is sufficiently small then the contrast function goes to the covariation between X and \tilde{Y} .



Setup

- We consider 300 simulations of the Bachelier model with synchronous equispaced data with period Δ_n .
- $t \in [0, 1]$, $\theta = 0.1$, $x_0 = \tilde{y}_0 = 0$, $\sigma_1 = \sigma_2 = 1$.
- The mesh size of the grid h_n satisfies $h_n = \Delta_n$.
- We consider the following variations :
 - Mesh size : $h_n \in \{10^{-3}(FG), 3.10^{-3}(MG), 6.10^{-3}(CG)\}$.
 - Correlation value : $\rho \in \{0.25, 0.5, 0.75\}$.

Results in the synchronous case

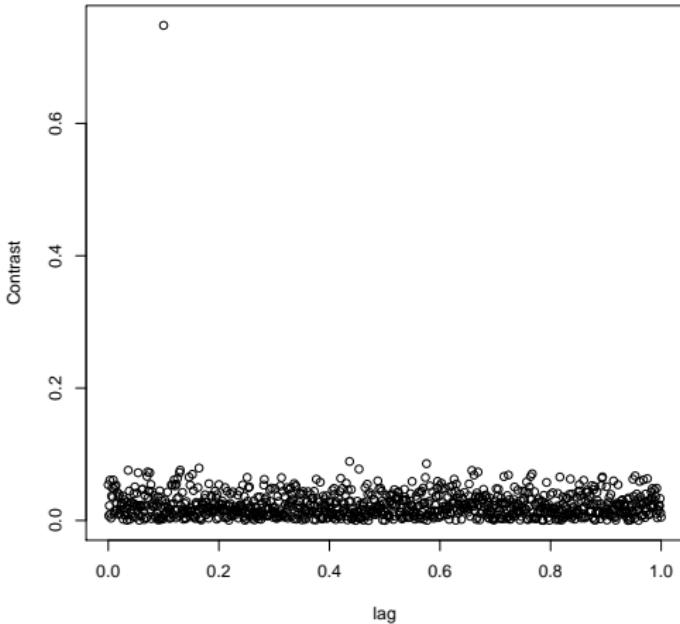
© Théo Jalabert

$\hat{\theta}_n$	0.096	0.099	0.1	0.102	Other
FG, $\rho = 0.75$	0	0	300	0	0
MG, $\rho = 0.75$	0	300	0	0	0
CG, $\rho = 0.75$	1	0	0	299	0
FG, $\rho = 0.50$	0	0	300	0	0
MG, $\rho = 0.50$	0	299	0	1	0
CG, $\rho = 0.50$	13	0	0	280	7
FG, $\rho = 0.25$	0	0	300	0	0
MG, $\rho = 0.25$	0	152	0	11	137
CG, $\rho = 0.25$	10	0	0	66	124

Table 1 : *Estimation of $\theta = 0.1$ on 300 simulated samples for $\rho \in \{0.25, 0.5, 0.75\}$.*

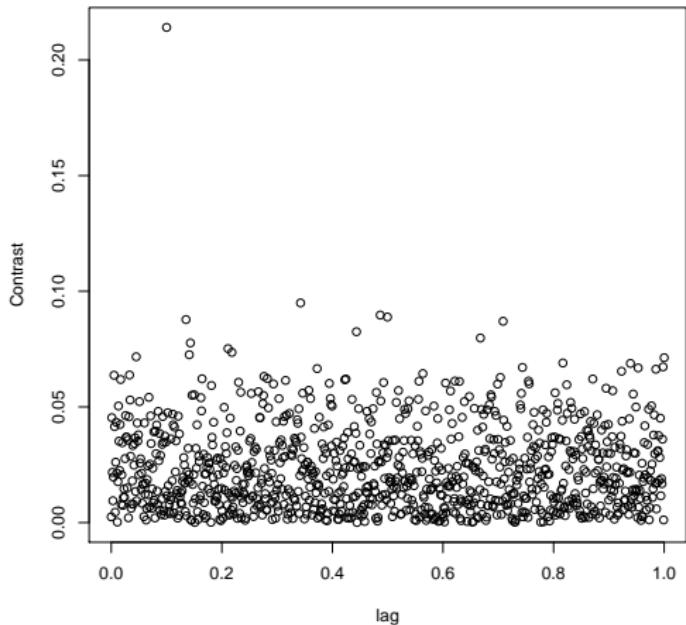
One sample path, FG, $\rho = 0.75$

© Théo Jalabert



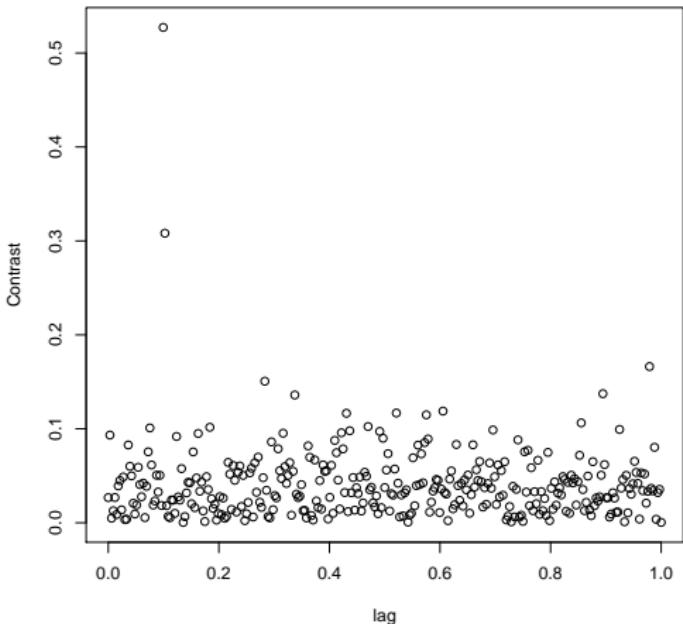
One sample path, FG, $\rho = 0.25$

© Théo Jalabert



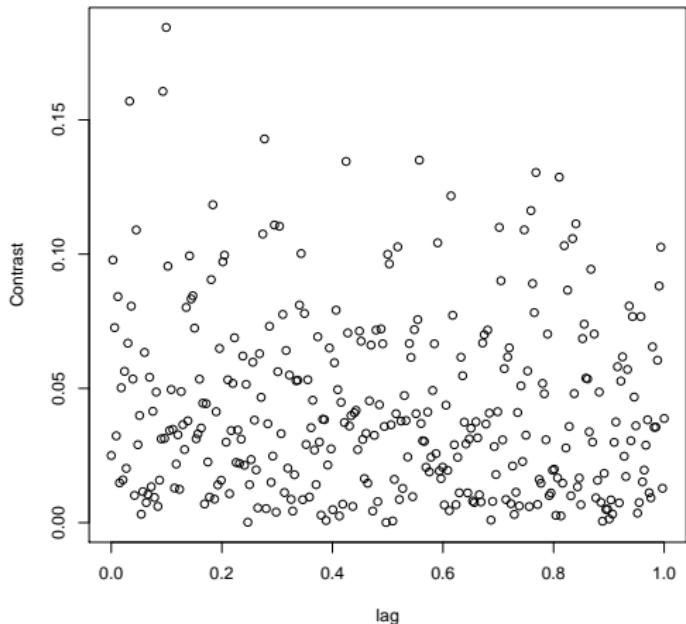
One sample path, MG, $\rho = 0.75$

© Théo Jalabert



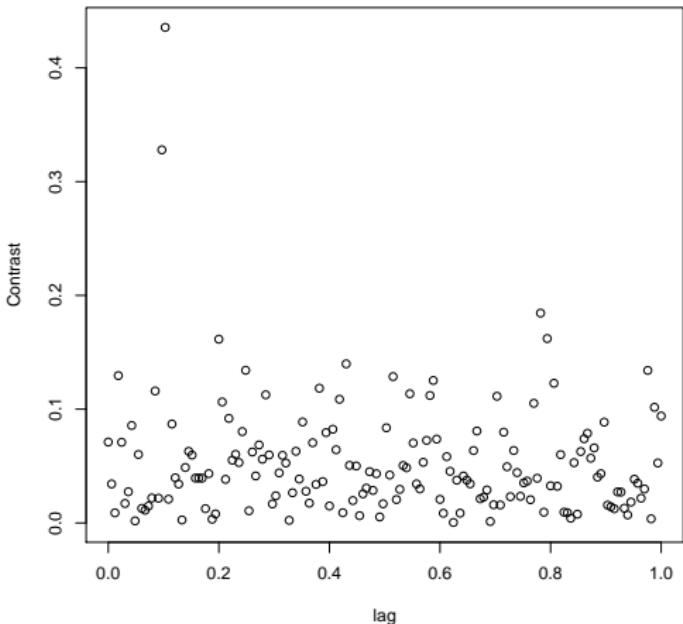
One sample path, MG, $\rho = 0.25$

© Théo Jalabert



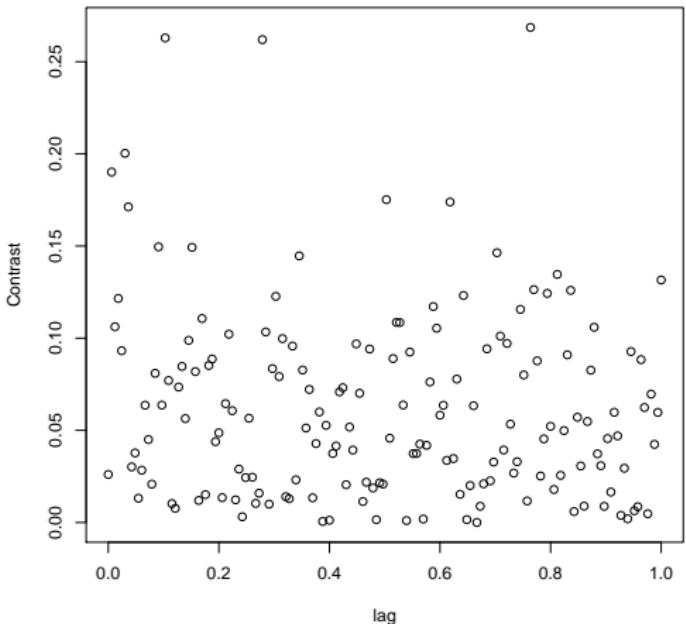
One sample path, CG, $\rho = 0.75$

© Théo Jalabert



One sample path, CG, $\rho = 0.25$

© Théo Jalabert





Setup

- We randomly pick 300 sampling times for X over $[0, 1]$, uniformly over a grid of mesh size 10^{-3} .
- We randomly pick 300 sampling times for Y likewise, and independently of the sampling for X .
- Fine grid case, with $\theta = 0.1$ and $\rho = 0.75$.

Results for the non synchronous case

© Théo Jalabert



$\hat{\theta}$	0.099	0.1	0.101	0.102	0.103	0.104	0.105
FG, $\rho = 0.75$	16	106	107	46	19	4	2

Table 2 : *Estimation of $\theta = 0.1$ on 300 simulated samples for $\rho = 0.75$ and non-synchronous data.*



Dataset

We study here the lead-lag relationship between the two following assets :

- The future contract on the DAX index, with maturity December 2010,
- The Euro-Bund future contract, with maturity December 2010.



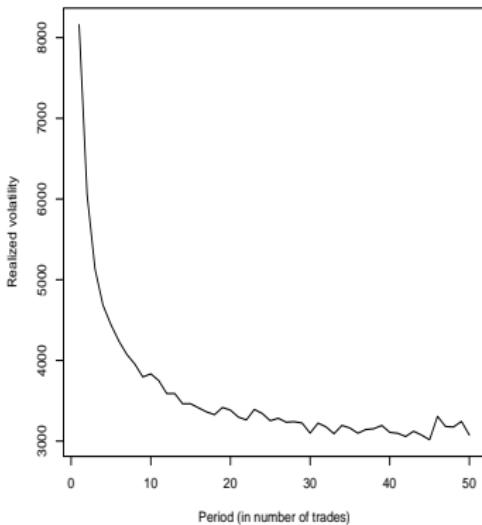
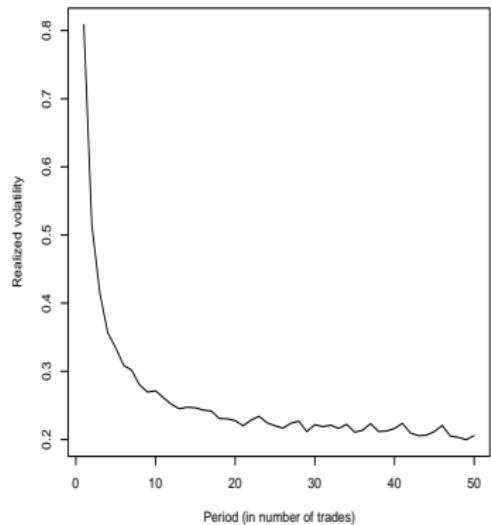
Methodology

- We want to use high frequency data.
- First approach : use of the model with uncertainty zones.
- Here we just use signature plots in trading times. This enables to take advantage of non synchronous data.
- We keep one trade out of twenty.
- We then compute the function $U_n(\theta)$ over these trades.

Signature plots, October 13, for Bund (left) and FDAX (right).

© Theo Jalabert

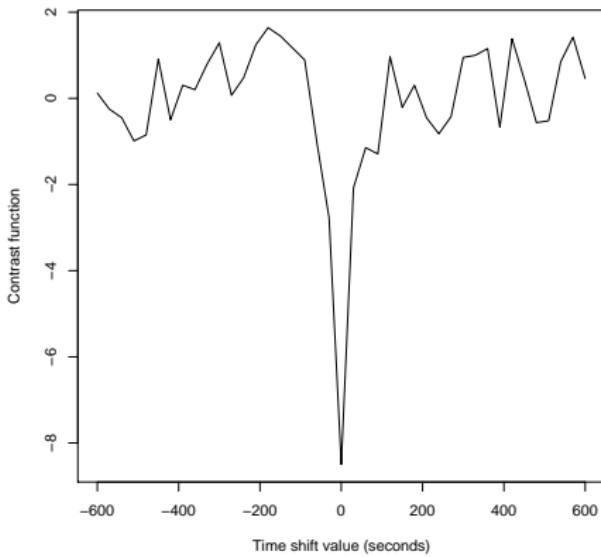
Theo Jalabert



Function U^n , October 13, between -10 and 10 minutes, mesh=30 seconds

© Theo Jalabert

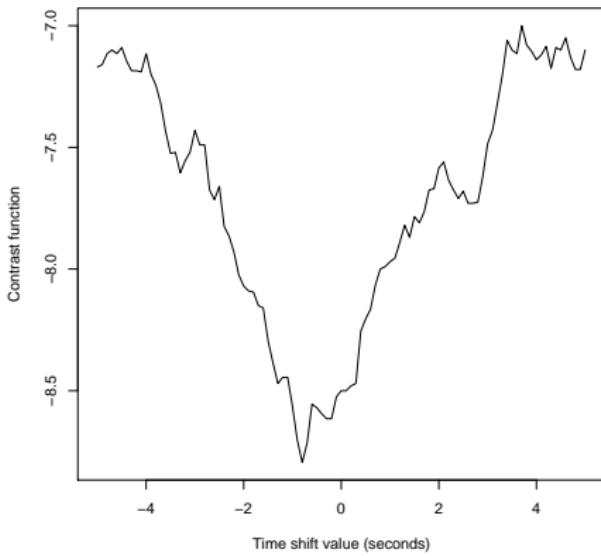
Theo
Jalabert



Function U^n , October 13, between -5 and 5 seconds,
mesh=0.1 second

© Théo Jalabert

Théo Jalabert



Bund and DAX, lead-lag estimation

© Théo Jalabert



Jour	Vol.(Bund)	Vol.(FDAX)	LL.	J.	Vol.B.	Vol.F.	LL
1 Oct.	2847	4215	-0.2	18 Oct.	1727	2326	-2.1
5 Oct.	2213	3302	-1.1	19 Oct.	2527	3162	-1.6
6 Oct.	2244	2678	-0.1	20 Oct.	2328	2554	-0.5
7 Oct.	1897	3121	-0.5	21 Oct.	2263	3128	-0.1
8 Oct.	2545	2852	-0.6	22 Oct.	1894	1784	-1.2
11 Oct.	1050	1497	-1.4	25 Oct.	1501	2065	-0.4
12 Oct.	2265	3018	-0.8	26 Oct.	2049	2462	-0.1
13 Oct.	2018	3037	-0.8	27 Oct.	2606	2864	-0.6
14 Oct.	2057	2625	0.0	28 Oct.	1980	2632	-1.3
15 Oct.	2571	3269	-0.7	29 Oct.	2262	2346	-1.6