

Physique des marchés

Faits stylisés

[<https://s.42l.fr/PhM>]

[damien.challet@centralesupelec.fr]

25 février 2025

Plan du cours

date	heure	sujet	salle
24.02	8 : 15	faits stylisés	MI.106
04.03	8 : 15	stratégies	sd.206
11.03	8 : 15	prévisibilité et stratégies	amphi IV
25.03	8 : 15	prévisibilité et stratégies	sd.101
31.03	8 : 15	modèles d'agents	Amphi III
01.04	8 : 15	dynamique de la prévisibilité	EB.118
08.04	8 : 15	dynamique de la prévisibilité	?
15.04	8 : 15	contrôle	?

Organisation du cours

1. Habituellement, 50% théorie, 50% TDs
2. Contrôle : essentiellement des sujets des TPs
 - 2.1 dynamique de la prévisibilité : apparition / disparition
 - 2.2 modèles d'agents vus au cours
 - 2.3 stratégies spéculatives : par exemple critiquer un article

Physique des marchés

- But du cours

Comprendre la dynamique des marchés

- Ingrédients

1. stratégies d'investissement
2. agents + apprentissage
3. interaction entre agents

- Corollaires :

1. étude de la spéculation
2. dynamique de la prévisibilité
3. origine des grandes fluctuations

Approche dynamique des marchés

Dynamique de prix

$$P_{t+1} = F_t(\Omega_{t+1}), \quad \Omega_t = \text{information disponible au temps } t$$

- F_t ?
- Variables importantes ?
- Unité de temps ?
- Propriétés dynamiques ?

Finance traditionnelle

$$\Omega_{t+1} = \{P_t, D_{t+1}, \epsilon_{t+1}\}$$

$$P_{t+1} = (1+r)P_t + D_{t+1} + \epsilon_{t+1}$$
$$\epsilon \sim \mathcal{N}(0, \sigma^2)$$

- r : taux d'intérêt
- D_{t+1} : dividendes connues au temps $t+1$
- $\epsilon_{t+1} \sim \mathcal{N}(0, \sigma^2)$: nouvelles connues au temps $t+1$
- unité de temps : semaine, mois, etc. (basse fréquence)

Mathématiques financières

volatilité stochastique :

$$\Omega_{t+1} = \{P_t, D_{t+1}, \epsilon_{t+1}, \sigma_t\}$$

$$P_{t+1} = (1+r)P_t + D_{t+1} + \epsilon_{t+1}$$

$$\epsilon_{t+1} \sim \mathcal{N}(0, \sigma_t^2)$$

$$\sigma_{t+1} = F_\sigma(\{P_{t-k}\}_{k \geq 0}, \{\epsilon_{t-k'}\}_{k' \geq 0})$$

Boucle de rétroaction entre changements de prix passés et future volatilité

F_σ est fixé par le modèle.

Physique des Marchés

1. Données → faits stylisés
2. Mécanismes → faits stylisés
 - 2.1 approche microscopique : agents
 - 2.2 agent = investisseur = stratégies + apprentissage

Approche dynamique des marchés

Dynamique de prix

$$P_{t+1} = F_t(\Omega_t)$$

- Finance traditionnelle : $E(P_{t+1})$
- Mathématiques financières : e.g. quel modèle de volatilité pour déterminer le prix d'une option ?
- Ce cours : quels mécanismes produisent une F_t qui reproduit la dynamique de P_t ?

Mécanisme principal : instabilité intrinsèque

Point de vue canonique

principale source de fluctuations de prix : nouvelles
autre source : noise traders

Point de vue réaliste

fluctuations endo-/exogènes
réactions en cascade
sur-réactions
instabilités
non-linéarités

Vision traditionnelle

Les marchés

calcurent

sont sans danger

sont intelligents

sont efficents

CONFiance !

Vue moderne et future

Les marchés sont

désordonnés

complexes

instables

(in)efficients

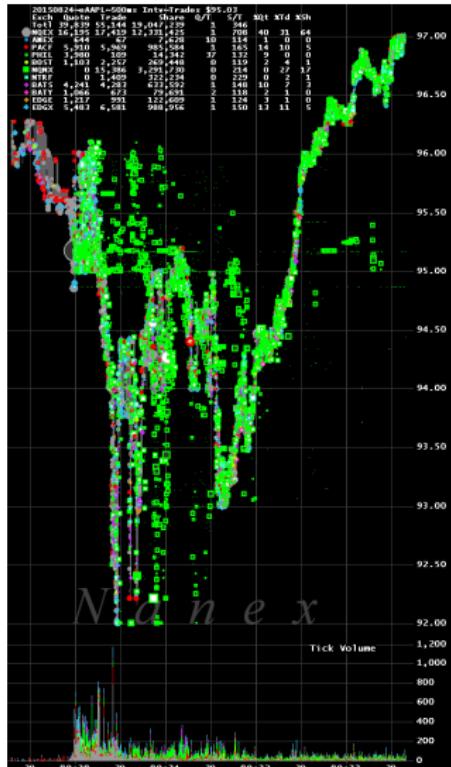
intéressants !

MÉFIANCE !

Efficients ?

- 1. Exploitation de la prévisibilité : très bons
 - 2. Prix justes : à un facteur 2 près, 90% du temps [Black 1986] [Bouchaud et al. 2017]
 - 3. Equilibre physique : exceptionnellement
 - 4. Equilibre compétitif : beaucoup de bruit
 - 5. Rendements gaussiens : 60-95%

Efficients ?



source [nanex.com]

Inefficients ?

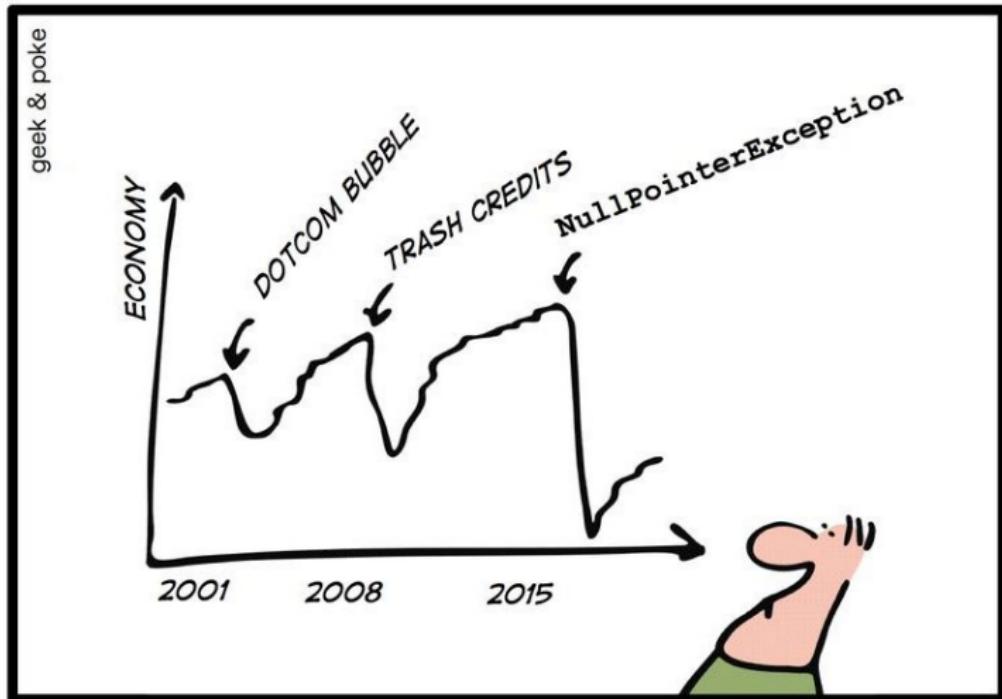
1. Êtres humains

- 1.1 finance comportementale
- 1.2 contraintes, lois

2. Ordinateurs

- Trading algorithmique
- HFT (High Frequency Trading)
- “Computers are useless. They can only give you answers.”
(Pablo Picasso)
- “To err is human, but to really foul things up you need a computer.”
(Paul Ehrlich)

Trading algorithmique



*MODERN HIGH-FREQUENCY TRADING WOULD
BE IMPOSSIBLE WITHOUT SUPER
SOPHISTICATED SOFTWARE*

What could happen?

Knight Shows How to Lose \$440 Million in 30 Minutes][SEC report]

Knight Shows How to Lose \$440 Million in 30 Minutes

By Matthew Philips

August 3, 2012, 12:10 AM GMT+2



Thomas Joyce talks about yesterday's trading errors on Bloomberg Television. Source: Bloomberg

Talk about a bad day. In the mother of all computer glitches, market-making firm Knight Capital Group lost \$440 million in 30 minutes on Aug. 1 when its

Hypothèses

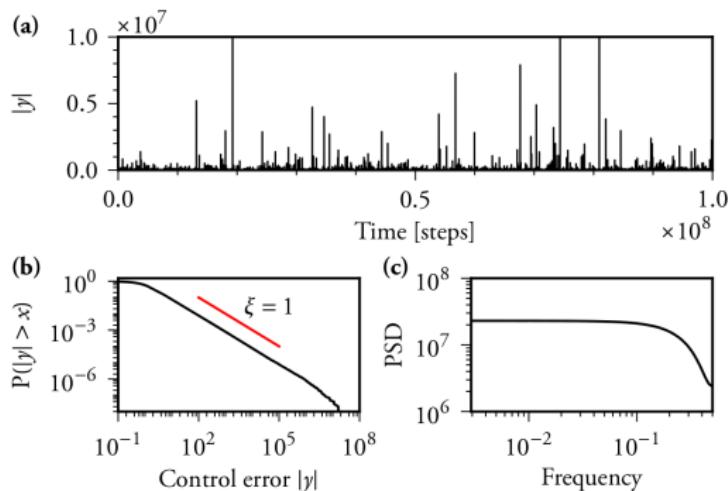
1. Trading algorithmique : majoritaire sur les marchés
2. Influence des stratégies : primordiale

$$P_{t+1} = F_t(\{P_{t-k}\}, \{\text{stratégies}\}, \text{nouvelles, tout le reste})$$

- Quelles stratégies ?
- Dynamique d'utilisation des stratégies ?
- Bouchaud et Challet (2016) [lien] : comportement du marché dû au binôme stratégies + apprentissage

Exemple : apprentissage et boucle de rétro-action

1. Equilibrage d'un bâton, J. L. Cabrera and J. G. Milton, On-off intermittency in a human balancing task, Phys. Rev. Lett. 89 158702 (2002).
2. Suppression optimale d'un signal, Pazelt and Pawelzik : "critically adaptive control" (2011)



Finance quantitative : régularité dans les données

Dynamique des prix

$$P_{t+1} = F_t(\Omega_t), \quad \Omega_t = \text{passé}$$

Propriétés de P_t ?

Régularités \equiv faits stylisés \leq lois naturelles

Faits stylisés : prix $p(t)$

- Incrément

$$\delta_\tau P(t) = P(t) - P(t - \tau)$$

- Rendement [%]

$$R_\tau(t) = \frac{P(t) - P(t - \tau)}{P(t - \tau)}$$

- Log-rendement

$$r_\tau(t) = \log P(t) - \log P(t - \tau) = \log(1 + R_\tau(t))$$

- % et log

$$R_\tau(t) = e^{r_\tau(t)} - 1$$

- Cumulation temporelle : log-rendements

$$r_\tau(t) = \sum_{t-\tau < t' \leq t} r_1(t')$$

- Pondération de titres : rendements

$$r_\tau^{(\text{portfolio})}(t) = \sum_{i=1}^N w_i R_\tau^{(i)}(t)$$

Quel prix ?



Au temps t

- meilleure offre d'achat $b(t)$ (bid)
- meilleure offre de vente $a(t)$ (ask)
- offre moyenne $m(t) = \frac{b(t)+a(t)}{2}$ (midprice)
- prix de la dernière transaction $P(t)$ (last)

Quels prix ?

Carnets d'ordres

- toutes les offres d'achats
- toutes les offres de vente

NESN (NESTLE N) Orderbook

Last trade / volume	CHF 44.66 / 500	Date / Time	18.11.2008 / 09:32:50
Daily change / absolute	0.09% / +0.04		
Cum. Volume	Bid Volume	Price	Ask Volume
		44.90	1'000 (1)
		44.86	8'539 (3)
		44.84	10'000 (1)
		44.82	5'576 (6)
		44.80	10'551 (3)
		44.78	3'677 (1)
		44.76	6'800 (2)
		44.70	9'061 (3)
		44.68	6'593 (4)
		44.66	6'573 (5)
2'049 (3)	2'049 (3)	44.60	
7'225 (4)	5'176 (1)	44.54	
8'305 (5)	1'080 (1)	44.50	
13'305 (6)	5'000 (1)	44.48	
30'081 (8)	16'776 (2)	44.46	
32'081 (9)	2'000 (1)	44.44	
47'430 (12)	15'349 (3)	44.42	
58'757 (16)	11'327 (4)	44.40	
60'757 (17)	2'000 (1)	44.38	
68'557 (20)	7'800 (3)	44.30	
Cum. Volume	Bid Volume	Price	Ask Volume

The values in brackets in the column Ask volume / Bid volume show the amount of sell-respectively buy orders. In the columns Cum. volume the cumulated amount of sell- respectively buy orders are stated.

Recent example : BTC-USD

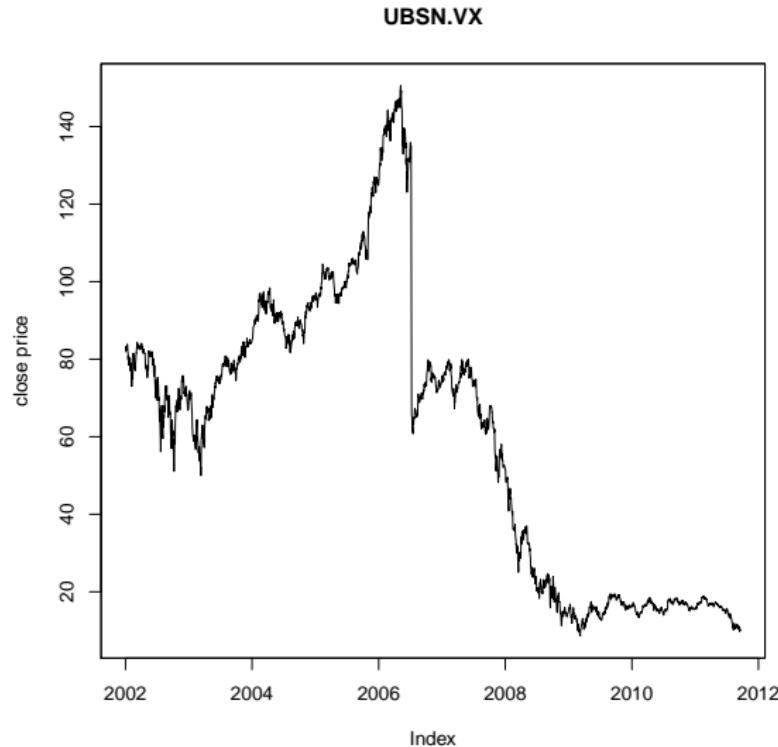
BITFINEX  For Traders  For Lenders  About Us  UNUS SED LEO Affiliate Program

ALL USD EUR GBP JPY BTC ETH USDT CNHT EURT XAUT MIM MXNT TRY 

ORDER BOOK BTC/USD						
AMOUNT	TOTAL	PRICE	PRICE	TOTAL	AMOUNT	
0.36275664	0.36275664	16,829.0000	16,830.0000	3,9338	3,9338	
0.48751605	0.85027269	16,826.0000	16,831.0000	5,1334	1,1997	
1.9825	2,8328	16,825.0000	16,832.0000	6,4511	1,3176	
2.3031	5,1359	16,824.0000	16,833.0000	9,3835	2,9324	
2.2861	7,4220	16,823.0000	16,834.0000	12,3409	2,9574	
0.85893164	8,2809	16,822.0000	16,835.0000	18,4170	6,0761	
1.5403	9,8212	16,821.0000	16,836.0000	20,2060	1,7890	
2.1685	11,9897	16,820.0000	16,837.0000	20,4534	0,24741000	
3.6239	15,6136	16,818.0000	16,838.0000	22,3052	1,8517	
0.60780000	16,2214	16,817.0000	16,839.0000	23,3791	1,0740	
1.9069	18,1283	16,816.0000	16,840.0000	25,6801	2,3010	
0.59887581	18,7272	16,815.0000	16,841.0000	26,9664	1,2863	
4.2230	22,9502	16,814.0000	16,842.0000	29,8027	2,8362	
0.03471473	22,9849	16,813.0000	16,843.0000	43,2992	13,4965	
1.1471	24,1320	16,812.0000	16,844.0000	44,1779	0,87869311	

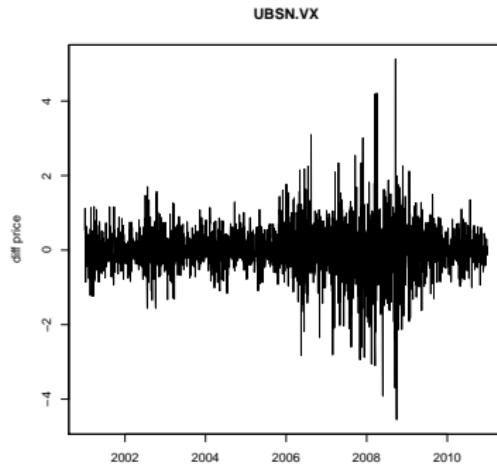
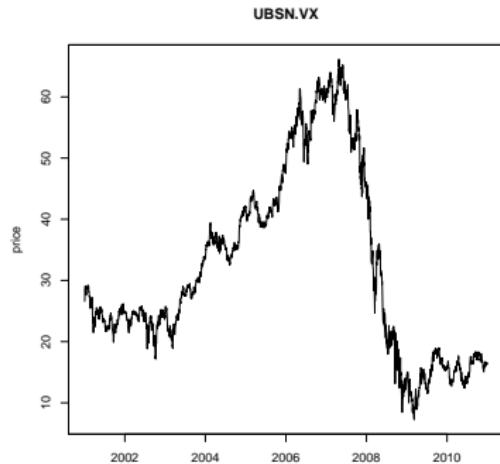
Échelles de temps : mois

UBSN.VX



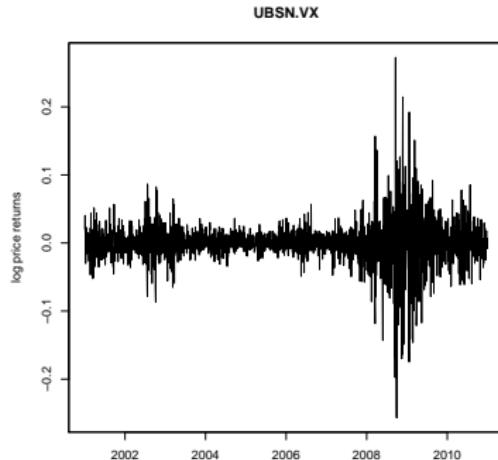
Échelles de temps : mois

UBSN.VX : prix ajustés dividendes / splits



Échelles de temps : mois

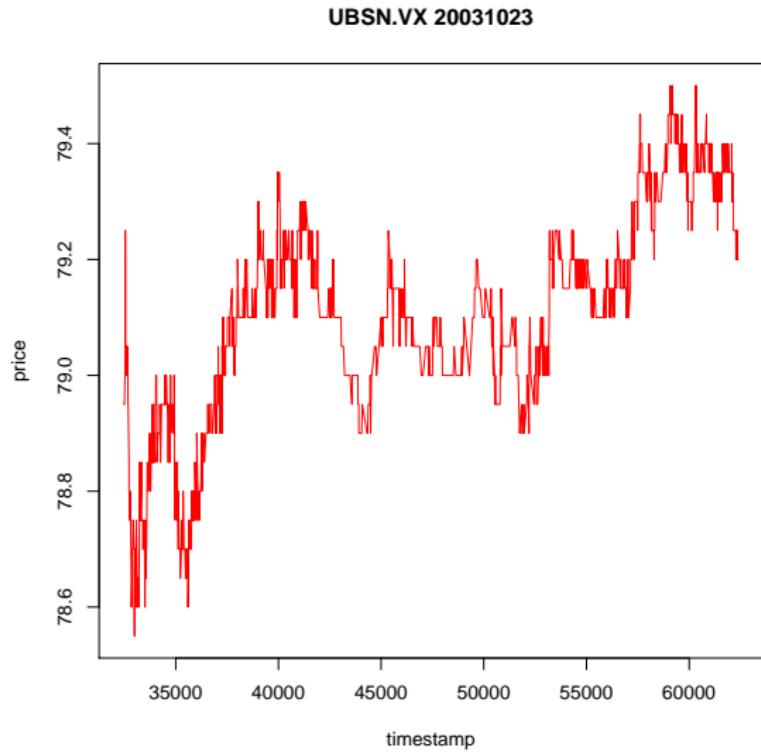
UBSN.VX : log-prix ajustés dividendes / splits



Fractale (Mandelbrodt) ?

Échelles de temps : jour

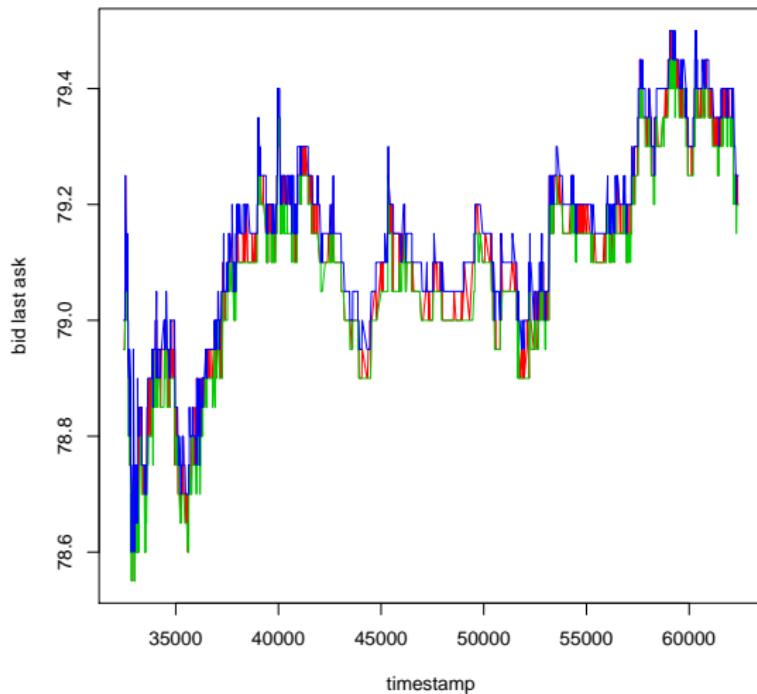
UBSN.VX : prix de la dernière transaction



Échelles de temps : jour

UBSN.VX : bid ask and last

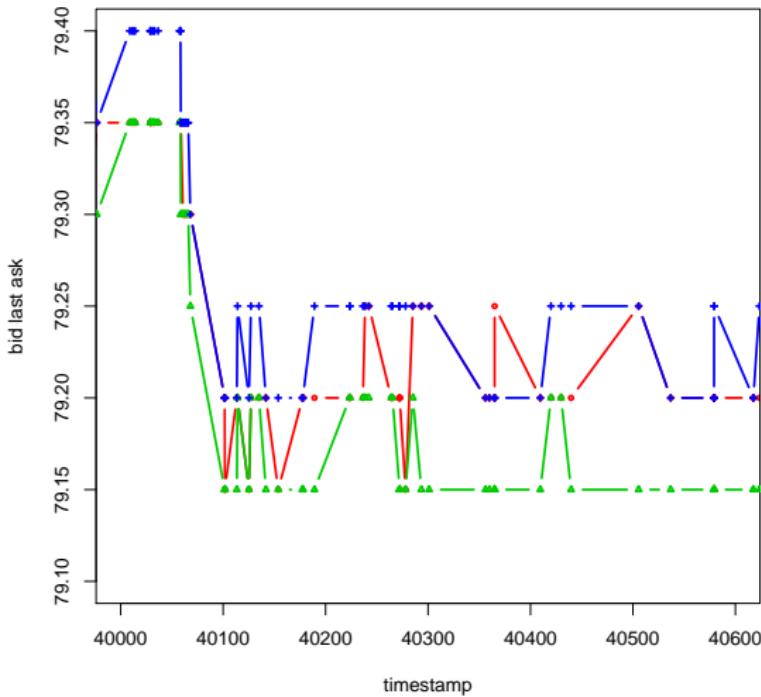
UBSN.VX 20031023



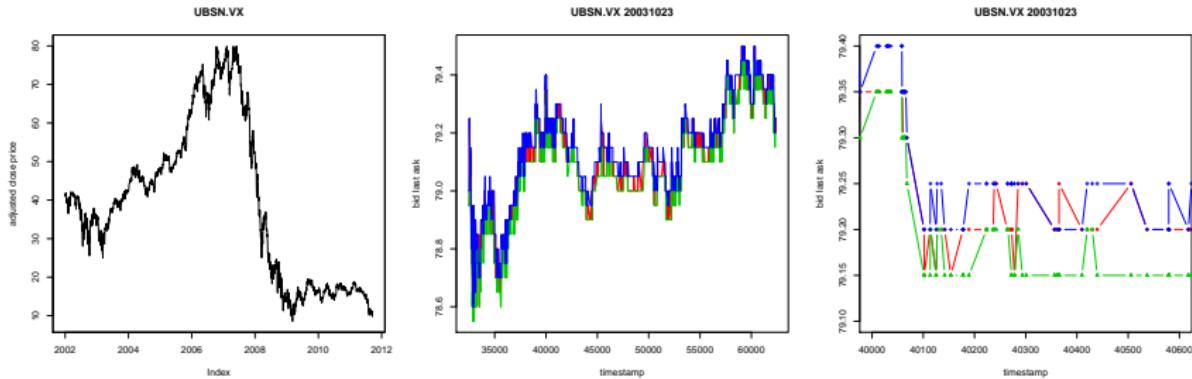
Échelles de temps : jour

UBSN.VX : tic(k), bid/ask bounce

UBSN.VX 20031023



Échelles de temps



- incréments de prix

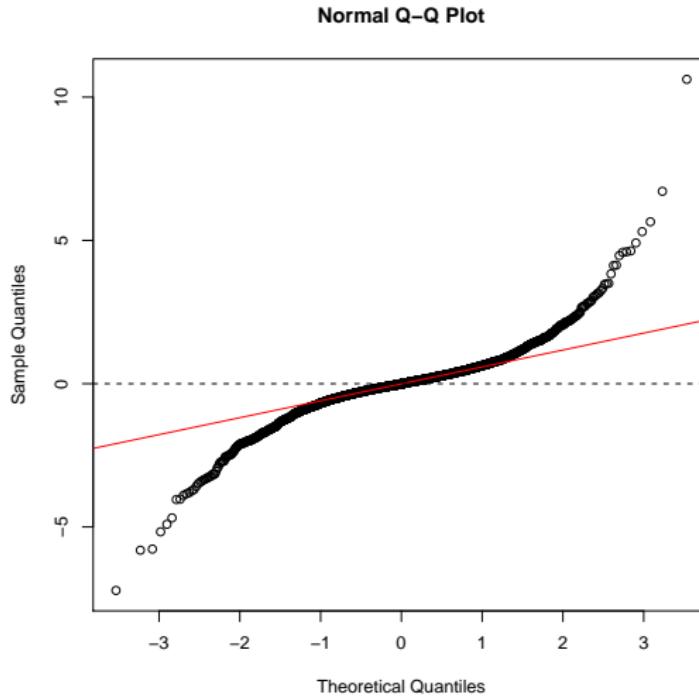
$$\delta_\tau P = \sum_{t-\tau < t' \leq t} \delta P_1(t')$$

- log-rendement

$$r_\tau(t) = \sum_{t-\tau < t' \leq t} r_1(t')$$

Statistique des prix : rendements, qqplot

Distribution UBSN.VX, rendements normalisés



Distributions à queues lourdes

- Queue

$$P(x) |x| > x_0$$

$$P(X > x_0) = P_{>}(x_0) < \epsilon$$

- Queue lourde

$$\lim_{x \rightarrow \infty} P(x) e^{\lambda x} = \infty \quad \forall \lambda > 0$$

- Queue semi-lourde

$$\exists \lambda_0 > 0 \text{ t.q. } \lim_{x \rightarrow \infty} P(x) e^{\lambda_0 x} = c < \infty, \quad c > 0$$

Queues lourdes : détection graphique

- Queue semi-lourde

$$P_>(x) \propto e^{-\lambda_0 x}$$

$$\log P_>(x) = -\lambda_0 x + \text{cst}$$

- Queue lourde $\iff P_>(x)$ décroît (suffisamment) plus lentement que toute exponentielle

$\implies \log P_>(x)$ est une fonction (suffisamment) convexe de x

Queues lourde : détection statistique

1. Calculer $x = |r| / E(|r|)$
2. (tracer les quantiles q correspondant à la fonction cumulatrice de x en fonction des quantiles théoriques d'une exponentielle de moyenne 1.)
3. Calculer $\max_q [q(\phi) - q_{exp}(\phi)]_+$, $\phi \in]0, 1[$

Comparaison de distributions

- N échantillons $\{x_i\}$
- Distributions 1 and 2 avec paramètres K_1 et K_2
→ log-vraisemblance \mathcal{L}_1 and \mathcal{L}_2
- Vuong's closeness test : z-score

$$Z = \frac{\mathcal{L}_1 - \mathcal{L}_2 - \frac{(K_1 - K_2)}{2} \log N}{\sqrt{N} \omega_N}$$

$$\text{où } \omega_N^2 = \frac{1}{N} \sum_i \left(\log \frac{P_1(x_i)}{P_2(x_i)} \right)^2$$

- $Z > 0$: distribution 1 est préférable → p-valeur
- Queue lourde ? Test de Vuong entre distributions exponentielle et à queue grasse .

Statistique des prix : rendements, index

Bouchaud et Potters, Theory of Financial Risk and Derivative Pricing, 2e édition (2003)

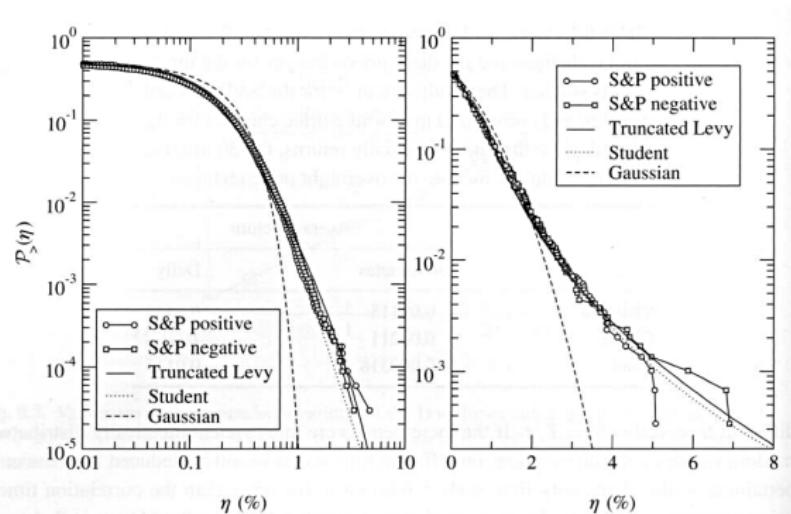


Fig. 6.5. Elementary cumulative distribution $P_{1>}(\eta)$ (for $\eta > 0$) and $P_{1<}(\eta)$ (for $\eta < 0$), for the S&P 500 returns, with $\tau = 30$ min (left, log-log scale) and 1 day (right, semi-log scale). The thick line corresponds to the best fit using a symmetric TLD $L_\mu^{(t)}$, of index $\mu = \frac{3}{2}$. We have also shown the best Student distribution and the Gaussian of same RMS

Statistique des prix : rendements, contrat à terme

Bouchaud et Potters, Theory of Financial Risk and Derivative Pricing, 2e édition (2003)

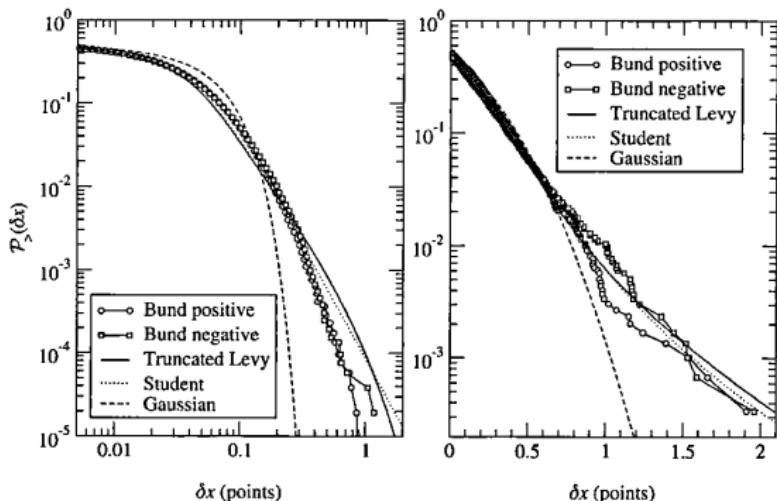
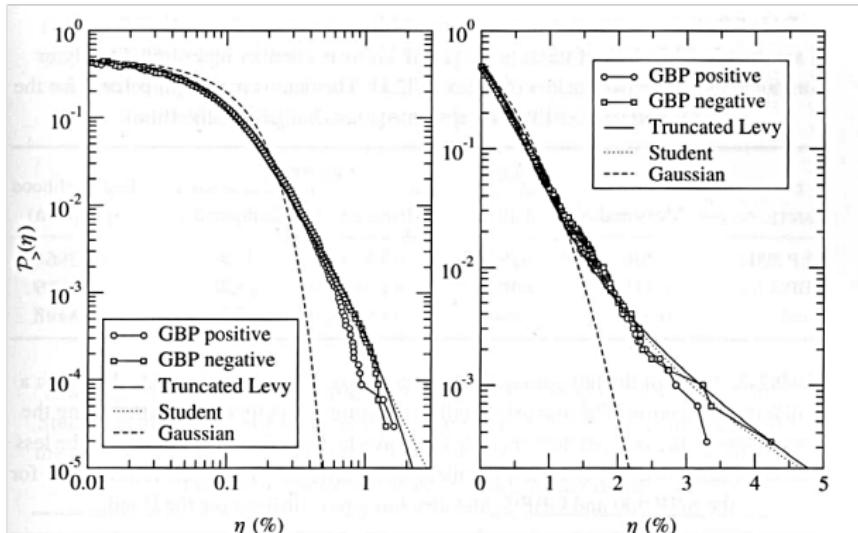


Fig. 6.7. Elementary cumulative distribution for the price increments (i.e. not the returns) of the Bund, for $\tau = 30$ min (left) and 1 day (right), and best fit using a symmetric TLD $L_\mu^{(0)}$, of index $\mu = \frac{3}{2}$, Student and Gaussian. In this case, the TLD and Student are not very good fits. A better fit is obtained using a TLD with a smaller μ .

Statistique des prix : rendements, marché des changes

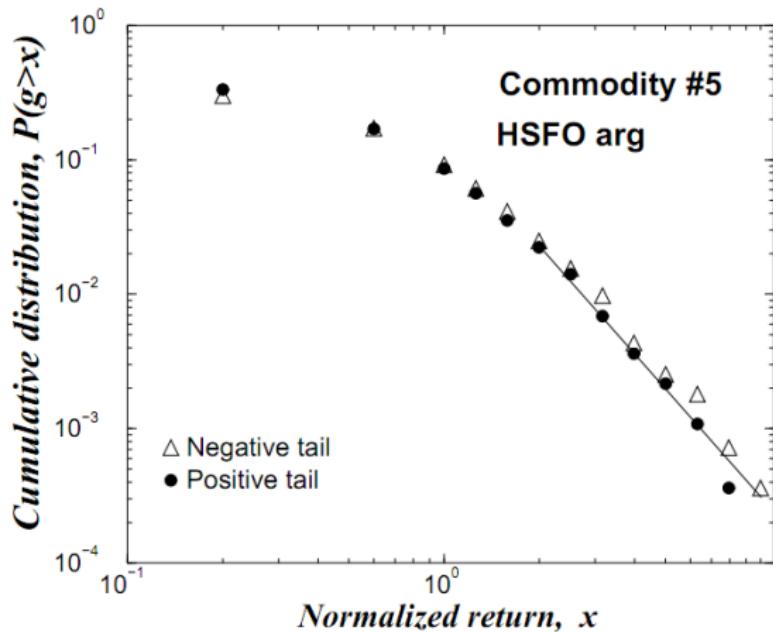
Bouchaud et Potters, Theory of Financial Risk and Derivative Pricing, 2e édition (2003)



g. 6.6. Elementary cumulative distribution for the GBP/\$ returns, for $\tau = 30$ min (left, log-log scale) and 1 day (right, semi-log scale), and best fit using a symmetric TLD $L_\mu^{(t)}$, of index $\mu = \frac{3}{2}$. We again show the best Student distribution and the Gaussian of same RMS.

Statistique des prix : rendements, matières premières

Matia et al. (2002)



Théorème central limite : gaussiennes

- $P_1(x)$, moyenne m , variance $\sigma^2 < \infty$
- x_i , $i = 1, \dots, N$, iid
- Somme $X_N = \sum_{i=1}^N x_i$

$$\lim_{N \rightarrow \infty} P\left(u_1 \leq \frac{X_N - mN}{\sigma\sqrt{N}} \leq u_2\right) = \int_{u_1}^{u_2} du \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$$

$$u_1, u_2 < \infty$$

- TCL : CENTRE de la distribution
- Gaussienne : minimum d'information

Théorème central limite : Levy

- $L_\mu(x)$, symétrique
- x_i , $i = 1, \dots, N$, iid
- Somme $X_N = \sum_{i=1}^N x_i$

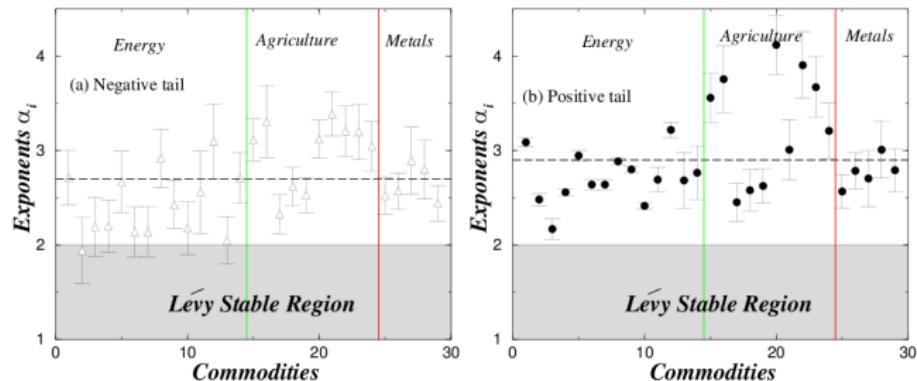
$$\hat{L}(k) = \exp(-a_\mu |k|^\mu) \begin{cases} \mu < 2 & L_\mu(x) \propto \frac{\mu A^\mu}{|x|^{1+\mu}}, \quad x \gg 1 \\ \mu = 2 & L_2(x) = \mathcal{N}(0, \sigma) \\ \mu > 2 & \text{pas une fct caractéristique} \end{cases}$$

- Distr. de Lévy : attracteurs de convolutions
- $\mu < 2$: variance ∞
- $\mu \leq 1$: moyenne non définie.
 - $\mu = 1$: loi de Cauchy $L_1(x) = \frac{A}{x^2 + \pi^2 A^2}$

Calibration

- maximum de vraisemblance pour l'exposant
 - Levy-stable ?
 - Gaussien ?

[Matia et al. (2002)]



Rendements : Levy ou Gauss ?

Levy dans marchés «extrêmes», i.e., exceptionnellement p. e. Citigroup.

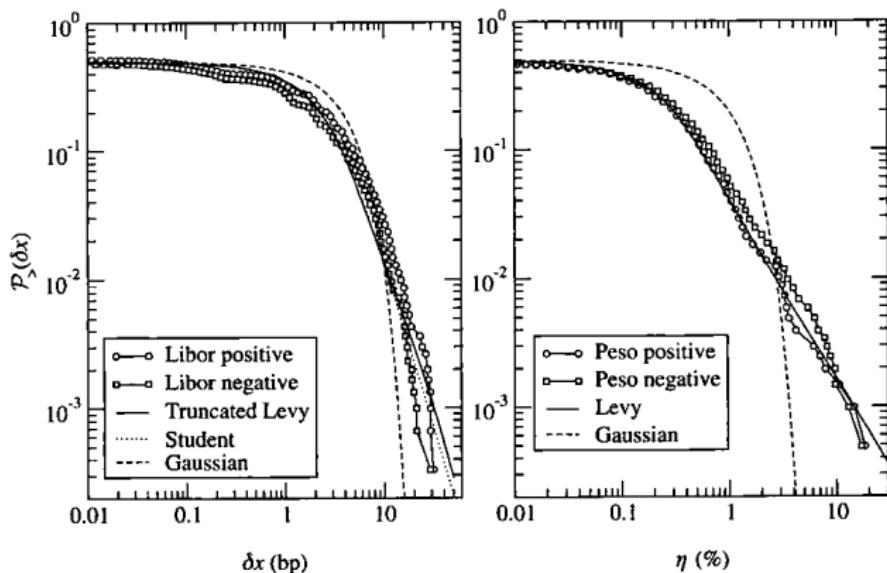
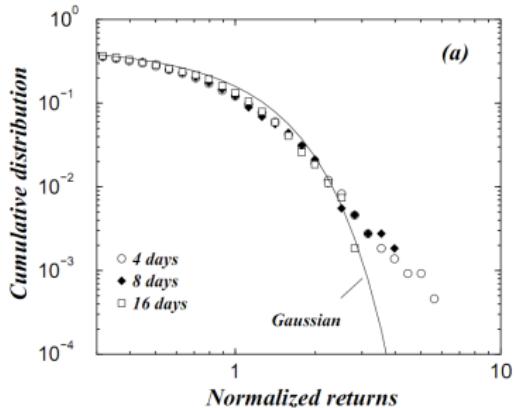
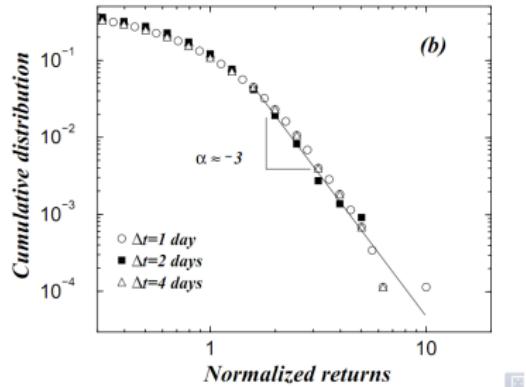
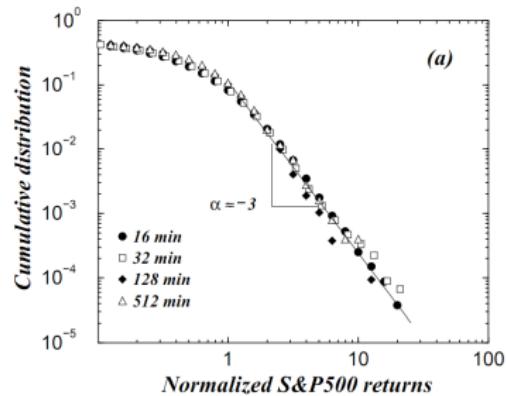


Fig. 6.11. Eurodollar front-month (i.e. USD Libor futures, left panel) and Mexican Peso (vs. USD, right panel) daily returns cumulative distributions. Note that neither the TLD nor the Student fit are very good for Libor. For the Mexican peso a pure Lévy distribution $\mu = 1.34$ works remarkably well

Convergence Student → Gauss



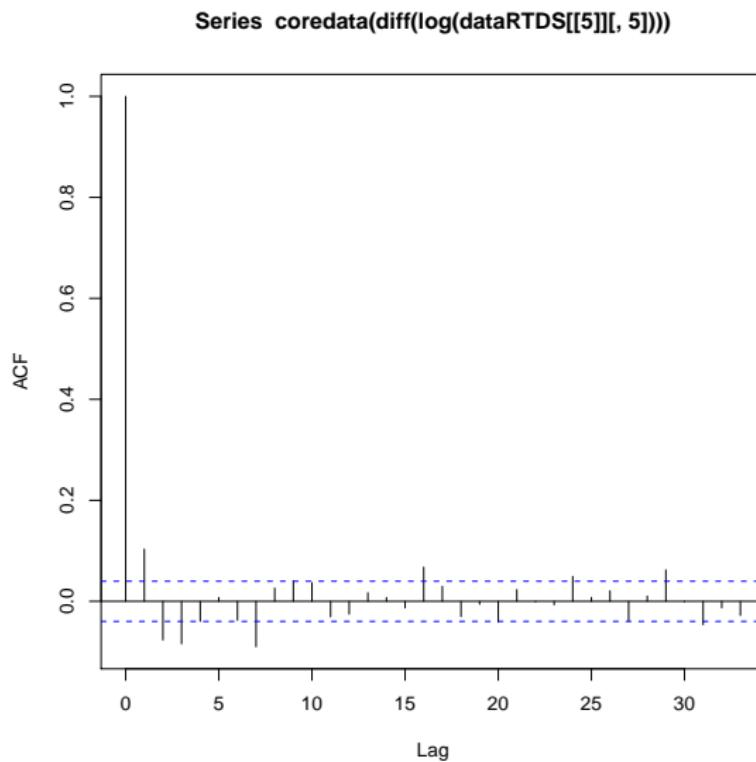
Vitesse de convergence
Taille de la région centrale

- Moments finis $\propto N^{2/3}$
- Student $\propto \sqrt{N \log N}$

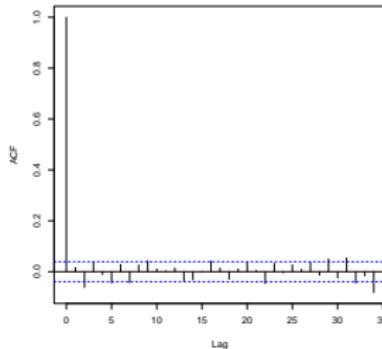
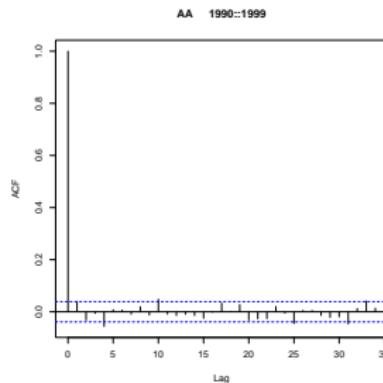
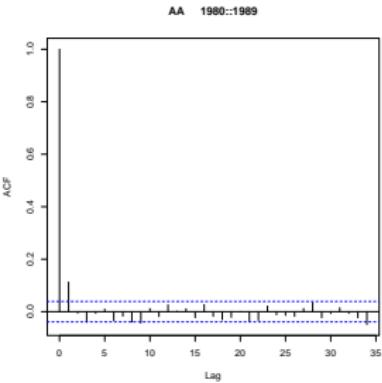
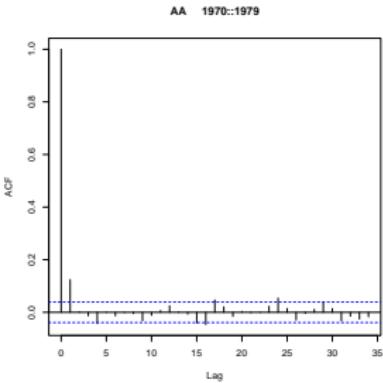
Problem : $P_>(x)$ only asymptotically power-law

- Hill estimator : hyp $P_>(x) = \frac{x_{\min}^{\mu-1}}{x^\mu}, \quad x \geq x_{\min}$
 1. choose starting point x_{\min}
 2. for all $x_i \geq x_{\min}$ compute $\frac{1}{\hat{\mu}} = \frac{1}{N} \sum_i \log(x_i/x_{\min})$
- How to choose x_{\min} ?
→ Shalizi & Newman (2007) [link]
 1. fit a power-law for all values of $x_{\min} \in \{x_i\}$
 2. x_{\min} minimises cost function (e.g. K-S statistics)

Traditionally : no price return autocorrelation.

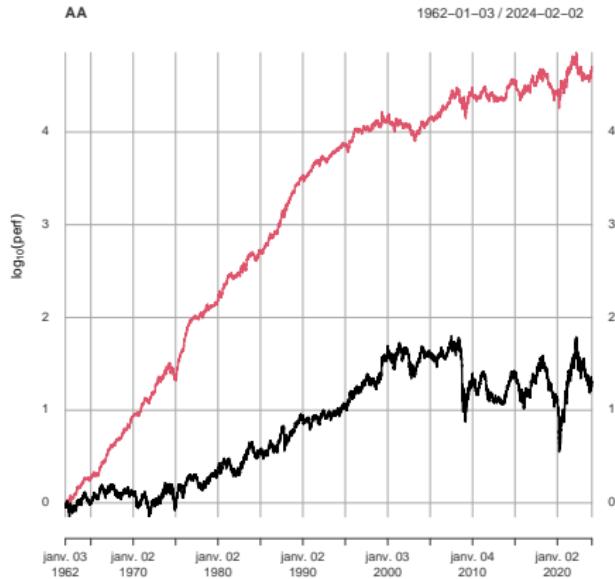


This course : dynamics of predictability

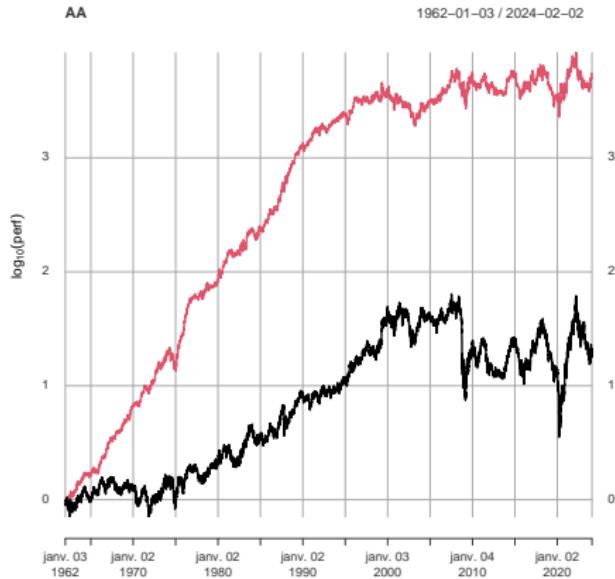


Long if past return > 0 : Alcoa

transaction cost 0



3bps (0.03%)



Long if past return > 0 : Bitcoin

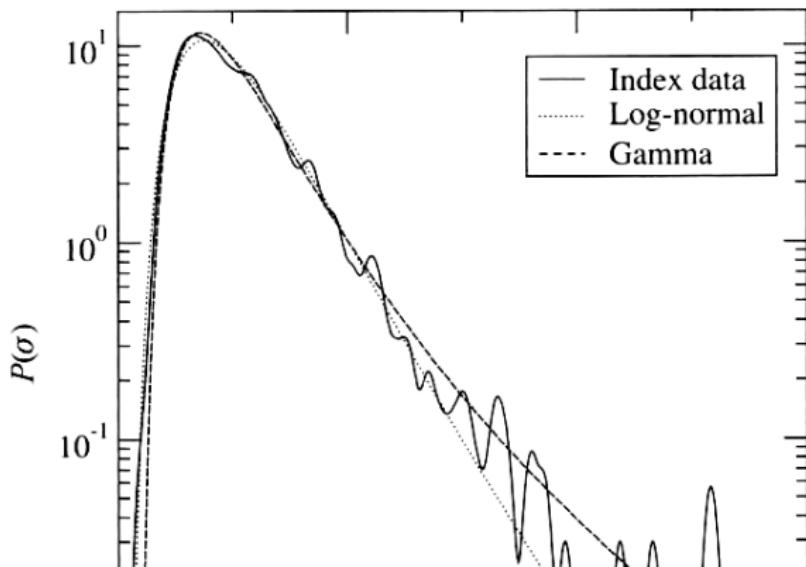


Volatilité : distribution

De Bouchaud et Potters (2003), S&P500, 5-min

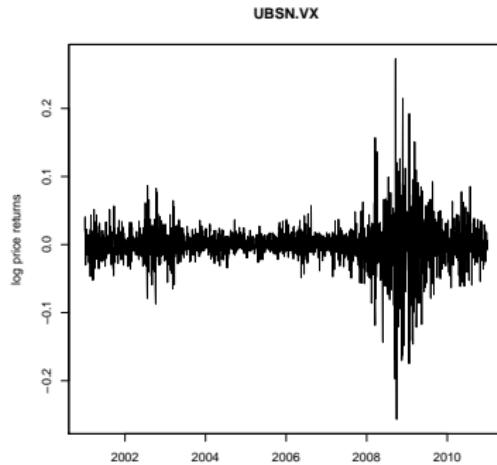
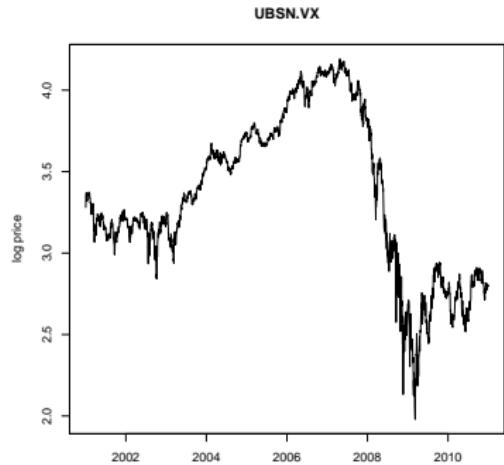
- Volatilité à haute fréquence

$$\sigma_{hf} = \frac{1}{N_{5\text{-min}}} \sum_k \langle |r_{5\text{-min}}(k)| \rangle$$



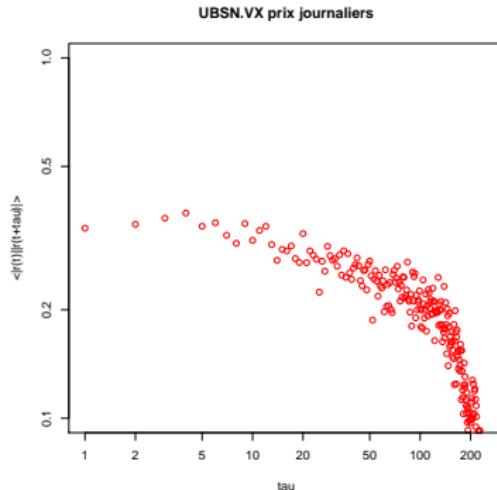
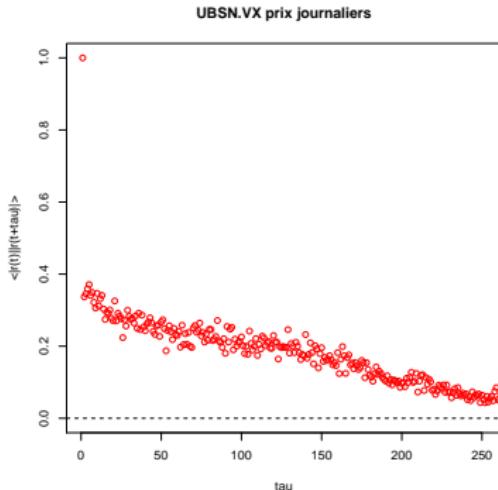
Volatilité : autocorrelation

“Volatility clustering”



Volatilité : autocorrelation

“Volatility clustering” : décroissante TRÈS LENTE de l'autocorrélation de $|r|$



Volatilité à très haute fréquence : signature plot

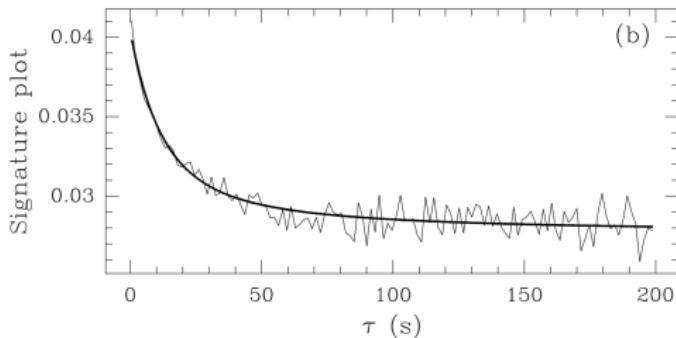
Prix milieu

$$m(t) = \frac{b(t) + a(t)}{2}$$

Origines des fluctuations

1. changement d'offres
2. transactions

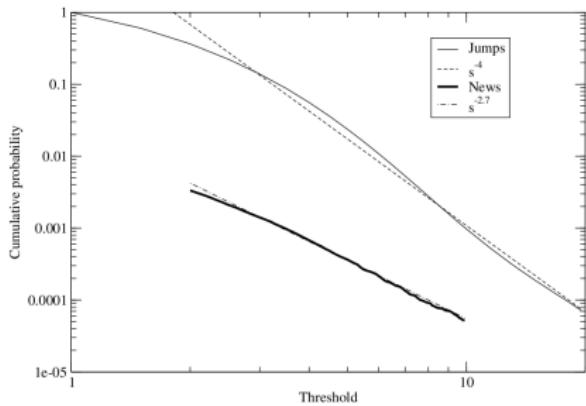
tiré de Bacry et al. (2010) (Bund)



Fluctuations intraday

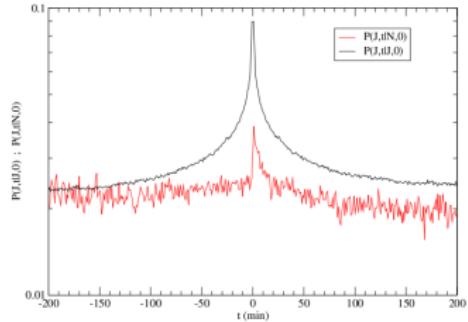
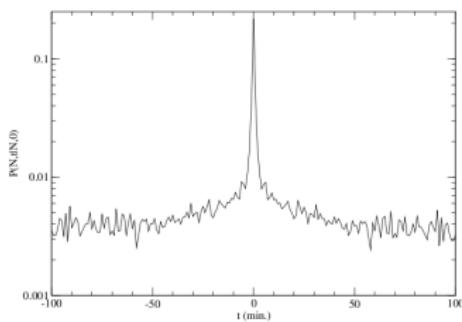
Joulin *et al.* (2008) :

- 1-min return $r(t)$,
- $m(t) = \frac{1}{120} \sum_{t'=t-120}^t |r(t')|$
- Saut si
 $|r(t)| > sm(t-1)$
- $P(\#jumps|s) \sim s^{-4}$
- 166 stocks during 149 days
- 93,698 news, 1 every 5 days per stock
- $s = 4$:
 - 177,674 jumps
 - 8 per day per stock



Jumps : exogenous vs endogenous

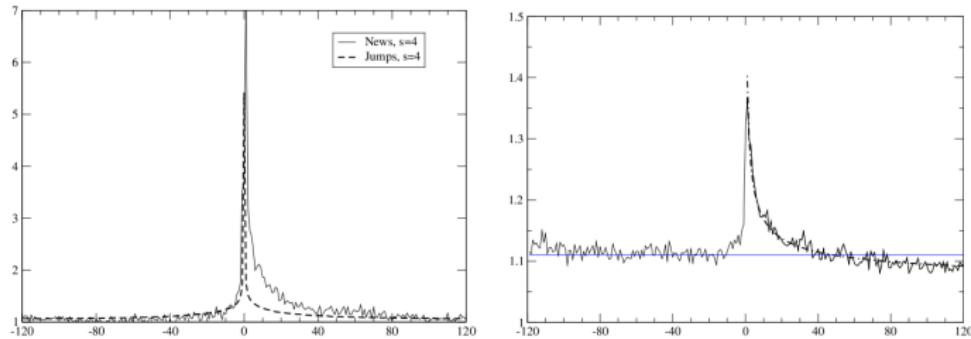
Event clustering



After news, prob jump smaller

Jumps : exogenous vs endogenous

Volatility decay :



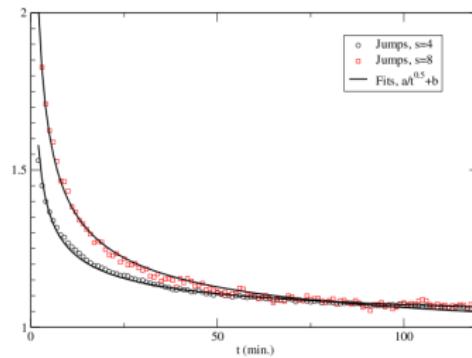
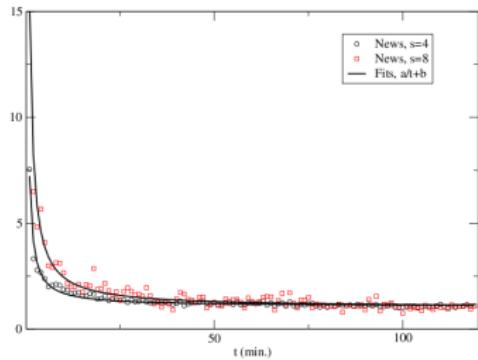
After average news, volatility decreases below previous baseline

Fluctuations

Most intraday
large price changes
not due to news.

Jumps : exogeneous vs endogeneous

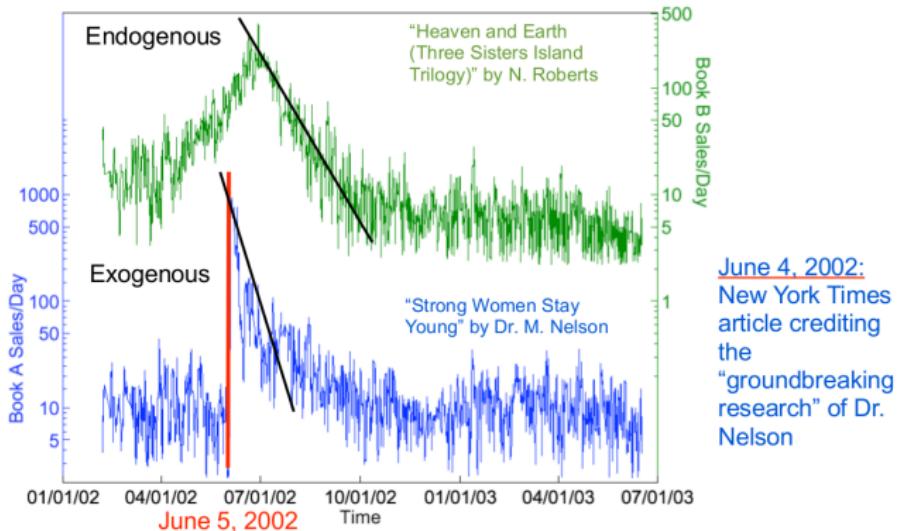
Volatility decay :



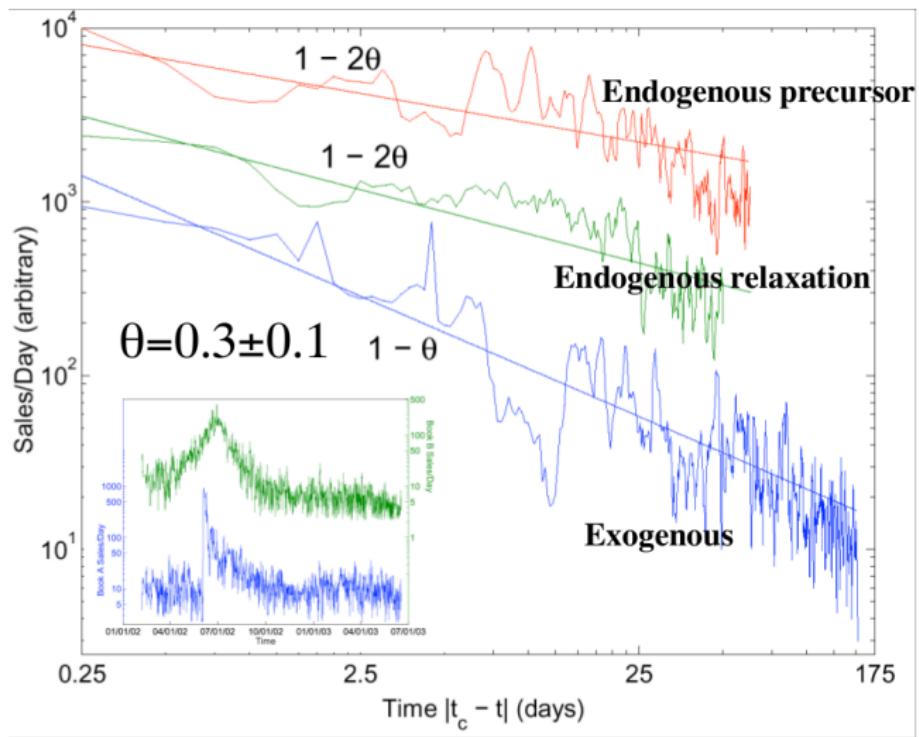
SLOW DECAY

Amazon

Taken from Sornette



Amazon



Fluctuations

Excess volatility : Shiller (2000)

Fundamental vol = vol related to dividends



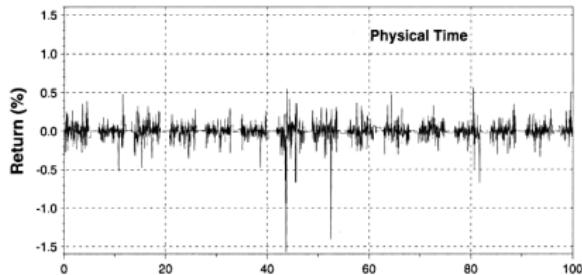
Figure 30.2 Excess volatility: ratio of price to fundamental volatility.

Source: DrKW Macro research.

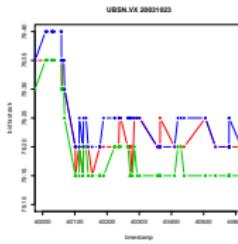
$p(t)$: temps physique

Trois problèmes :

1. Meso : saisonnalités



2. Micro : événements discrets



3. Rendements non gaussiens

Temps intrinsèques : subordination

- événements du carnet
 - déposition
 - annulation
 - transaction
- temps discret n : subordonné à ensemble d'événements
 - transactions
 - modifications des meilleures offres
- “business time” : moyenne mobile du nombre de ticks
- transformation $p(t) = p[t(n)]$: SUBORDINATION

Subordination : business time

Dacorogna et al (1993), Dacorogna et al (2001)

- FX : somme des activités (tics) des USA, JPY et EUR

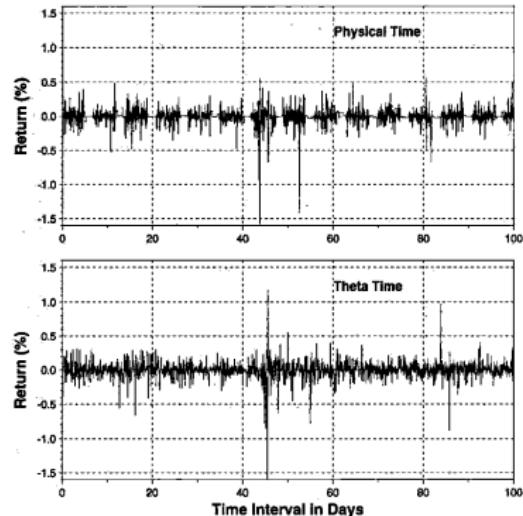
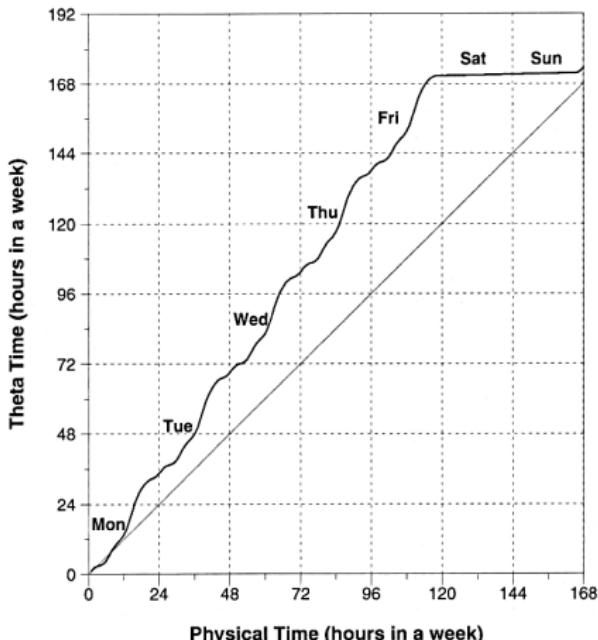


FIGURE 6.7 The hourly returns for USD-DEM from June 3, 1996, 00:00:00 to September 11, 1996, 00:00:00 are plotted using the physical time scale and the ϑ -time scale. Note also the extreme events that are clearly visible on both graphs.

Subordination : volatilité ou nombre de tics

- Dynamique du prix

$$P(r_{\Delta t}) \propto \exp\left(-\frac{r_{\Delta t}^2}{2\sigma\Delta t}\right)$$

- Volatility stochastique $\rightarrow P(\sigma) \rightarrow$ rendements non gaussiens (cf. Clark 1970, 1973)

$$P(r_{\Delta t}) = \int_0^\infty d\sigma P(\sigma) \mathcal{N}(r_{\Delta t}, \sigma)$$

- Directement accessible : nombre de tics
- hypothèse : variance \propto # transactions
- hypothèse equiv : 1 transaction \rightarrow rendement gaussien

$$P(r_{\Delta t}) = \int_0^\infty dN_{\Delta t} P(N_{\Delta t}) \mathcal{N}(r_{\Delta t}, N_{\Delta t})$$

Distribution du nombre de tics : Yakovenko

Silva and Yakovenko, Stochastic volatility of financial markets as the fluctuating rate of trading : an empirical study (2006)

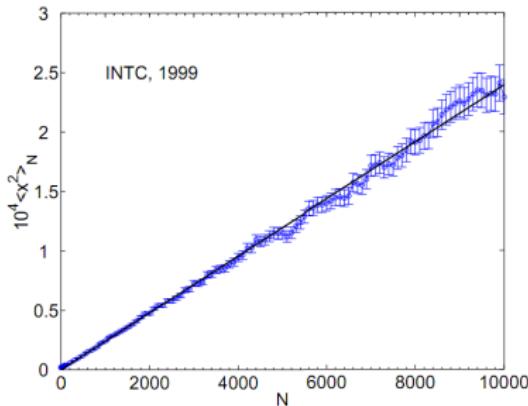


FIG. 1: The variance of log-returns $\langle x^2 \rangle_N$ after N trades plotted vs. N .

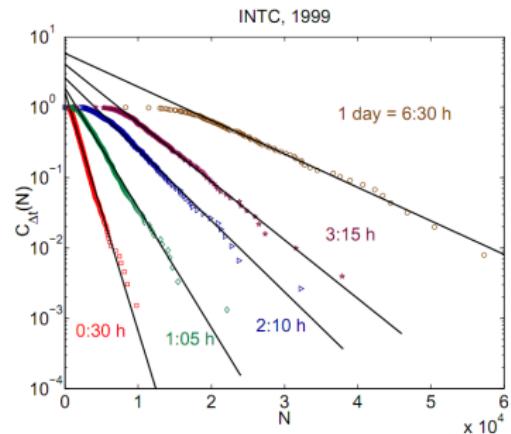


FIG. 6: Empirical cumulative distribution $C_{\Delta t}(N)$ for the number of trades N during the time interval Δt , plotted in the log-linear scale. The solid lines illustrate exponential tails.

Distribution du nombre de tics : Yakovenko

Si $P(N_{\Delta t})$ est exponentielle

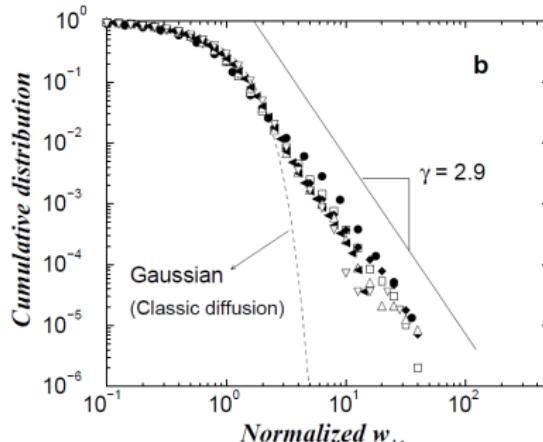
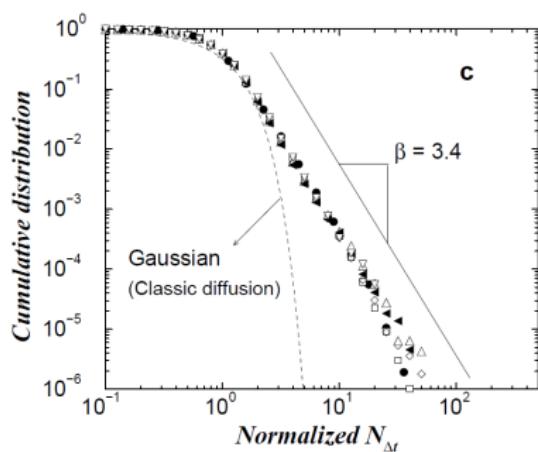
$$\begin{aligned} P(r_{\Delta t}) &= \int_0^{\infty} dN_{\Delta t} P(N_{\Delta t}) \mathcal{N}(r_{\Delta t}, N_{\Delta t}) \\ &= \frac{1}{\sqrt{2\theta\Delta t}} \exp\left(-|r_{\Delta t}| \sqrt{\frac{2}{\theta\Delta t}}\right) \end{aligned}$$

Distribution du nombre de tics : Stanley

Plerou et al, Economic Fluctuations and Diffusion (1999)

- $\Delta_t = 30$ mins, 1000 titres
- $w_{\Delta t}$: volatilité (locale) par transaction

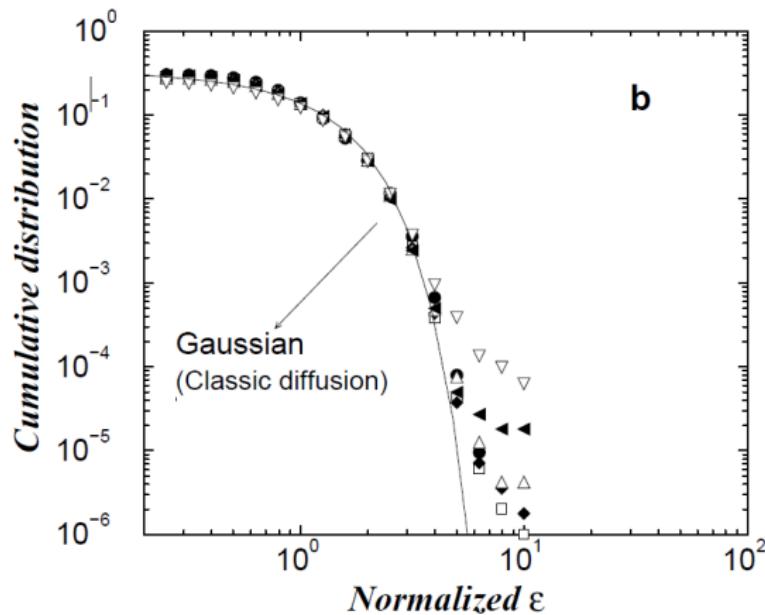
$$|r_{\Delta t}(t)|^2 = N_{\Delta t}(t)w_{\Delta t}(t)^2$$



Subordination : Stanley

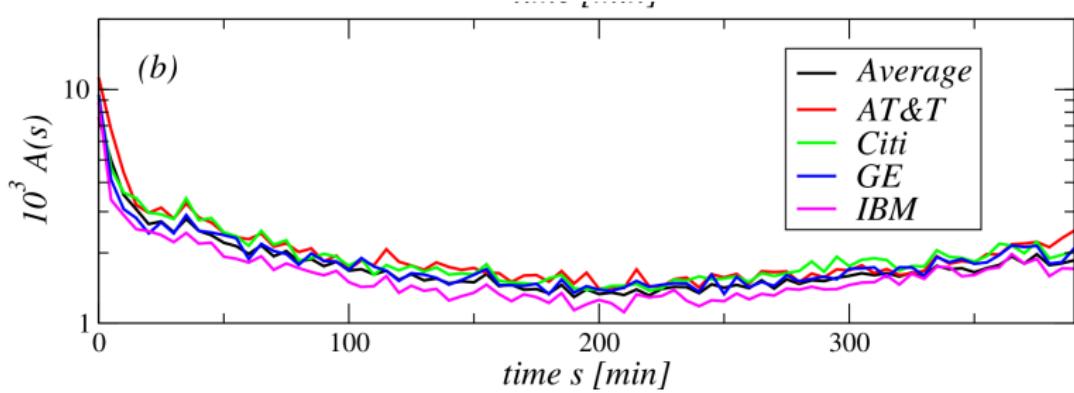
Réchelonnage :

$$|r_{\Delta t}|^2 = N_{\Delta t} w_{\Delta t}^2 \rightarrow \epsilon_{\Delta t}(t) = r_{\Delta t}(t) / [w_{\Delta t}(t) \sqrt{N_{\Delta t}(t)}]$$



Intraday U-pattern

E.g. Wang et al (2005)

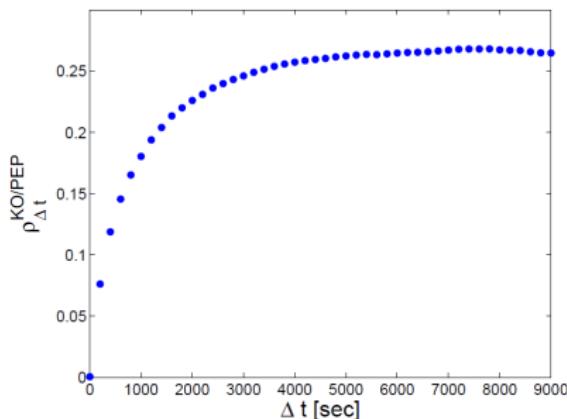


Correlations : Epps

Correlation entre deux titres i et j

$$C_{ij}(\Delta t) = \frac{\Delta t}{T} \sum_{t=1}^{T/\Delta t} r_{i,\tau}(t)r_{j,\tau}(t)$$

- Epps (1979) : $C_{ij}(\tau)$ croît avec τ en intraday



(tiré de Toth et Kertesz 2008)

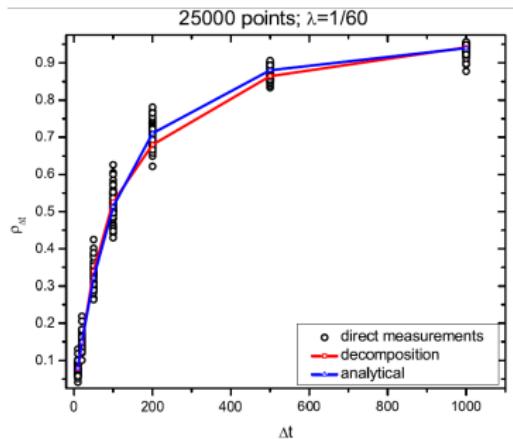
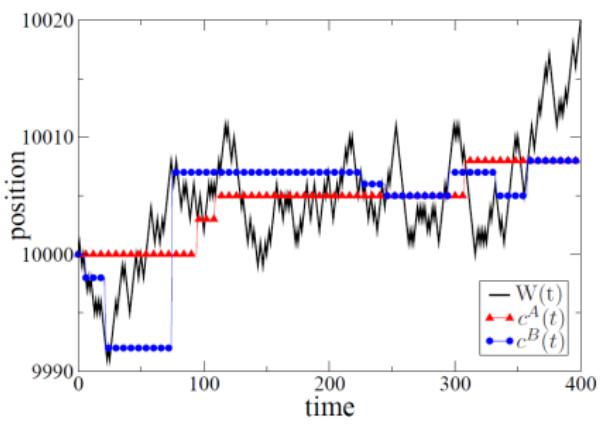
Causes : asynchronicité des rendements, lead-lag

Littérature : A. Lo and C. MacKinlay (1990), Toth and Kertesz (2008), Marsili et al (2010), etc.

Epps : asynchronicité

Ex : Toth et Kertesz (2008)

Une seule marche aléatoire, 2 échantillonages aléatoires $\lambda = 1/60$



$$\rho(\lambda) = 1 - \frac{1}{\lambda \Delta t} [1 - \exp(-\lambda \Delta t)]$$

Hayashi-Yoshida

Takaki Hayashi, Nakahiro Yoshida, On covariance estimation of non-synchronously observed diffusion processes, (2005)

- Causalité : Extension d'un rendement de prix sur l'intervalle de temps précédent
- Estimateur de covariance cumulée : intervalles avec recouvrement

$$\hat{C}_{HY} = \sum_i^{n_1} \sum_j^{n_2} (p_{1,t_{1,i}} - p_{1,t_{1,i-1}})(p_{2,t_{2,j}} - p_{2,t_{2,j-1}}) K_{i,j}$$

$$K_{i,j} = \begin{cases} 1 & \text{if } \max(t_{1,i-1}, t_{2,j-1}) < \min(t_{1,i}, t_{2,j}) \\ 0 & \text{otherwise} \end{cases}$$

$$\lim_{\Delta t \rightarrow 0} \hat{C}_{HY} = C \text{ pour des rendements gaussiens}$$

Covariance/correlation : lead-lag

Jusqu'à présent : correlation à temps identique. Lead-lag :
position du maximum du pic
Toth et Kertesz (2008)

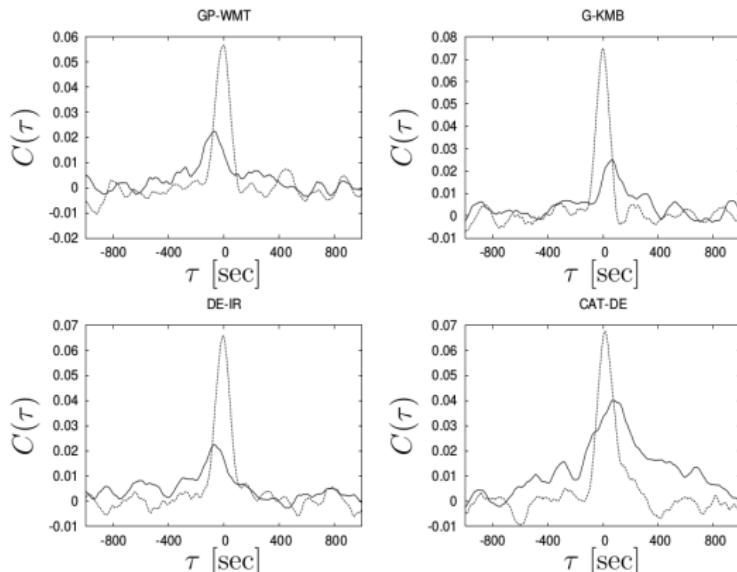


Fig. 2. Example plots showing the changes in the correlation functions. The functions are plotted for the years 1993 (solid) and 2003 (dashed). The peaks have become much higher and the maximum positions have moved towards zero.

lead-lag 2 :

Toth et Kertesz (2008)

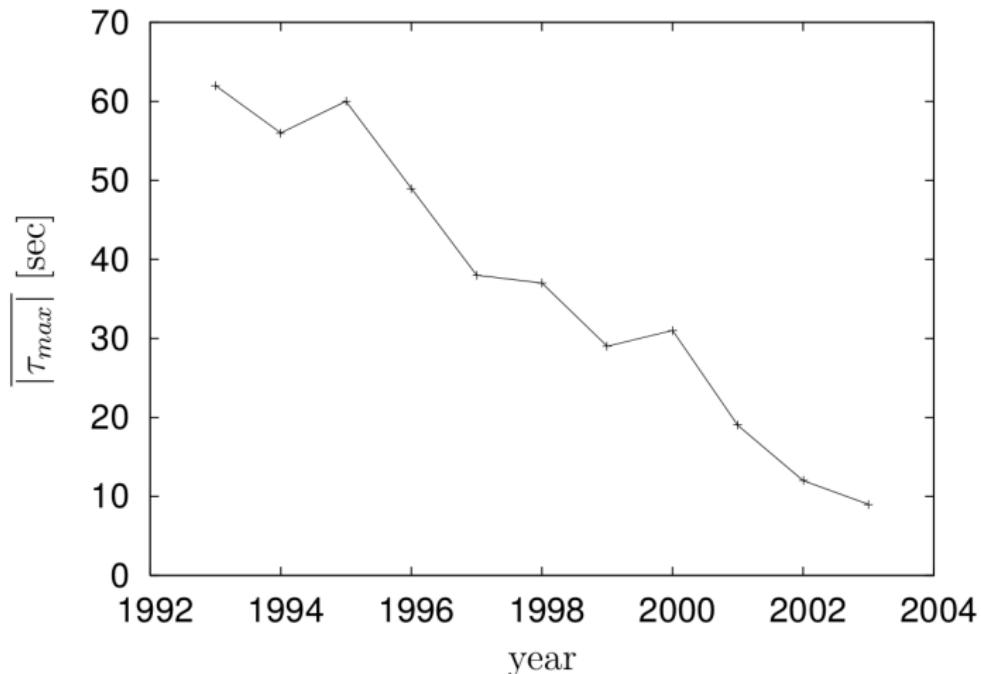


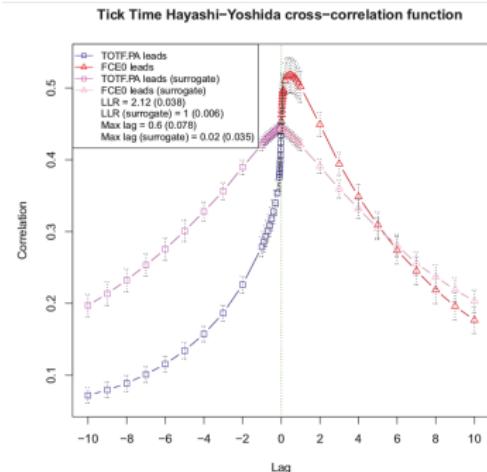
Fig. 3. The average time shift of the correlation functions taken over every pair examined, as a function of time.

Autre mesure de lead-lag 1

Huth et Abergel (2014)

lead – lag : asymétrie de la fonction de corrélation vs lag

$$LL := \frac{\int_0^T C(\tau)^2 d\tau}{\int_0^T C(-\tau)^2 d\tau}$$

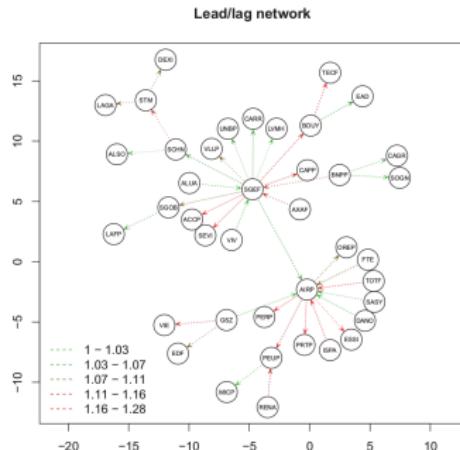


Autre mesure de lead-lag 1

Huth et Abergel (2014)

lead – lag : asymétrie de la fonction de corrélation vs lag

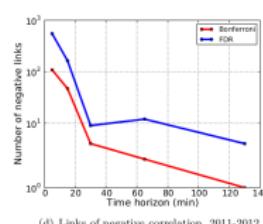
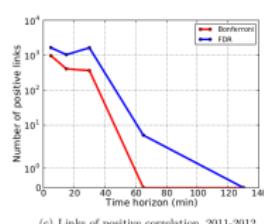
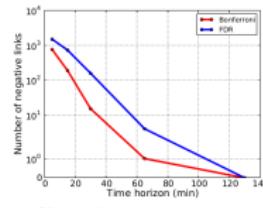
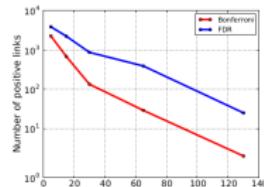
$$LL := \frac{\int_0^T C(\tau)^2 d\tau}{\int_0^T C(-\tau)^2 d\tau}$$



Autre mesure de lead-lag 2

Curme et al (2015)

- Hypothèse zero : pas de corrélation
- Estimateurs de lead-lag : complexes → p-valeur ?
- Plus précisément : statistique de permutation, déterminer p-valeur d'un estimateur de lead-lag.
→ casser une partie de la structure des rendements :



Lead-lag : prediction ?

Challet, Pelletier, Bongiorno (2021) [link]

Prediction networks (daily data)

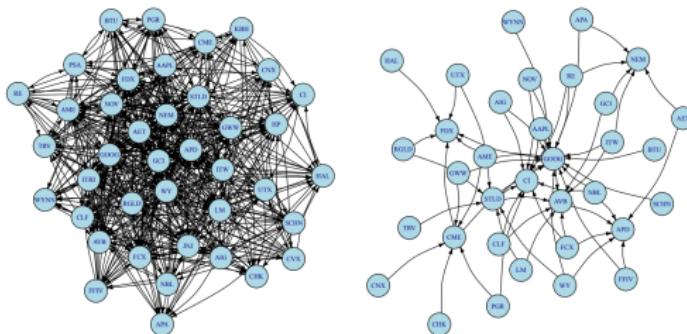


Figure 3: Examples of prediction network inferred with knockoffs. $T_{in} = 300$ trading days, period ending on 2006-03-13. Left plot: FDR= 0.3 and 100 runs per asset; right plot: FDR= 0.2 and 10 runs per asset. Selection via random forest variable importance.

Lead-lag : prediction ?

Challet, Pelletier, Bongiorno (2021) [link]

Use predicted returns in mean-variance portfolios

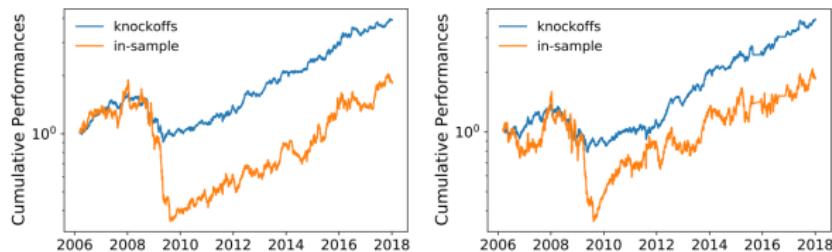


Figure 6: Cumulative performance from prediction networks with mean-variance portfolios with target daily return equal to 0.005, blue curve uses the knockoff prediction for the returns, the orange curve uses the average in-sample returns over the past 300 days. Left plot: FDR= 0.3; right plot FDR= 0.2.

Leverage effect

Corrélation prix-volatilité

$$\mathcal{L}_S(\tau) = \frac{\langle r(t)(r(t+\tau))^2 \rangle}{\langle r(t)^2 \rangle^2} \simeq -A_S \exp(-\tau/T_s)$$

Bouchaud et al., The leverage effect in financial markets : retarded volatility and market panic (2001)

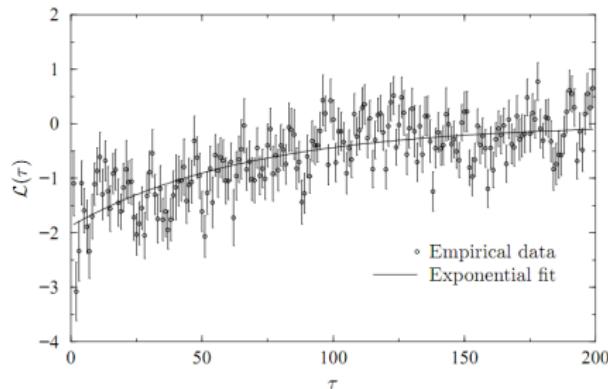


Figure 1: Return-Volatility correlation for individual stocks. Data points are the empirical correlation averaged over 437 US stocks, the error bars are two sigma errors bars estimated from the inter-stock variability. The full line shows an exponential fit (Eq. (3)) with $A_S = 1.9$ and $T_S = 69$ days. Note that $\mathcal{L}(\tau = 0)$ is

Leverage effect

Bouchaud et al., The leverage effect in financial markets : retarded volatility and market panic (2001)

$$\mathcal{L}_I(\tau) = \frac{\langle r(t)(r(t+\tau))^2 \rangle}{\langle r(t)^2 \rangle^2} \simeq -A_I \exp(-\tau/T_I)$$

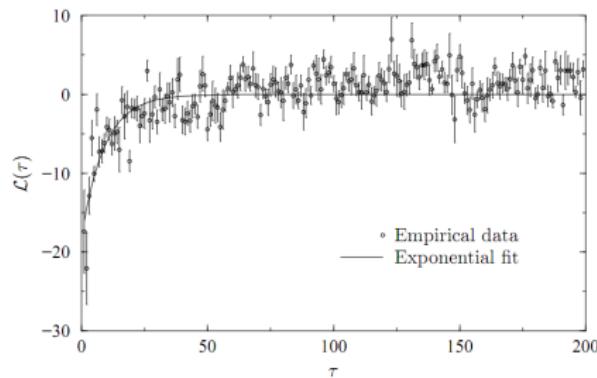


Figure 2: Return-Volatility correlation for stock indices. Data points are the empirical correlation averaged over 7 major stock indices, the error bars are two sigma errors bars estimated from the inter-index variability. The full line shows an exponential fit (Eq. (3)) with $A_I = 18$ and $T_I = 9.3$ days.

Résumé

1. Rendements

- non corrélés
- queues grasses
- convergence vers Gauss

2. Volatilité

- corrélation à très long terme
- queue grasse

3. Subordination

4. Corrélations : effet d'Epps

- asynchronicité
- lead-lag