Microstructure of Extreme Events

Fabrizio Lillo

Scuola Normale Superiore di Pisa, University of Palermo (Italy) and Santa Fe Institute (USA)

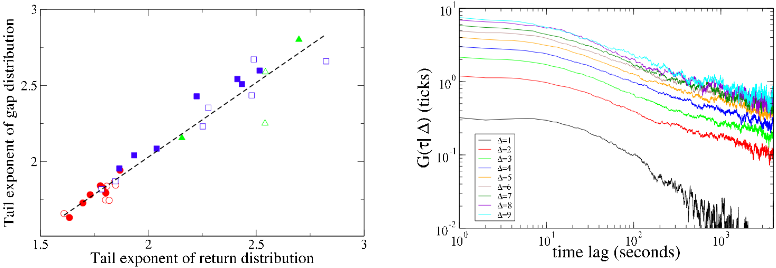
Perm - February 6, 2013

Motivations

- Financial markets are intrinsically unstable and display large price fluctuations
- Often these fluctuations occur on very short time scales, and sometimes revert quickly.
- In recent year there has been an increasing frequency of these events
- Neither idiosyncratic news nor market wide news can explain the frequency and amplitude of price jumps (Bouchaud et al. 2008)
- Stock price jump: the difficulty of a definition

Stock price jumps and liquidity fluctuations

- Large transaction volumes are not responsible for large price jumps (Farmer et al. 2004, Bouchaud et al 2008)
- On microstructural time scales, liquidity fluctuations are responsible of large price jumps
- The spread and other limit order book quantities decay on average to the normal value by following a very slow dynamics



Left. Tail exponent of return vs tail exponent of gap between best and second best (Farmer et al 2004). Right. Dynamics of average excess spread conditional to a spread increase of Δ ticks (Ponzi et al 2009).

Stock price jumps: which time scale?

- Microshocks: To qualify as a down(up)-draft candidate, the stock had to tick down (up) at least 10 times before ticking up (down)– all within 1.5 seconds and the price change had to exceed 0.8% (Nanex)
- Mesoshocks: We choose to compare the absolute size $|r(t)|$ of a one minute bin return to a short term (120 minutes) flat moving average $m(t)$ of the same quantity, in order to factor in slow modulations of the average volatility. An s-jump is such that $|r(t)| > sm(t)$ (Bouchaud et al. 2008).

Are microshocks “embedded” in mesoshocks?

Table 2. Micro shocks embedded into meso shocks

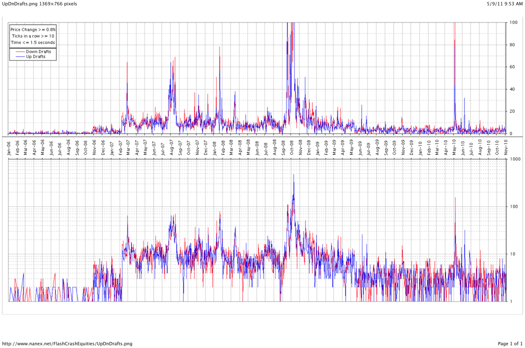
Event type	Meso events quantity	Time bounds	Micro events in time bounds	Total micro events quantity	Share of embedded events
Up-draft	3354	±1 min	42	157	26%
Down-draft	3406	±1 min	44	212	21%

with M. Frolova and S. Ivliev (2012, submitted)

Stock price jumps and regulation

It has been recently suggested (Nanex) that

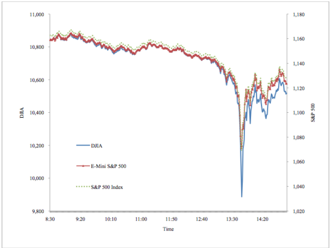
- There has been a sharp increase of price jump frequency
- This is related to changes in market regulation



Dynamics of the number of price shocks detected (2006-2010) in a large universe of US stocks

Regulation NMS, Hybrid Phase IV - completed rollout February 27, 2007.

- There is growing concern about the possible role of High Frequency Trader as responsible of large price jumps
- Debate partly driven by the Flash Crash of May 6, 2010
- “when rebalancing their positions, High Frequency Traders may compete for liquidity and amplify price volatility” (Kirilenko et al. 2010)

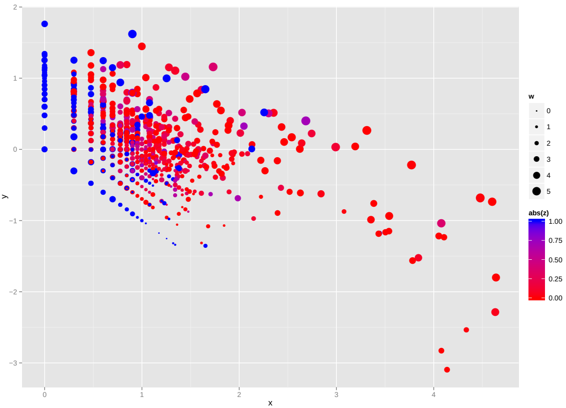


This figure presents end-of-minute transaction prices of the Dow Jones Industrial Average (DJIA), S&P 500 Index, and the June 2010 E-Mini S&P 500 futures contract on May 6, 2010 between 8:30 and 15:15 CT.

Market shocks: should High Frequency Traders be blamed? (with M. Frolova and S. Ivliev)

- Agent resolved data of stocks traded in the Russian stock market
- We present here the results of the investigation of the HFTs behavior when trading a stock traded in the Russian stock exchange
- We identified 22 8-sigma down shocks and 27 up-shocks
- We focus our analysis on a 4 hour window around each shock
- The number of active IDs (trading account) is 44,240
- Information on individual orders (not only executed ones)

Market shocks: should High Frequency Traders be blamed? (with M. Frolova and S. Ivliev)



Each point is an agent. x-axis is the decimal logarithm of the number of limit orders, y-axis is the decimal logarithm of the fraction of executed limit orders. Size of points is proportional to $\log_{10}(\textit{traded volume})$ and color codes the fractional position accumulated at the end of the day.

Market shocks: should High Frequency Traders be blamed? (with M. Frolova and S. Ivliev)

- We consider one minute subintervals
- We focus our attention on a four minute interval around the shock (t=0)
- Given a variable X, for each shock s and each active HFT k, we compute the quantity $sgn(X_t^{k,s} - X_{t-1}^{k,s})$ and we test the null hypothesis that it is binomial distributed with $p = 0.5$ for each t.

Variable	t = -1	t = 0	t = +1
Trading volume	-	↑**	↓**
Net inventory	-	↑**	↓**
Limit order arrival rate	-	↑**	↓**
Cancellation rate	-	↑**	-
Cancellation/limit order ratio	-	↓**	↑**
Weighted distance of LO from best	-	↑**	-
Limit order rate wrt the market	-	↓*	-
Cancellation rate wrt the market	-	-	-

Result of a statistical test searching for change of trading behavior of 9 HFTs in a 2 min interval around 49 price jumps.

Market shocks: should High Frequency Traders be blamed? (with M. Frolova and S. Ivliev)

- At the minute of the down shock, HFTs
  - trade more (hot potato)
  - buy
  - increase limit order and cancellation rate, but the cancellation/limit order ratio decreases and relative to the market the rates remain constant
  - place order deeper in the book
- The next minute, they revert most of these variations
- No sign of significant change at the minute before (probably one minute is too long)
- Similar results for up shocks

Stock price jumps and systemic events

- Flash crash was not only about E-Mini S&P 500 futures (where everything started)
- Because of arbitration, when the e-Mini changes price with high volume,
  - Many ETFs are repriced (quotes updated, trades executed).
  - The component stocks of ETFs are also repriced, along with many indexes.
  - All the option chains for the ETFs, their components and indexes are also repriced.
- Trades being executed at irrational prices as low as one penny or as high as $ 100,000
- Flash crash: a 20 millesecond cascade (Nanex)
- HFT as means of contagion (Gerig 2012)
- Motivation for our study: how frequent are systemic events? How can we model them?

This paper: stock price jumps in a systemic context

An econometric approach to detection and modeling of the dynamics of stock price jumps in a set of stocks

- Stock price jumps and sensitivity to the detection method
- A point process approach to stock price jumps $\rightarrow$ Few events
    - Time clustering of jumps
    - Cross stocks excitation of jumps
    - Co-jumps: speed of contagion
    - Systemically relevant jumps

Bormetti et al., Modelling systemic cojumps with Hawkes factor models, arXiv:1301.6141

This project has been developed in QuantLab:

G. Bormetti
L.M. Calcagnile
F. Corsi
S. Marmi

SCUOLA NORMALE SUPERIORE PISA

M. Treccani

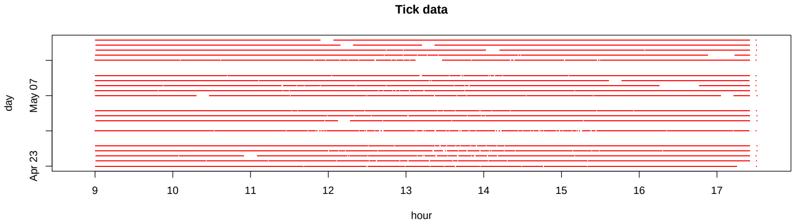
LIST

QUANT Lab

QuantLab is a joint Research Lab between SNS & List Spa.
See www.quantlab.it

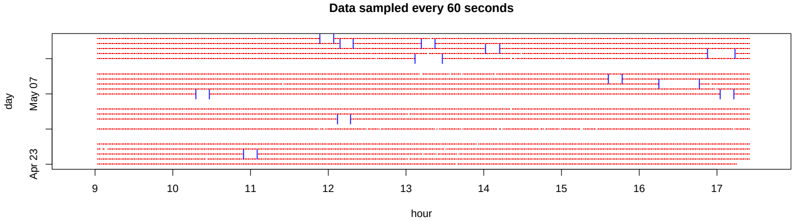
Data Sample

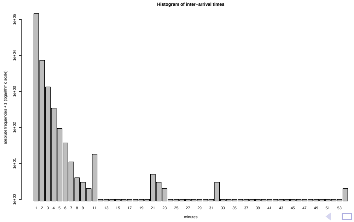
- Tick by Tick Data of FTSE 40 Italy, traded at Milan stock exchange
- 20 high liquidity stocks
- 88 Days of Executions (March-June 2012)
- One minute time scale
- Log Return of the Prices



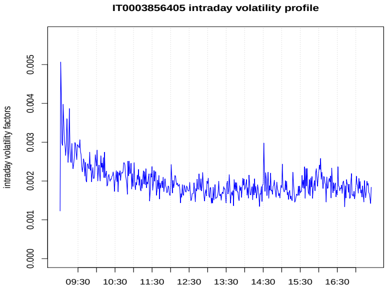
Data cleaning

- Jump detection methods are obviously very sensitive to outliers, data anomalies, etc.
- Outlier removal: Brownlees - Gallo (2006) algorithm
- Merging / Splitting and Volatility auctions are automatically subtracted





Removing intraday volatility pattern



$$\tilde{r}_{d,t} = \frac{r_{d,t}}{\lambda_t}, \qquad \lambda_t = \frac{1}{N_{\text{days}}} \sum_d \frac{|r_{d,t}|}{s_d},$$

where t is the index of intraday times, d is the index of days and s_d is the standard deviation of absolute returns of day d.

- Most of the jump identification methods are essentially estimating the local volatility and then testing that the ratio between absolute return and volatility is above a given threshold
- This approach implicitly requires the definition of a time scale used to compute returns (in our case, one minute)
- We define
  Jump:

$$\frac{|r|}{\sigma} > \theta \quad \text{(we choose } \theta = 4 \text{ as in Bouchaud } \textit{et al.}\text{, 2008)}$$

- The volatility estimation is the most critical step and differentiates methods
- We tested other approaches (such as wavelets), but without clear results

Asymptotic results:

$$\mathrm{p}-\lim_{\delta\to 0}\delta^{\frac{1}{2}}\sum_{i=1}^{t/\delta}|r_i(t)| = \mu_1\int_0^t \sigma(s)\,\mathrm{d}s$$

$$\mathrm{p}-\lim_{\delta\to 0}\sum_{i=1}^{t/\delta-1}|r_i(t)||r_{i+1}(t)| = \mu_1^2\int_0^t \sigma^2(s)\,\mathrm{d}s,$$

where $\mu_1 = \mathrm{E}(|u|) = \sqrt{\frac{2}{\pi}} \simeq 0.797885$, $u \sim \mathcal{N}(0,1)$.

Estimators:

$$\hat{\sigma}_{\mathrm{abs},t} = \mu_1^{-1}\overline{|r|} = \mu_1^{-1}\alpha\sum_{i>0}(1-\alpha)^{i-1}|r_{t-i}|$$

$$\hat{\sigma}^2_{\mathrm{bv},t} = \mu_1^{-2}\overline{|r_i||r_{i+1}|} = \mu_1^{-2}\alpha\sum_{i>0}(1-\alpha)^{i-1}|r_{t-i}||r_{t-i-1}|,$$

with $\alpha = 0.032$ (gives 50% of weight to the closest 22 observations).

Barndorff-Nielsen and Shephard (2003, 2004), Bouchaud et al. (2008)

Long series of zero returns can generate automatically (and spuriously) jumps.
How to handle (long) series of zero returns?
Three possibilities:

1. take last price and introduce zero returns
2. leave returns as *not available* and take the over-gap return
3. leave returns as *not available* and rescale the over-gap return with the square root of the number of lacking observations

Example

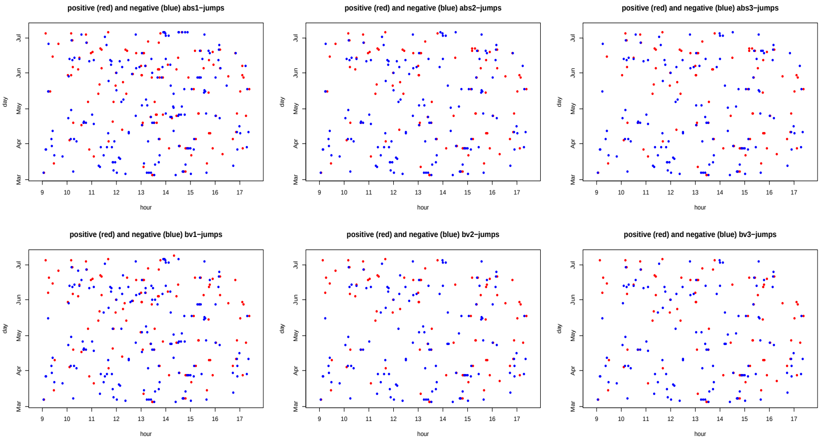
Price series:

12:00	12:01	12:02	12:03	12:04	12:05	12:06	12:07
p_0	p_1	–	–	–	–	p_6	p_7

Return series:

1. $\log\frac{p_1}{p_0},\ 0,\ 0,\ 0,\ 0,\ \log\frac{p_6}{p_1}, \log\frac{p_7}{p_6}$
2. $\log\frac{p_1}{p_0}, \mathtt{NA}, \mathtt{NA}, \mathtt{NA}, \mathtt{NA},\ \log\frac{p_6}{p_1}, \log\frac{p_7}{p_6}$
3. $\log\frac{p_1}{p_0}, \mathtt{NA}, \mathtt{NA}, \mathtt{NA}, \mathtt{NA},\ \frac{1}{\sqrt{5}}\log\frac{p_6}{p_1}, \log\frac{p_7}{p_6}$

Finding jumps



Number of jumps:

	1	2	3	$\cap$
σ_{abs}	281	228	217	205
σ_{bv}	260	208	200	190
$\cap$	239	196	186	178

ISIN	jumps	jumps up	jumps down	single jumps	cojumps
IT0000062072	103	48 (47%)	55 (53%)	53 (51%)	50 (49%)
IT0000062957	63	29 (46%)	34 (54%)	38 (60%)	25 (40%)
IT0000064482	121	60 (50%)	61 (50%)	97 (80%)	24 (20%)
IT0000068525	93	46 (49%)	47 (51%)	56 (60%)	37 (40%)
IT0000072618	127	67 (53%)	60 (47%)	55 (43%)	72 (57%)
IT0001063210	59	28 (47%)	31 (53%)	44 (75%)	15 (25%)
IT0001334587	178	**73** (41%)	**105** (59%)	150 (84%)	28 (16%)
IT0001976403	123	61 (50%)	62 (50%)	76 (62%)	47 (38%)
IT0003128367	188	**81** (43%)	**107** (57%)	107 (57%)	81 (43%)
IT0003132476	155	**66** (43%)	**89** (57%)	95 (61%)	60 (39%)
IT0003487029	70	28 (40%)	42 (60%)	41 (59%)	29 (41%)
IT0003497168	129	74 (57%)	55 (43%)	79 (61%)	50 (39%)
IT0003856405	95	50 (53%)	45 (47%)	74 (78%)	21 (22%)
IT0004176001	74	41 (55%)	33 (45%)	46 (62%)	28 (38%)
IT0004231566	103	50 (49%)	53 (51%)	72 (70%)	31 (30%)
IT0004623051	115	**47** (41%)	**68** (59%)	85 (74%)	30 (26%)
IT0004644743	100	51 (51%)	49 (49%)	65 (65%)	35 (35%)
IT0004781412	118	**49** (42%)	**69** (58%)	57 (48%)	61 (52%)
LU0156801721	59	27 (46%)	32 (54%)	32 (54%)	27 (46%)
NL0000226223	86	39 (45%)	47 (55%)	51 (59%)	35 (41%)
total	2159	**1015**	**1144**		
average	108.0	50.8 (47%)	57.2 (53%)		

The true Log Price Y_t follows

$$Y_t = X_t + \varepsilon_t \qquad \varepsilon_t \in \mathcal{N}(0, 10^{-4})$$

with

$$X_t = \int_0^t \mu_s \mathrm{d}s + \int_0^t \sigma_s dW_s + \sum_{l=1}^{N_t} L_t$$

and Spot Vol following Geometric O-U Process:

$$\mathrm{d}\log\sigma_t^2 = -(0.682 + 0.10\log\sigma_t^2)\mathrm{d}t + 0.25\mathrm{d}W_{1,t} \qquad \mathrm{Corr}(W_t, W_{1,t}) = -0.62$$

It has been harmonized with real data:

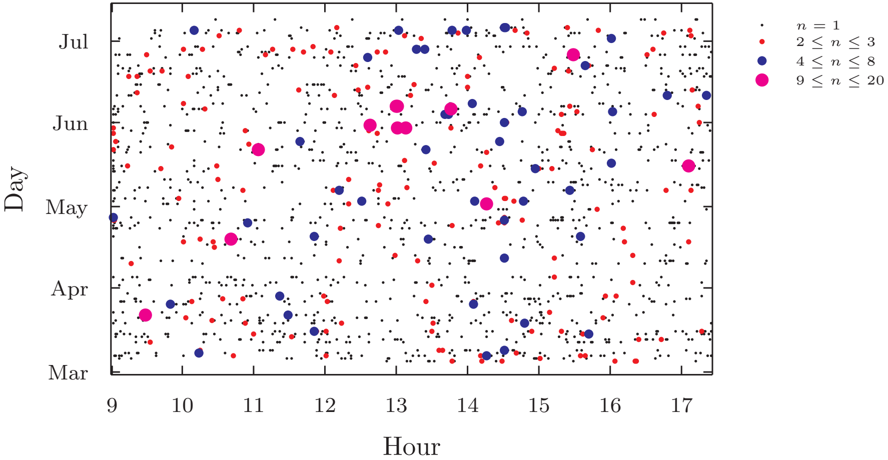
- Forced to have 3 daily jumps with intensity 5 times Spot Vol
- Introduced seasonality

False Negatives & False Positives

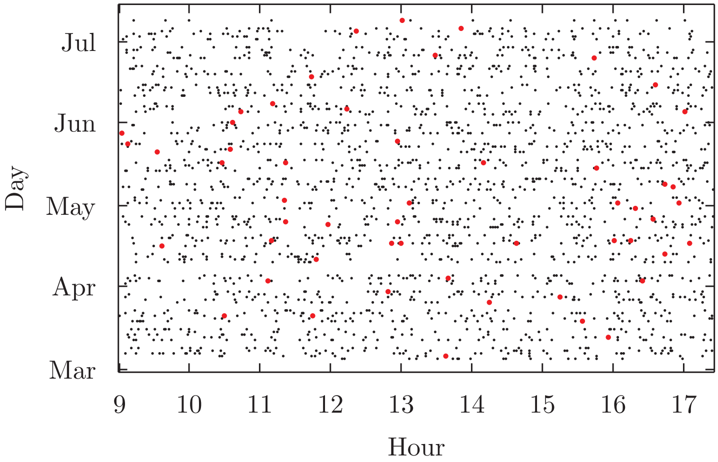
Est/gaps	Zeros	Not Rescaled	Rescaled
Abs	16%|50%	26%|30%	27%|9%
Bi-Power	14%|51%	24%|32%	29%|9%

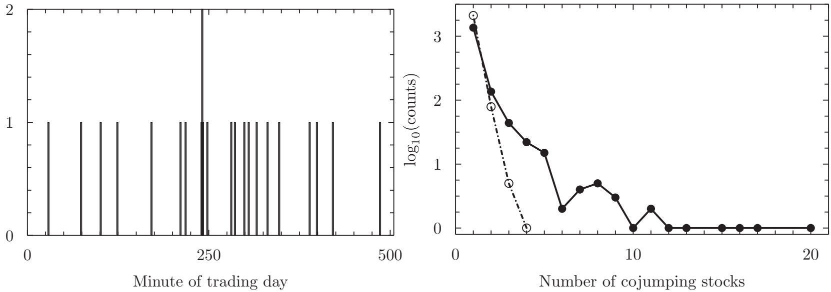
⇒ On the model, the intersection of the different methods gives a percentage of false positives well below 10%

Time series of the number of stocks n co-jumping simultaneously



Preserving the total number of jumps for each stock independently
(≡ independent Poisson processes)





Left. Intraday distribution of co-jumps with more than 5 stocks. $\Rightarrow$ No sign of intraday periodicity

Right. Histogram of the number of stocks jumping simultaneously in one minute. Red $\rightarrow$ Real data. Black $\rightarrow$ 0.01% confidence intervals under the null

Comments

- There is a large number of jumps per stock: ≈ 1.2 jumps per stock per day
- At the daily scale, the number of jumps is Poisson distributed
- There is a large number of “simultaneous” jumps of groups of stocks, not explained by Poisson
- Can we identify evidence of time clustering of jumps on the same or on different assets?
- Few events, thus robust statistical methods for identifying dependencies are needed

A **self-cojump** of a stock is the event of *at least two jumps* of the stock's price occurring inside a time window.
A **cross-cojump** between two stocks is the event of *at least one jump* in the series of both stocks occurring inside a given time window.

- s_i the number of jumps in window i
- estimator for self-cojumping probability in a window of length w

$$\hat{p}^{s}_{w} = \frac{\sum_{i=1}^{\lfloor \frac{N}{w} \rfloor} \mathbf{1}_{s_i \geq 2}}{\lfloor \frac{N}{w} \rfloor},$$

- estimator for cross-cojumping probability in a window of length w between stock l and k

$$\hat{p}^{c}_{w} = \frac{\sum_{i=1}^{\lfloor \frac{N}{w} \rfloor} \mathbf{1}_{s_i^l \geq 1} \mathbf{1}_{s_i^k \geq 1}}{\lfloor \frac{N}{w} \rfloor}.$$

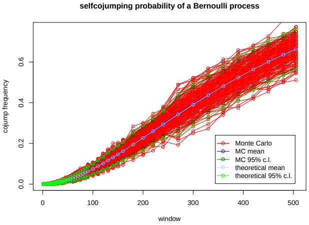
Under the independence assumption,

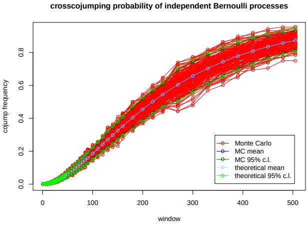
$$\mathbb{E}[\hat{p}^s_w] = p_{w,\lambda} \qquad \mathrm{Var}[\hat{p}^s_w] = \frac{p_{w,\lambda} - p^2_{w,\lambda}}{\left\lfloor \frac{N}{w} \right\rfloor},$$

where $p_{w,\lambda} = \mathrm{P}(\{s \geq 2\}) = 1 - Pois_{\lambda w}(s = 0) - Pois_{\lambda w}(s = 1)$.
Under the cross-independence assumption,

$$\mathbb{E}[\hat{p}^c_w] = q_{w,\lambda_l} q_{w,\lambda_k} \qquad \mathrm{Var}[\hat{p}^c_w] = \frac{q_{w,\lambda_l} q_{w,\lambda_k} - q^2_{w,\lambda_l} q^2_{w,\lambda_k}}{\left\lfloor \frac{N}{w} \right\rfloor},$$

where $q_{w,\lambda_i} = \mathrm{P}(\{s^i \geq 1\}) = 1 - Pois_{\lambda_i w}(s = 0)$.





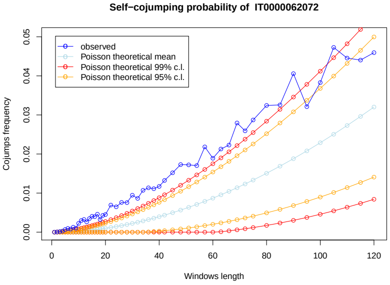
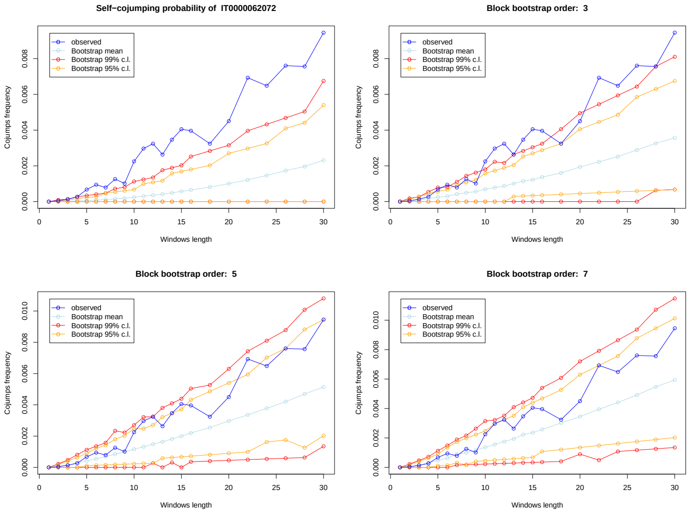


Figure: Self cojumping probability test for one typical Italian assets, $\lambda^{\mathrm{P}}_{62072} = 2.4 \cdot 10^{-3}$

The blue line corresponds to the empirical level of the estimator, the lightblue line to the theoretical mean, while the red and orange lines correspond to the boundaries of the 99% and 95% confidence bands respectively.

In our data sample the null is rejected for 18 stocks out of 20.
$\Rightarrow$ Strong evidence of time clustering of jumps and violation of the univariate Poisson model.

Block Bootstrap for identification of the time scale of clustering



Evidence of a time scale of several minutes

Hawkes Processes (Hawkes, 1971)
A univariate point process N_t is called a *Hawkes process* if it is a *linear self-exciting process*, defined by the intensity

$$I(t) = \lambda(t) + \int_{-\infty}^{t} \nu(t-u)\mathrm{d}N_u \,,$$

where λ is a deterministic function called the *base intensity* and ν is a positive decreasing weight function.
The most common parametrization of ν is given by $\nu(t) = \sum_{j=1}^{P} \alpha_j \mathrm{e}^{-\beta_j t}$, for $t > 0$, where $\alpha_j \geq 0$ are scale parameters, $\beta_j > 0$ controls the strength of decay, and $P \in \mathbb{N}$ is the order of the process.
The associated log-likelihood function for $P = 1$ can be recursively computed as

$$\mathcal{L}(t_1, \ldots, t_n) = (1-\lambda)t_n - \frac{\alpha}{\beta}\sum_{i=1}^{n}\left(1 - \mathrm{e}^{-\beta(t_n - t_i)}\right) + \sum_{i=1}^{n} \ln(\lambda + \alpha R_i)\,,$$

where $R_1 = 0$, and $R_i = \mathrm{e}^{-\beta(t_i - t_{i-1})}(1 + R_{i-1})$ for $i \geq 2$.

Test self-cojumps under Null Hawkes

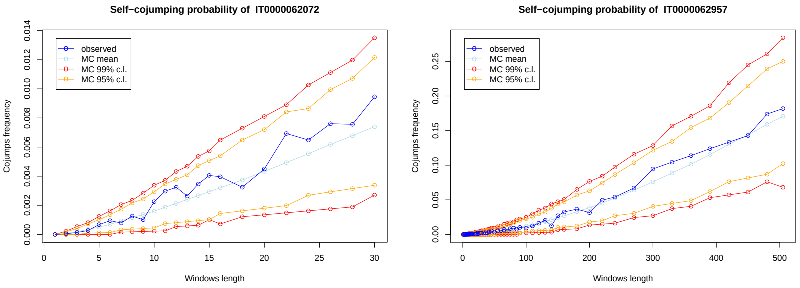


Figure: Self cojumping probability test for two Italian assets. Hawkes model parameters: $\lambda^{\mathrm{H}}_{62072} = 2.1 \cdot 10^{-3}$, $\alpha^{\mathrm{H}}_{62072} = 31 \cdot 10^{-3}$, $\beta^{\mathrm{H}}_{62072} = 0.25$, and $\lambda^{\mathrm{H}}_{62957} = 1.4 \cdot 10^{-3}$, $\alpha^{\mathrm{H}}_{62957} = 8.0 \cdot 10^{-3}$, $\beta^{\mathrm{H}}_{62957} = 0.28$; the number of Monte Carlo paths is 10^3.

$\Rightarrow$ Univariate Hawkes processes are able to capture the dynamics and time clustering of jumps of real data

The blue line corresponds to the empirical level of the estimator, the lightblue line to the Monte Carlo mean, while the red and orange lines correspond to the boundaries of the 99% and 95% confidence bands respectively.

Can a multivariate Hawkes process capture both the self- and the cross-cojumps?

A *K-dimensional Hawkes Processes* is a *linear self-exciting process*, defined by the multivariate intensity $\mathbf{I}(t) = \left(I^1(t), \ldots, I^K(t)\right)'$ where the k-type intensity using an exponential kernel with $P = 1$ is given by

$$I^k(t) = \lambda^k(t) + \sum_{m=1}^{K} \sum_{t_i^m < t} \alpha_{km} \mathrm{e}^{-\beta_{km}(t - t_i^m)} .$$

The process is *stationary* if the spectral radius of the matrix $\Gamma = \left(\frac{\alpha_{km}}{\beta_{km}}\right)_{k,m=1,\ldots,K}$ is strictly smaller than one. In this case the unconditional expected intensity is $\mathbb{E}[\mathbf{I}(t)] = (1_K - \Gamma)^{-1} \boldsymbol{\lambda}(t))$.

For the one-dimensional case this implies that, if stationarity holds and $\lambda(t)$ is constant, then the expected number of jumps in a time interval of length T is

$$\mathbb{E}\left[\int_{t_0}^{t_0+T} \mathrm{d}N_t\right] = \frac{\lambda}{1 - \alpha/\beta} T .$$

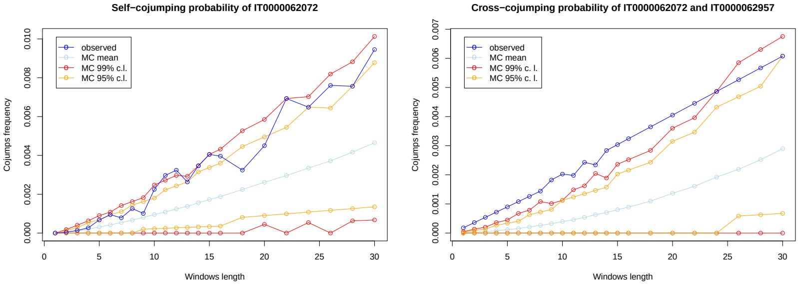


Figure: Self and Cross cojumping probability test against the null 2-dim Hawkes with parameters: $\lambda_1^H = 2.0 \cdot 10^{-3}$, $\alpha_{11}^H = 1.7 \cdot 10^{-2}$, $\beta_{11}^H = 0.30$, $\beta_{12}^H = 0.15$, and $\lambda_2^H = 1.3 \cdot 10^{-3}$, $\alpha_{22}^H = 8.6 \cdot 10^{-5}$, $\beta_{22}^H = 0.48$, $\beta_{21}^H = 0.60$ (β_{12}^H and β_{21}^H are not stat significant.) The number of Monte Carlo paths is 10^3.

The parameters describing the mutual excitation are often statistically significant

$\Rightarrow$ However multivariate Hawkes processes fail to describe both the single stock and the cross stock time clustering of jumps observed in real data

$\Rightarrow$ When calibrated on real data, Hawkes processes strongly underestimate the "simultaneous" jumps of stocks.

Main idea: A single factor model of jumps of a set of stocks

- There is one (unobserved) market factor point process describing the jumps
- When the factor jumps, stock i jumps with probability P_i
- A stock can jump also because of an idiosyncratic point process

A toy example: Pick 2 Stocks and calibrate a 1 Factor Model

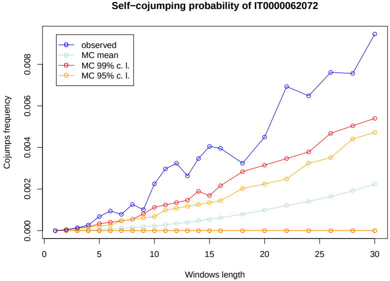
$$\begin{aligned} P_1 P_2 \lambda_F T &= n_{12} \\ P_1 \lambda_F T &= n_1 \\ P_2 \lambda_F T &= n_2 \end{aligned} \quad \Longrightarrow \quad P_1, P_2, \lambda_F$$

where n_1 and n_2 are the realized number of jumps of the stock S_1 and S_2, while n_{12} represents the realized number of cojumps between S_1 and S_2 within the one minute sampling interval.

- Poisson scenario: $\lambda_{Poisson} \equiv \lambda_F$

  Easy to implement with 2 stocks

  No unique way to extend to multi stocks



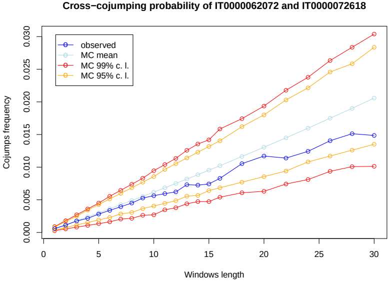


Figure: Self and Cross cojumping probability test against a null model specifically designed for two stocks: single stock jumps are generated by the thinning of a systemic factor driven by a Poisson process. The number of Monte Carlo paths is 10^3.

$\Rightarrow$ The factor term explains the cross-cojumps, but the Poissonianity of the factor leads to underestimation of self-cojumps (as expected).

One Factor Hawkes scenario (no idiosyncratic term)

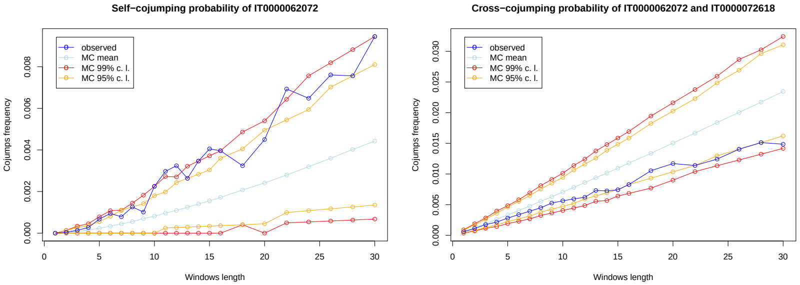


Figure: Self and Cross cojumping probability test against a null model specifically designed for two stocks: single stock jumps are generated by the thinning of a systemic factor driven by an exponential Hawkes process. The number of Monte Carlo paths is 10^3.

$\Rightarrow$ The Hawkes structure of the factor allows to explain both the self- and the cross-cojumps.

Consider the whole set of 20 Stocks and define

- $\{t_i^s\}$ for $i = 1, \dots, n_s$, set of event times for the s-th stock
- $\{t_j^F\}$ for $j = 1, \dots, n_F$, set of event times when the realized number of cross-cojumps rejects the null hypothesis of cross independence
- $\{t_{i'}^s\} \subseteq \{t_i^s\}$ for $i' = 1, \dots, n_{s'}$, subset of event times for the s-th stock compatible with the null of cross-independence

Systemic factor

Estimate λ_F, α_F, and β_F from the data set $\{t_j^F\}$.

Idiosyncratic components

$$\begin{aligned} P_1 \frac{\lambda_F}{1-\alpha_F/\beta_F} T &= n_1 - n_{1'} \\ &\vdots \\ P_{20} \frac{\lambda_F}{1-\alpha_F/\beta_F} T &= n_{20} - n_{20'} \end{aligned} \quad \Longrightarrow \quad P_1, \dots, P_{20}$$

Estimate λ_s, α_s, and β_s from the data set $\{t_{i'}^s\}$, for $s = 1, \dots, 20$.

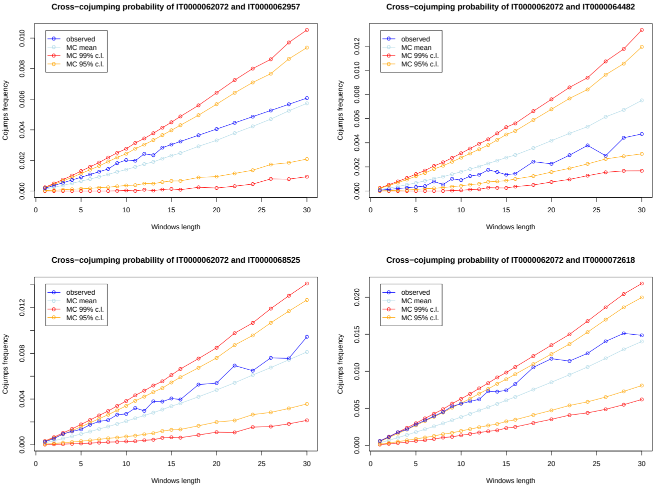


Figure: Cross cojumping probability test against a null model designed for arbitrary $N \geq 2$ stocks. Single stock jumps are the result both of the thinning of a systemic Poisson factor and the evolution of an exponential Hawkes process. The number of Monte Carlo paths is 10^3.

Test self-cojumps under Null One Factor Hawkes + Idiosyncratic Hawkes

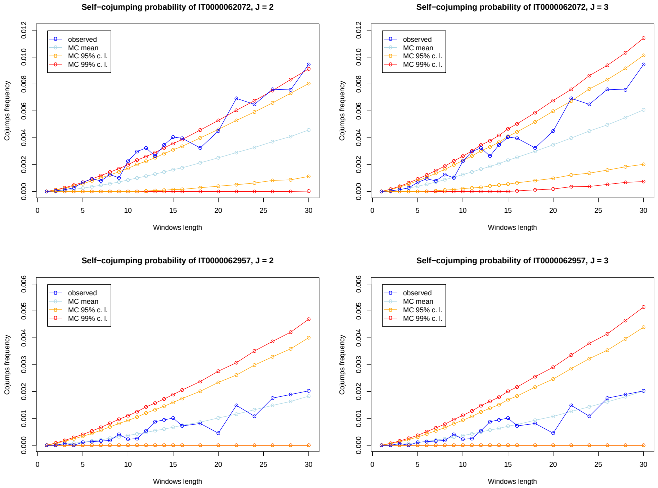


Figure: Self cojumping probability test against a null model designed for arbitrary $N \geq 2$ stocks. Single stock jumps are the result both of the thinning of a systemic Hawkes factor and the evolution of an exponential Hawkes process. Left column $J = 2$, right columns $J = 3$. The number of Monte Carlo paths is 10^3.

Conclusions

- A large number of jumps are present in financial markets
- The detection of jumps shows some dependence on the details of the detection method. We use intersection of several methods to improve the reliability of the events.
- On individual stocks, jumps are clearly not described by a Poisson process, but display time clustering well described by Hawkes processes
- We identify a very large number of simultaneous and systemic co-jumps, i.e. sizable sets of stocks "simultaneously" jumping
- A multivariate Hawkes process does not provide a satisfactory description of these systemic jumps
- We propose a one factor point process model which is able to describe
    1. The time clustering of jumps on individual stocks
    2. The time lagged cross excitation between different stocks
    3. The large number of "simultaneous" systemic jumps