CENTRE FOR ECONOMETRIC ANALYSIS
CEA@Cass

http://www.cass.city.ac.uk/cea/index.html

Cass Business School
Faculty of Finance
106 Bunhill Row
London EC1Y 8TZ

Appendix to A Frequency-Specific Factorization to Identify Commonalities with an Application to the European Bond Markets

Simona Boffelli, Jan Novotný, and Giovanni Urga

CEA@Cass Working Paper Series

WP–CEA–04a-2014

Appendix to A Frequency-Specific Factorization to Identify Commonalities with an Application to the European Bond Markets

Simona Boffelli[a], Jan Novotný[b], Giovanni Urga[c]

[a] *Bergamo University, Italy.*
Via dei Caniana 2, Bergamo, 24127, Italy. Simona.Boffelli@unibg.it, Tel: +39 035 205 2677, Fax: +39 035 205 2549.

[b] *Cass Business School, City University London, UK and CERGE-EI, CZ.*
106 Bunhill Row, London, EC1Y 8TZ, UK. Jan.Novotny.1@city.ac.uk, Tel: +44 (0)20 7040 8089, Fax: +44 (0)20 7040 8881.

[c] *Cass Business School, City University London, UK and Bergamo University, Italy.*
106 Bunhill Row, London, EC1Y 8TZ, UK. G.Urga@city.ac.uk, Tel: +44 (0)20 7040 8698, Fax: +44 (0)20 7040 8881.

This version: 27 November 2014

A. Data

We describe the main characteristics of government bond yields used in this study in full details. We then review the macro-factors, news announcements and bond auctions employed as the explanatory variables. Finally, we present the testing procedures to identify and date jump arrivals.

A.1. Yields

We use data for the 10-year government bonds of Belgium, France, Germany, Italy, the Netherlands and Spain over the period from June 1, 2007 to May 31, 2012.[1] We consider bid data. The 10-year bonds are market benchmarks defined as the most active at that maturity. Data were provided by Morningstar and come at tick-by-tick frequency which we re-sampled at 5-minute frequency using calendar time and excluding time intervals with values missing for at least one country. The 5-minute frequency is robust to micro-structure noise while offering sufficiently high frequency to properly evaluate the impact of specific events. The trading period considered in this paper is from 8 a.m. to 3:30 p.m. (UTC).

[1] Attinasi et al. (2011) and Bikbov and Chernov (2010) provide evidence that long-term maturities are more sensitive to macro-factors relative to short term maturities.

We detect and remove outliers by applying a filter which is a modification of the procedure to remove outliers proposed in Brownlees and Gallo (2006) and further elaborated by Barndorff-Nielsen et al. (2011, p. 156). Let p_t be a tick-by-tick time series of yields at time t, where for the purpose of the setting the filtering procedure we denote time as $t = [d,i]$, d denotes the particular trading day and i the intraday position. Then, the observation is removed if the following condition holds true

$$|p_{[d,i]} - p_{[d,i]}\left(\tilde{k}^L\right)| > \max\left\{4MD_{[d,i]}(\tilde{k}), n\gamma\right\} \wedge |p_{[d,i]} - p_{[d,i]}\left(\tilde{k}^R\right)| > max\left\{4MD_{[d,i]}(\tilde{k}), n\gamma\right\} \tag{A.1}$$

where $\tilde{k}$ is the bandwidth, $p_{[d,i]}\left(\tilde{k}^L\right)$ and $p_{[d,i]}\left(\tilde{k}^R\right)$ are sample medians of the $\tilde{k}/2$ observations before (L for left) and after (R for right) the time $t = [d,i]$, respectively, $MD_{[d,i]}\left(\tilde{k}\right)$ is the mean absolute deviation from the median of the whole neighbourhood, $\wedge$ denotes logical conjunction and n is a multiplier.

The advantage of this rule lies in the separate comparison of the $[d,i]$-th trade against its left and right neighbours while the measure of dispersion is calculated on the whole group of $\tilde{k}$ trades. This approach is specifically designed to avoid detection of outliers as false jumps. Finally, we also remove the first return of the day that occurs at 8 a.m., as it largely reflects the adjustment to information accumulated overnight and hence exhibits a spurious excess volatility compared to any other five-minute interval. In Table A.1, we report some descriptive statistics of the data.

For each time series, we report the overall number of ticks available, from which we remove holidays, weekends and trades occurring outside the trading period starting at 8 a.m. and ending at 3:30 p.m. UTC. We remove outliers following the description in (A.1) which leads us to detection of the percentage of outliers ranging from 0.09% for Germany to 0.16% for Belgium. In addition, we also report some descriptive statistics providing insights about market liquidity such as the mean number of trades per day and the time elapsed between two consecutive trades. Both statistics indicate that German market is the most liquid, with a daily average number of trades of 2,345 and a trade duration of 14.2 seconds, followed by France (828 trades, 38 seconds), Spain (764 trades, 38 seconds), Italy (736 trades, 43 seconds), Belgium (659 trades, 47 seconds) and the Netherlands

Table A.1: Data selection and descriptive statistics on government bond yields and spreads.

	DE	IT	FR	ES	BE	NL
No. ticks	3,077,442	978,261	1,096,247	978,357	841,854	657,249
Limiting trading time	2,928,107	917,455	1,027,268	969,129	831,094	645,773
No. trades per day	2,345	736	828	764	659	513
	[1889]	[526]	[596]	[512]	[481]	[378]
Trade duration: (sec)	14.2	42.9	38	38.1	47.0	60.4
	[44.4]	[97.1]	[88.6]	[90.3]	[115.7]	[123.4]
5-minute intervals	115,110	115,110	115,110	115,110	115,110	115,110
Exclude 1st daily obs	113,831	113,831	113,831	113,831	113,831	113,831
Bid YTM						
Mean (%)	3.18	4.66	3.61	4.58	4.01	3.48
	[0.82]	[0.69]	[0.58]	[0.65]	[0.47]	[0.75]
Median (%)	3.20	4.57	3.56	4.41	4.08	3.54
	(1.48 - 4.64)	(3.76 - 6.99)	(2.52 - 4.78)	(3.76 - 6.38)	(2.99 - 4.96)	(1.98 - 4.79)
Bid-Ask Spread of YTM						
Mean (bps)	0.63	0.64	0.78	0.75	1.00	0.72
	[0.05]	[0.05]	[0.08]	[0.05]	[0.06]	[0.05]
Median (bps)	0.62	0.64	0.79	0.75	1.00	0.72
	(0.56 - 0.76)	(0.51 - 0.80)	(0.66 - 0.94)	(0.67 - 0.89)	(0.89 - 1.11)	(0.65 - 0.85)

Note: Values in the square brackets are the standard deviation, while the pair of numbers in the round brackets denotes the 1st and 99th percentiles, respectively.

(513 trades, 60 seconds). After re-sampling at the 5-minute frequency and removing the 8 a.m. time interval for each day, we end up with 113,831 returns, covering 1,279 days, corresponding to 89 observations per day. In Table A.1, we also report descriptive statistics for yields: Italy has the highest average yield equal to 4.66%, while Germany has the lowest, equal to 3.18%. Note that the information the average indicator offers is limited, as the government bond yields vary significantly throughout our sample period, as can be seen from Figure A.1.

Government bond yields move very closely until May 2010, when markets start to pay more attention to sovereign debt risk in response to the Greek crisis. In May 2010, Greek government deficit was revised and estimated at 13.6% of GDP, causing a a correspondent decrease in international confidence in Greece's ability to repay its sovereign debt. As a consequence, despite the first rescue package approved by Eurozone countries and the IMF, concerns about the solvency of Euro countries began to rise together with yields.

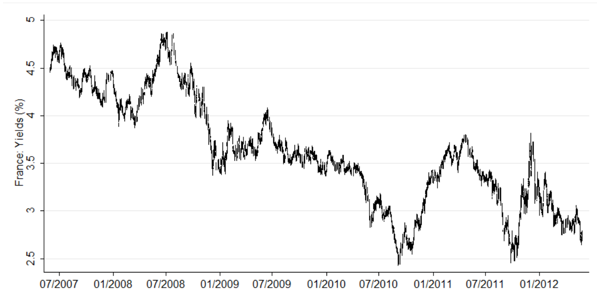
A.2. Macro-factors

We employ two real economy indicators: unemployment and industrial production levels, and a forward looking indicator, the Economic Sentiment. Our choice is motivated

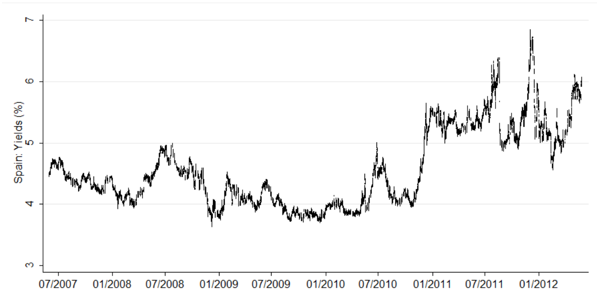
Figure A.1: Government 10-years bond: YTM.

(a) Italy

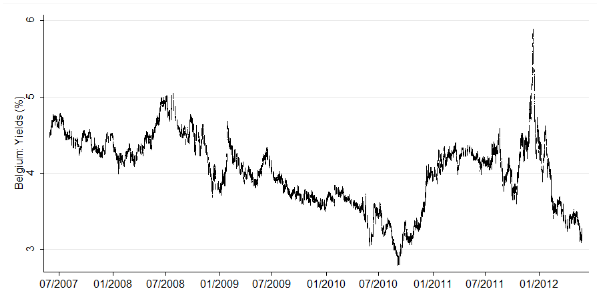
(b) France



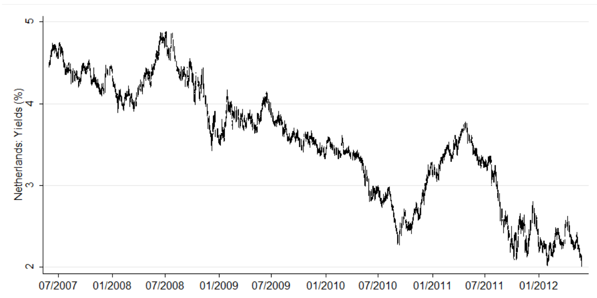
(c) Spain



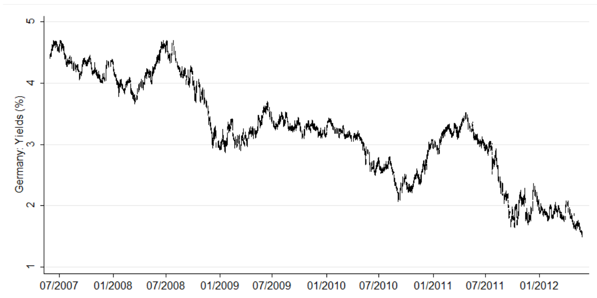
(d) Belgium



(e) the Netherlands



(f) Germany



Note: The figure reports the 10-year government bonds bid YTM at 5-minute frequency.

by the existing literature, for example Mody (2009) and Aizenman et al. (2013); industrial production is often found to be particularly relevant to asset behaviour in a number of studies, for example Schwert (1989), Ludvigson and Ng (2009) and Lustig et al. (2014). Data come from Eurostat.

Figure A.2 depicts the three macro-factors chosen in this analysis for every country on a monthly basis. The upper panel reports the unemployment rate, the middle panel captures the level of the industrial production, and the lower panel depicts the economic sentiment.

A.3. Macro-announcements

We consider macro-announcements related to the US, the Euro area, Belgium, France, Germany, Greece, Italy, the Netherlands, Portugal and Spain. In some cases, we are unable to use all available macro-announcements as they are released when markets are still closed. For instance, this is the case of France, with releases occurring between 6:30 and 7:45 a.m. UTC. In the case of Spain, though macro-announcements are released at 8:00 a.m. UTC, we keep the indicators, shifting them to 8:05 a.m. in order to match them with the trading period considered for bonds. Data related to macro-announcements are median expected value by survey of panellists E, forecasts standard deviation σ and actual value of the release denoted as A. Data were collected from Bloomberg. In our application, we adopt the standard surprise measure ζ defined as

$$\zeta = \frac{(A - E)}{\sigma}\,. \tag{A.2}$$

Table A.2 provides a complete list of macro-announcements considered in this paper together with brief characteristics.[2]

Table A.2: Macroeconomic news with pre-scheduled announcements

Country	Macro-announcement	Freq.	Release time (UTC)	No.	Cat.	$S(\sigma_S)$

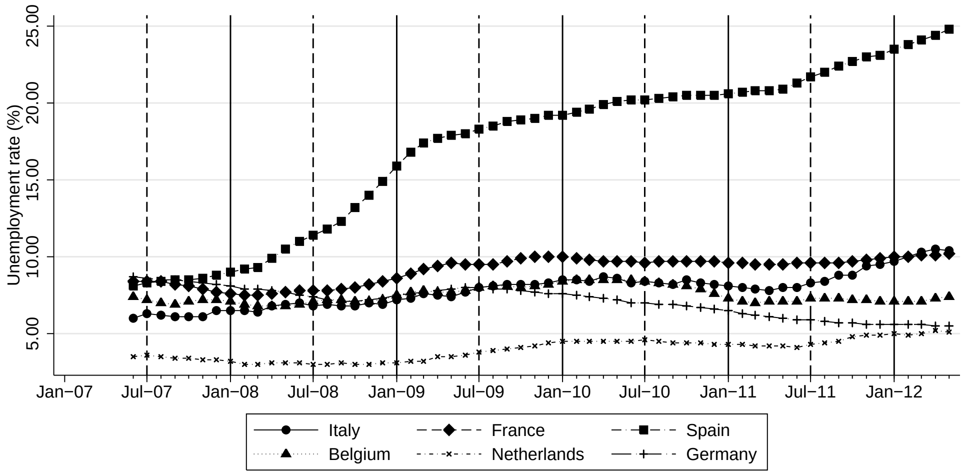
[2]Note: The table reports the details on macro-economic announcements released in the period from June 1, 2007 to May 31, 2012. In some cases the release time changes according to the daylight-saving time. In addition, the change of daylight-saving time is not perfectly synchronized for Europe and US. FL stands for Forward Looking, P for Price and RE for Real Economy macro-announcement categories. Surprise is defined as $\zeta = \frac{(A-E)}{\sigma}$ with A being the actual released value, E the surveyed value by panellists and σ the standard deviation of forecasts.

Country	Macro-announcement	Freq.	Release time (UTC)	No.	Cat.	$S(\sigma_S)$
US	Business inventories	M	15:00	55	RE	-0.38 (1.97)
	Chicago PMI	M	14:45	60	FL	0.80 (2.73)
	Consumer confidence	M	15:00	57	FL	-0.55 (3.29)
	CPI	M	13:30	57	P	0.14 (1.31)
	Durable goods	M	13:30	58	FL	-0.31 (2.98)
	Factory orders	M	15:00	59	FL	0.28 (1.30)
	GDP advance	Q	12:30 / 13:30	20	RE	1.78 (4.56)
	GDP preliminary	Q	12:30 / 13:30	19	RE	-0.31 (1.48)
	GDP final	Q	12:30 / 13:30	20	RE	-0.30 (2.14)
	Industrial production	M	14:15	57	RE	-0.29 (1.78)
	Initial jobless claim	W	13:30	251	RE	0.06 (1.45)
	Non-farm payroll	M	13:30	58	RE	-0.30 (2.25)
	Philadelphia FED Index	M	15:00	51	FL	-0.16 (3.65)
	PPI	M	13:30	57	P	0.14 (1.81)
	Retail sales	M	13:30	57	RE	0.06 (1.92)
	University of Michigan	M	14:55	59	FL	0.86 (1.71)
EA	Business climate	M	09:00	57	FL	-0.03 (2.09)
	Consumer confidence	M	10:00	56	FL	-0.09 (1.80)
	Flash HICP	M	10:00	58	P	0.12 (1.37)
	HICP	M	10:00	41	P	0.00 (0.00)
	Industrial production	M	10:00	59	RE	-0.05 (1.51)
	Introductory Statement	M	13:30	58	RE	-
	M3	M	09:00	57	P	-0.49 (2.77)
	Monthly Bulletin	M	10:00	60	RE	-
	PMI flash	M	09:00	59	FL	0.44 (2.14)
	PMI final	M	09:00	60	FL	0.46 (2.59)
	PPI	M	10:00	59	P	0.01 (0.95)
	Retail sales	M	10:00	59	RE	-0.58 (1.41)
	Unemployment	M	10:00	54	RE	0.41 (1.71)

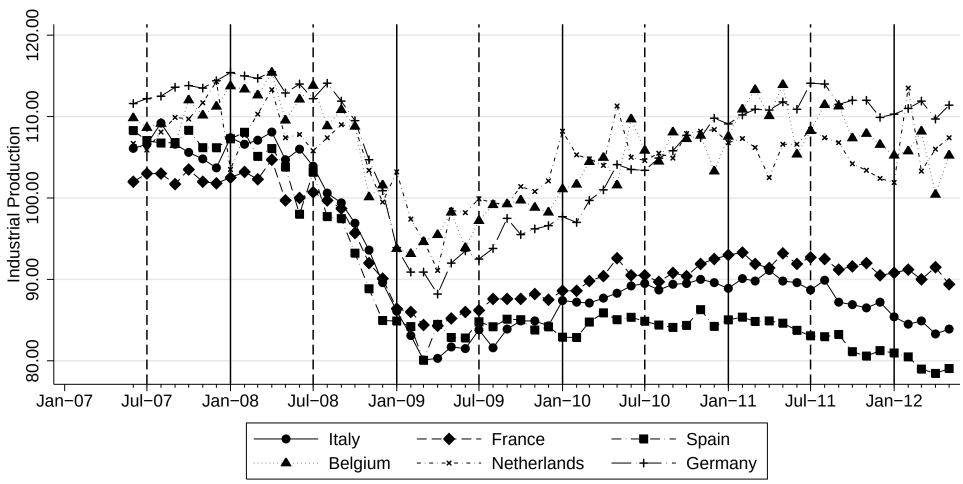
Country	Macro-announcement	Freq.	Release time (UTC)	No.	Cat.	$S(\sigma_S)$
DE	CPI preliminary	M	13:00	51	P	0.05 (1.34)
	IFO: business confidence	M	09:00	57	FL	0.59 (2.75)
	Industrial production	M	11:00	57	RE	-0.12 (2.18)
	Unemployment	M	08:55	58	RE	-0.45 (2.17)
	ZEW	M	10:00	58	FL	0.32 (2.44)
IT	Business confidence	M	08:30 / 09:00	59	FL	-0.23 (2.86)
	CPI preliminary	M	10:00	58	P	0.28 (2.12)
	CPI final	M	09:00 / 10:00	56	P	-1.05 (3.67)
	GDP preliminary	Q	09:00 / 10:00	17	RE	-1.04 (2.47)
	GDP final	Q	09:00 / 10:00	18	RE	-0.06 (0.80)
	Industrial production	M	09:00	58	RE	-0.04 (2.03)
FR	Industrial production	M	07:45 / 08:45	2	RE	6.67 (25.93)
ES	CPI	M	08:00	59	P	-0.01 (0.67)
	GDP preliminary	Q	08:00	20	RE	0.10 (1.74)
	GDP final	Q	08:00	19	RE	-0.11 (0.81)
	Industrial production	M	08:00	57	RE	-0.45 (3.06)
	Unemployment	Q	08:00	20	RE	0.96 (1.75)
PT	CPI	M	10:00	58	P	-0.85 (2.52)
	GDP preliminary	Q	10:00	19	RE	4.68 (9.58)
	GDP final	Q	11:00	20	RE	-0.28 (1.09)
NL	CPI	M	08:30	58	P	-0.02 (1.18)
	Industrial production	M	08:30	53	RE	-2.29 (5.63)
	Unemployment	M	08:30	54	RE	-0.48 (1.93)
BE	Business confidence	M	14:00	59	P	-0.26 (2.92)
GR	CPI	M	10:00	58	P	-0.58 (2.91)
	GDP preliminary	Q	08:30 / 10:00	8	RE	-0.19 (1.03)
	GDP final	Q	08:30 / 10:00	7	RE	-3.00 (2.38)
	Unemployment	M	10:00	39	RE	-0.17 (2.75)

Figure A.2: Macroeconomic factors.

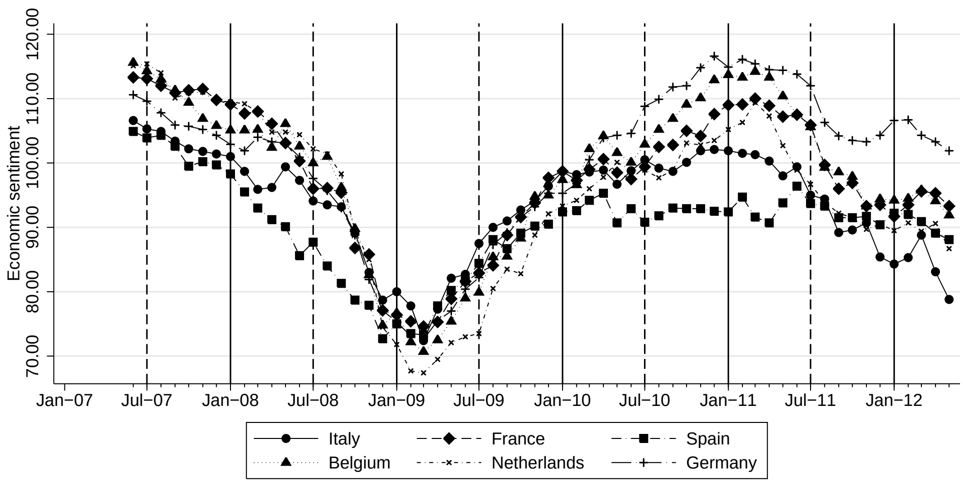
(a) Unemployment rate (%)



(b) Industrial production



(c) Economic sentiment



Note: In the figure, the upper panel reports the unemployment rate, the middle panel the level of the industrial production, and the lower panel the economic sentiment. Sample period: June 2007 - May 2012. Data were obtained from the Eurostat website.

Table A.3: Government bond auctions

	Total number of bond auctions	Number of 10-year bond auctions	Average yield (%) Mean (SD)
Austria	42	21	3.64 (0.68)
Belgium	145	32	3.98 (0.55)
Finland	8	4	2.75 (0.46)
France	366	52	3.61 (0.59)
Germany	261	44	3.04 (0.84)
Greece	69	4	4.78 (0.09)
Italy	272	64	4.76 (0.75)
the Netherlands	193	22	3.42 (0.74)
Portugal	129	22	5.01 (0.79)
Spain	193	33	4.85 (0.81)

Note: The table reports details on government bond auctions held from June 1, 2007 to May 31, 2012. Average yield refers to the yield at which the government allocated the bonds issued in an auction, and it is collected only for 10-year bond auctions. Values in brackets denote the standard deviation.

A.4. Auctions

We take into consideration auctions of European countries issuing Euro-denominated bonds: Austria, Belgium, Finland, France, Germany, Greece, Italy, the Netherlands, Portugal and Spain. Most auctions take place between 8 and 10 a.m. UTC. To capture the performance of an auction, we consider the average yield at which the government sells the bonds. Average yields were collected just for auctions relative to 10-year bonds as they not only correspond to the maturity of the bonds analysed but they even represent the most relevant ones. In Table A.3, we report the total number of auctions, the number of 10-year bond auctions and the mean and standard deviation of the average yield. Average yields reflect country-specific sovereign risk: safer countries such as Finland and Germany succeed in selling bonds at higher prices and lower returns, with an average yield of 2.75% and 3.04% respectively, while riskier countries such as Italy, Spain and Portugal allocate their bonds at an average yield of 4.76%, 4.85% and 5.01%, respectively.

B. Identification and Testing for Jumps

In this section, we specify the identification procedure for price jump arrivals and the testing for a presence of price jumps in a given sample. There is a growing strand of literature devoted to the identification of price jumps: see, among others, Aït-Sahalia (2004), Barndorff-Nielsen and Shephard (2006), Andersen et al. (2007), Jiang and Oomen

(2008), Aït-Sahalia and Jacod (2009), Lee and Mykland (2008) or Lee and Hannig (2010). In this work, we employ tests proposed by Lee and Mykland (2008) and Barndorff-Nielsen and Shephard (2006), respectively.

B.0.1. Identification of Jump Arrivals

In order to analyse the arrivals and co-arrivals, we have to specify the mapping, which allows identifying the exact time of jump arrivals. To this purpose, we use the Lee and Mykland (2008) test (LM henceforth). The LM test suffers from the limitation that spot volatility is assumed constant over the local window; therefore, we use the correction for the intraday volatility pattern as suggested by Andersen and Bollerslev (1998). Thus, the LM test corrected for the intraday volatility pattern at time t is defined as

$$fLM_t = \frac{|r_t|}{\widehat{\sigma}_t \cdot \Xi_{[d,i]}} \, ,$$

where the local integrated variance at time t is $\widehat{\sigma_t}^2 = \frac{1}{(n-2)\mu_1^2} \sum_{j=n-2}^{1} |r_{t-j}||r_{t-(j-1)}|$, with $\mu_1 = \sqrt{2/\pi}$ and $n = 81$ corresponding to 90% of the number of observations at 5-minutes sampling frequency as suggested by Lee and Mykland (2008); $\Xi_{[d,i]}$ is the correction term at time t indexed in terms of the trading day d and its intraday position i and defined as

$$\Xi_{[d,i]} = \frac{\exp\left(f\left(\theta_{TML}; x_{[d,i]}\right)\right)}{\sqrt{\frac{1}{N_i}\sum_{i=1}^{N_i}\left[\exp\left(f\left(\theta_{TML}; x_{[d,i]}\right)\right)\right]^2}} \, , \tag{B.1}$$

and estimated via the Truncated Maximum Likelihood (TML) estimator proposed by Boudt et al. (2011), who show that the filtered jump test statistic increases the accuracy of intraday jump detection methods, where

$$\begin{aligned} f\left(\theta_{TML}; x_{[d,i]}\right) \quad = \quad & \delta_0 + \delta_{0,1}\frac{i}{N_1} + \delta_{0,2}\frac{i^2}{N_2} + \sum\nolimits_{j=1}^{J} \lambda_j S_{[d,i],j} + \sum\nolimits_{k=1}^{K} \bar{\lambda}_k \bar{S}_{[d,i],k} + \sum\nolimits_{j=1}^{4} \theta_j W_j + \\ & + \sum\nolimits_{p=1}^{P}\left(\delta_{c,p} cos\left(\frac{2\pi P}{N} i\right) + \delta_{s,p} sin\left(\frac{2\pi P}{N} i\right)\right) + \varepsilon_{d,i} \, , \end{aligned} \tag{B.2}$$

where $N_1 = \frac{N_i+1}{2}$ and $N_2 = \frac{(N_i+1)(N_i+2)}{6}$ are normalizing constants; $S_{[d,i],j}$ is the surprise for macro-announcements; $\bar{S}_{[d,i],j}$ is the surprise for government bond auctions computed as the difference in average yield between current and previous 10-year auction; J is the number of macro-announcements available; K is the sum of all auctions considered; W_j is the dummy for a day of the week; λ_j and $\bar{\lambda}_k$ are specific loading coefficients; P is a tuning parameter determining the order of the expansion of the sinusoids; θ_{TML} is full parameter vector to be estimated.

Moreover, loading coefficients λ_j are modelled applying the Andersen and Bollerslev (1998) decay-structure, which allows the specific event to impact over a time window but with decaying weights. For macro-announcement surprises, we set the time window from 30 minutes prior to the release up to 90 minutes after it, as in Andersen and Bollerslev (1998). For government bond auctions, we use a window ranging from two hours before the end of the auction to one hour after, as we want to take into account the uncertainty in the markets during the auction period.

Then, following Lee and Mykland (2008), the ξ-statistic for jump tests is defined as:

$$\frac{\max_{t\in A_\nu} |fLM_t| - C_\nu}{S_\nu} \to \xi\,, \tag{B.3}$$

where A_ν is the tested region with ν observations, $C_\nu = (2\ln\nu)^{1/2} - \frac{\ln\pi + \ln(\ln\nu)}{2(2\ln\nu)^{1/2}}$, and $S_\nu = \frac{1}{(2\ln\nu)^{1/2}}$. The ξ-statistic follows the standard Gumbel distribution, which is characterized by the probability distribution function $P(\xi \leq x) = \exp\left(e^{-x}\right)$.

B.0.2. Testing for Jumps

In addition to the dating of the jump arrivals, we also need a test statistic which allows us to decide on the significance of jump(s) over a certain time interval $[T_{k-1}, T_k]$. We employ the $\hat{G}$-statistic as introduced by Barndorff-Nielsen and Shephard (2006). The $\hat{G}$-statistic for an univariate process specified by a 1-dimensional equation for the price dynamics over a time $[T_{k-1}, T_k]$ sampled into n equidistant time steps $T_{k-1} = t_0 < t_1 < \cdots < t_n = T_k$ reads

$$\hat{G} = n^{1/2}\frac{\widehat{IV}_n - \widehat{QV}_n}{\widehat{IQ}_n} \xrightarrow{\mathcal{D}} N\left(0, \theta\right)\,, \tag{B.4}$$

where $\widehat{IV}_n$ stands for the estimator of the Integrated Variance, $\widehat{QV}_n$ for the estimator of the Quadratic Variance, $\widehat{IQ}_n$ for the estimator of the Integrated Quarticity, and $\xrightarrow{\mathcal{D}}$ indicates stable convergence in law. The null is no jump(s) in the $[T_{k-1}, T_k]$.

In our study, we use the following estimators for the $\widehat{QV}_n$, $\widehat{BV}_n$, and $\widehat{IQ}_n$

$$\widehat{QV}_n : \widehat{RV}_n = \sum_{j=1}^{n} \left(Y_{t_j} - Y_{t_{j-1}} \right)^2 , \tag{B.5}$$

$$\widehat{IV}_n : \widehat{BV}_n = \mu_1^{-2} \left(\frac{n}{n-1} \right) \sum_{j=2}^{n} \left| Y_{t_j} - Y_{t_{j-1}} \right| \left| Y_{t_{j-1}} - Y_{t_{j-2}} \right| , \tag{B.6}$$

$$\widehat{IQ}_n : \widehat{MPQ}_n = \mu_1^{-4} n \left(\frac{n}{n-3} \right) \sum_{j=4}^{n} \prod_{k=0}^{3} \left| Y_{t_{j-k}} - Y_{t_{j-k-1}} \right| , \tag{B.7}$$

where Y_{t_j} is the t_j-th observation in the time interval $[T_{k-1}, T_k]$, $\mu_\alpha = E\left(|z|^\alpha\right)$, with $z \sim N(0,1)$, in particular, and $\mu_1 = \sqrt{2/\pi}$. $\widehat{IV}_n$ and $\widehat{IQ}_n$ estimators were introduced in Barndorff-Nielsen and Shephard (2004), Andersen et al. (2012), Christensen et al. (2010), Corsi et al. (2010), Aït-Sahalia and Jacod (2007), Mancini (2006) and all three estimators have the following asymptotic properties

$$\widehat{RV}_n \xrightarrow{p} QV \, , \, \widehat{BV}_n \xrightarrow{p} IV \, , \, \widehat{MPQ}_n \xrightarrow{p} IQ \, , \text{ as } n \to \infty .$$

In this set-up, the parameter θ in equation (B.4) takes the value $\theta \approx 0.61$.

C. Testing for Price Jumps Country-by-country

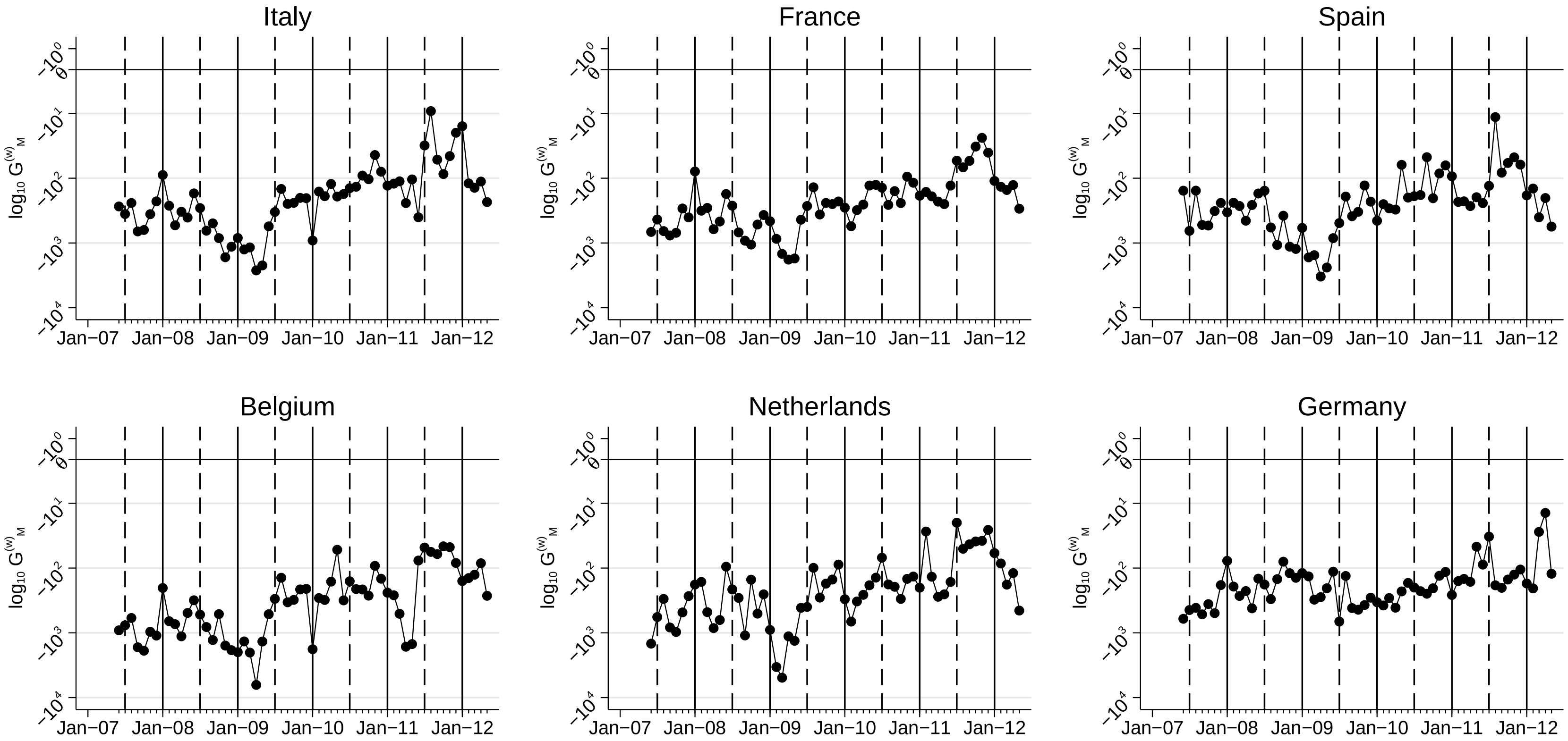
Figure B.1 depicts the $\log_{10}\left(G_n^{(w)}\right)$ estimated for each month and each country independently. For every instance, we can reject the null hypothesis of no jumps since all the $G_n^{(w)}$-statistic range between $-6,356.70$ and -9.18, which lie well below -2.11, the 5% critical value to reject the null hypothesis of no jumps.

D. Zero-coupon Bond Optimal Portfolio

The presence of the global co-jump raises a question: what will the $G_n^{(w)}$-statistic be for a portfolio constructed using the standard portfolio optimization techniques? In our case, price time series cannot be directly employed in constructing the market portfolio, as our analysis is not based on a specific bond for each country but rather on 10-year benchmarks. Those benchmarks are identified with respect to time to maturity and to liquidity criteria. Therefore, by analysing the time series of the prices, we would face issues connected to roll-over.

Instead, we use a "zero-coupon bond approximation". This means that we treat the

Figure B.1: $G_n^{(w)}$-statistic by countries on monthly basis.



Note: The figure reports the $\log_{10}$ transformation of the $G_n^{(w)}$-statistic estimated for every month and for every country. θ indicates the 5% critical value to reject the null hypothesis of no jumps.

set of 10-year bonds as if they were zero-coupon bonds, so that we can get prices through the very simple yield-to-price formula

$$100 = p_k (1 + y_k)^{10} \ ,$$

where y_k is the yield to (constant) maturity, which is fixed for every month k at 10 years. Then, for each month, we construct an optimal portfolio using the Markowitz optimization.[3] The optimal portfolio in terms of risk-return trade-off is given as a solution to

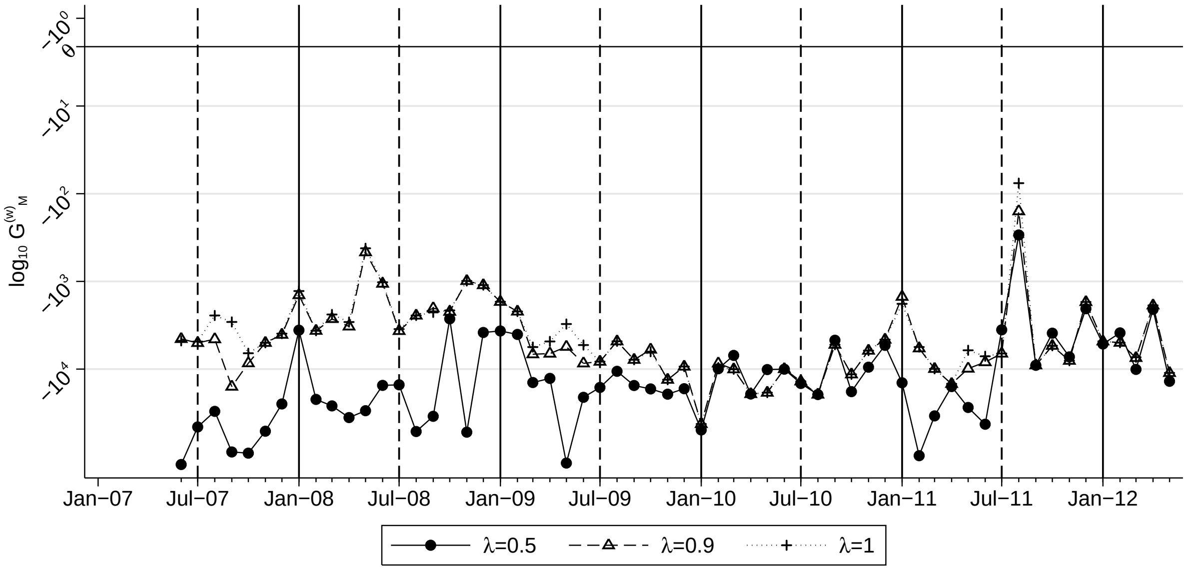
$$w_k = \arg\max_{\tilde{w}} E\left[\sum_{t\in k}\sum_{j=1}^{6} \tilde{w}^{(j)} \left(\frac{p_t^{(j)} - p_{t-1}^{(j)}}{p_t^{(j)}}\right)\right] - \frac{(1-\lambda)}{\lambda} SD_k \left[\sum_{j=1}^{6} \tilde{w}^{(j)} \left(\frac{p_t^{(j)} - p_{t-1}^{(j)}}{p_t^{(j)}}\right)\right] , \tag{D.1}$$

where the first term is the expected cumulative return for a month k of the weighted index and the second term is the risk part with $SD_k\left[\left(\frac{p_t - p_{t-1}}{p_t}\right)\right]$ denoting the standard deviation of portfolio returns rescaled to a monthly scale. The parameter λ reflects the risk-aversion where $\lambda = 1$ denotes a risk neutral agent, while for $\lambda \to 0$, the risk aversion is increases above all limits. We perform the exercise with three values of $\lambda = 1$, $\lambda = 0.9$, and $\lambda = 0.5$. Once we obtain the optimal weights w as a solution to (D.1), we calculate the $G_n^{(w)}$-statistic for a composite index based on "zero-coupon bond approximation" weights.

Figure D.1 depicts the $G_n^{(w)}$-statistic for three optimal portfolios with $\lambda = 1$, $\lambda = 0.9$, and $\lambda = 0.5$. The $G_{n,\lambda=0.5}^{(w)}$-statistic ranges between $-12.24 \cdot 10^4$ and -295.43, the $G_{n,\lambda=0.9}^{(w)}$-statistic ranges between $-42,223.74$ and -158.27, and the $G_{n,\lambda=1}^{(w)}$-statistic ranges between $-42,223.74$ and -75.76. In all three cases, we reject the null hypothesis of no jumps for every month at 5% confidence level, indicating that the optimal zero-coupon bond based portfolio does not lead to global co-jumps. In addition, we see that the most risk averse case of $\lambda = 0.5$ results in the lowest levels of $G_n^{(w)}$-statistic. This can be ascribed to the fact that the prices which rose the most were those of Germany, which is perceived as the least risky through the standard deviation measure. Therefore, even the three threshold indices constructed using the "zero-coupon bond approximation" do not lead to a situation

[3] Note that the portfolio based on the "zero-coupon bond approximation" is a proxy for what is actually traded.

Figure D.1: $G_n^{(w)}$-statistic for "zero-coupon bond approximation" portfolio.



Note: The figure reports the $\log_{10}$ transform of the $G_n^{(w)}$-statistic for three specifications of the optimal portfolio of approximated zero-coupon bonds with $\lambda = 0.5$, $\lambda = 0.9$, and $\lambda = 1$, plotted for every month. θ indicates the 5% critical value to reject the null hypothesis of no jumps.

in which jumps are disappearing, or are even close to disappearing.

E. Measure of Commonality and Pearson Binary Correlation

In this paper, we introduce the measure of commonality to describe the commonality of the price jumps in the portfolio of N assets and link the commonality to the prevailing economic conditions. In this section, we link the measure of commonality Q for $N = 2$ with the standard Pearson correlation coefficient ϕ for two time series with binary variables.

Let us denote $j_t^{(1)}$ and $j_t^{(2)}$ the binary time series, with $t = 1, \ldots, n$, given as

$$j_t^{(i)} = \begin{cases} 1 & \text{jump at } i\text{-th time series at time } t \\ 0 & \text{otherwise} \end{cases} .$$

The Pearson correlation coefficient for binary variables is then based on the contingency table captured in Table E.1. In particular, following Agresti (1996) the Pearson correlation coefficient ϕ is given by

$$\phi = \frac{n_{11}n_{00} - n_{10}n_{01}}{\sqrt{n_{\bullet 1}n_{\bullet 0}n_{1\bullet}n_{0\bullet}}}\,, \tag{E.1}$$

Table E.1: Contingency table.

	$j_t^{(2)} = 1$	$j_t^{(2)} = 0$	
$j_t^{(1)} = 1$	n_{11}	n_{10}	$n_{1\bullet}$
$j_t^{(1)} = 0$	n_{01}	n_{00}	$n_{0\bullet}$
	$n_{\bullet 1}$	$n_{\bullet 0}$	n

Note: In the table, the coefficient n_{11} indicates number of common jumps, n_{10} the number of idiosyncratic jumps at $j_t^{(1)}$, n_{01} the number of idiosyncratic jumps at $j_t^{(2)}$, n_{00} the number of cases without any jump, n the total number of observations, and coefficients $n_{\bullet 1}$, $n_{\bullet 0}$, $n_{1\bullet}$, and $n_{0\bullet}$ denote the marginals for both time series, respectively.

where all the variables n_{ij} are as defined in Table E.1. The Pearson correlation takes a value of 0 when there is no agreement between two time series, while the maximum positive/negative values correspond to the perfect agreement/disagreement between the two time series. In general, the coefficient ϕ is defined on $[-1, 1]$. For a fixed proportions $n_{\bullet 1}/n$ and $n_{1\bullet}/n$, however, the Pearson correlation ϕ lies in the interval $[\phi_{min}, \phi_{max}]$ with

$$\begin{aligned} \phi_{min} &= \max\left(-\left(\frac{n_{\bullet 1}/n \cdot n_{1\bullet}/n}{(1 - n_{\bullet 1}/n)\,(1 - n_{1\bullet}/n)}\right)^{1/2}, -\left(\frac{(1 - n_{\bullet 1}/n)\,(1 - n_{1\bullet}/n)}{n_{\bullet 1}/n \cdot n_{1\bullet}/n}\right)^{1/2}\right), \\ \phi_{max} &= \min\left(\left(\frac{n_{\bullet 1}/n\,(1 - n_{1\bullet}/n)}{n_{1\bullet}/n\,(1 - n_{\bullet 1}/n)}\right)^{1/2}, \left(\frac{n_{1\bullet}/n\,(1 - n_{\bullet 1}/n)}{n_{\bullet 1}/n\,(1 - n_{1\bullet}/n)}\right)^{1/2}\right). \end{aligned}$$

The Pearson correlation ϕ is a proper statistic and can be used to test the null of no incidence between the two time series with the asymptotic distribution of ϕ related to the χ^2 distribution as

$$\phi^2 = \frac{\chi^2}{n}.$$

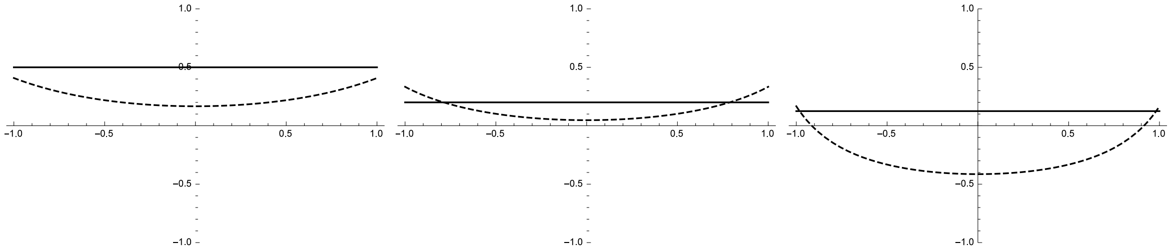
The measure of commonality, on the other hand, in the case of $N = 2$ can be expressed in terms of Contingency Table E.1 defined as

$$Q = \frac{n_{11}}{n - n_{00}}, \tag{E.2}$$

which clearly differs from E.1. In general, Q lies in $[0, 1]$, where for a fixed $n_{\bullet 1}/n$ and $n_{1\bullet}/n$, the measure of commonality is bound in $Q \in [Q_{min}, Q_{max}]$ with

Figure E.1: The comparison the measure of commonality and the Pearson correlation.

(a) $\frac{n_{11}}{n} = 0.4$ and $\frac{n_{00}}{n} = 0.2$. (b) $\frac{n_{11}}{n} = 0.1$ and $\frac{n_{00}}{n} = 0.5$. (c) $\frac{n_{11}}{n} = 0.1$ and $\frac{n_{00}}{n} = 0.2$.



Note: In the figure, the solid line depicts the measure of commonality, Q, and the dash line the Pearson correlation ϕ. The x-axis denotes the ratio r between n_{10} and n_{01}, namely $\frac{n_{10}}{n} = \left(1 - \frac{n_{11}}{n} - \frac{n_{00}}{n}\right)\frac{1+r}{2}$ and $\frac{n_{10}}{n} = \left(1 - \frac{n_{11}}{n} - \frac{n_{00}}{n}\right)\frac{1-r}{2}$.

$$
\begin{aligned}
Q_{min} &= \frac{\max\left(n_{1\bullet} + n_{\bullet 1} - n, 0\right)}{\min\left(n_{1\bullet} + n_{\bullet 1}, n\right)}, \\
Q_{max} &= \min\left(\frac{n_{\bullet 1}}{n_{1\bullet}}, \frac{n_{1\bullet}}{n_{\bullet 1}}\right).
\end{aligned}
$$

To understand the difference in the two information measures ϕ and Q, let us fix n, n_{11}, and n_{00}, i.e., we fix Q. In fact, we fix $n_{10} + n_{01}$ as well. However, the measure of commonality Q is insensitive to the asymmetry between n_{10} and n_{01}, respectively, while ϕ is a function of the ratio n_{10}/n_{01}. Figure E.1 depicts the measure of commonality and the Pearson correlation for three cases as function of the ratio r between n_{10} and n_{01}. The ratio is defined such that $\frac{n_{10}}{n} = \left(1 - \frac{n_{11}}{n} - \frac{n_{00}}{n}\right)\frac{1+r}{2}$ and $\frac{n_{10}}{n} = \left(1 - \frac{n_{11}}{n} - \frac{n_{00}}{n}\right)\frac{1-r}{2}$. The three cases are: $\frac{n_{11}}{n} = 0.4$ and $\frac{n_{00}}{n} = 0.2$, $\frac{n_{11}}{n} = 0.1$ and $\frac{n_{00}}{n} = 0.5$, and $\frac{n_{11}}{n} = 0.1$ and $\frac{n_{00}}{n} = 0.2$. The figures stress the difference between the two measures. In particular, in the last case, a particular combination of n_{10} and n_{01} exists such that the Pearson correlation suggests no incidence between the two time series, while the measure of commonality maintains fixed value.

With a given $n_{\bullet 1}/n$ and $n_{1\bullet}/n$ being fixed, the measure of commonality assesses the degree of association between the two time series. In particular, the highest possible overlap is reached if and only if $Q \to Q_{max}$. When $n_{\bullet 1}/n$ and $n_{1\bullet}/n$ are not specified, the low level of the measure of commonality means that there is a low association between the two time series, or that the two time series are not close enough in terms of the number of events. In particular, if $n_{\bullet 1}/n \ll n_{1\bullet}/n$ and $n_{11} = n_{\bullet 1}$, the measure of commonality is low in levels even though there is a perfect overlap between events.

In this paper, we do not aim to specify the asymptotic distribution of the measure of commonality Q under the null of no association between the two time series, as the main use of the measure of commonality is to assess the agreement for $N > 2$ and not to compete with the Pearson correlation. In addition, compared to the Pearson correlation, the measure of commonality is not properly centred, i.e., the mean under the null depends on $n_{\bullet 1}/n$ and $n_{1\bullet}/n$, respectively. In the main text, we assess the null of no association between the time series in the finite sample using the Monte Carlo procedure.

F. Extensions: Correlations and Temporal Aggregation

In this section, we finally discuss two natural extensions of our proposed framework. The first deals with whether a significant change exists in the global co-jump vector with respect to Germany during the emergence of the European debt crisis; the second refers to the sensitivity of our frequency-specific factorization approach to the time-aggregation frequency.

F.0.3. Correlations

We analyse the high-frequency pair-wise correlation between yields using the consistent and efficient estimator of the high-frequency covariance proposed by Aït-Sahalia et al. (2010). The estimator allows us to deal with situations when data are observed asynchronously and are possibly affected by the market micro-structure noise. Assuming that log-returns of the j-th asset $r_t^{(j)} = Y_{t_i}^{(j)} - Y_{t_{i-1}}^{(j)}$ follow an MA(1) process in the time interval $[0, T]$, where the sampling $t_1, t_2, \ldots$ of the time interval $[0, T]$ is not necessarily equidistant. The log-likelihood function for vector of observed returns r is

$$l(a^2, \sigma^2) = -\frac{1}{2} \log \det (\Omega) - (n/2) \log (2\pi) - \frac{1}{2} r' \Omega^{-1} r \,, \tag{F.1}$$

with

$$
\Omega = \begin{bmatrix} \sigma^2\Delta^i + 2\sigma_u^2 & -\sigma_u^2 & 0 & \dots & 0 \\ -\sigma_u^2 & \sigma^2\Delta^i + 2\sigma_u^2 & -\sigma_u^2 & \vdots & \vdots \\ 0 & -\sigma_u^2 & \sigma^2\Delta^i + 2\sigma_u^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & -\sigma_u^2 \\ 0 & \dots & 0 & -\sigma_u^2 & \sigma^2\Delta^i + 2\sigma_u^2 \end{bmatrix}. \tag{F.2}
$$

where Δ^i=t_i-t_{i-1}.

The covariance estimator can now be obtained from the polarisation formula

$$
\mathrm{Cov}\left(r^{(1)}, r^{(2)}\right) = \frac{1}{4}\left[\mathrm{Var}\left(r^{(1)} + r^{(2)}\right) - \mathrm{Var}\left(r^{(1)} - r^{(2)}\right)\right], \tag{F.3}
$$

where Var(•) denotes the Quasi-Maximum Likelihood Estimator (QMLE) of the quadratic variation that is obtained from (F.1).

Figure F.1 depicts the average daily correlation coefficients for every pair of countries calculated for each month. First, there is evidence of an overall decrease in a pair-wise correlation when the European debt crisis started during 2010. Second, there is a striking regime change for all pairs including Germany from December 2010 to May 2011, where correlations change sign and decrease in absolute value. It is worth noticing that this corresponds to the period of change in the global co-jump vectors. The change in the global co-jump vectors thus occurs due to the correlation.

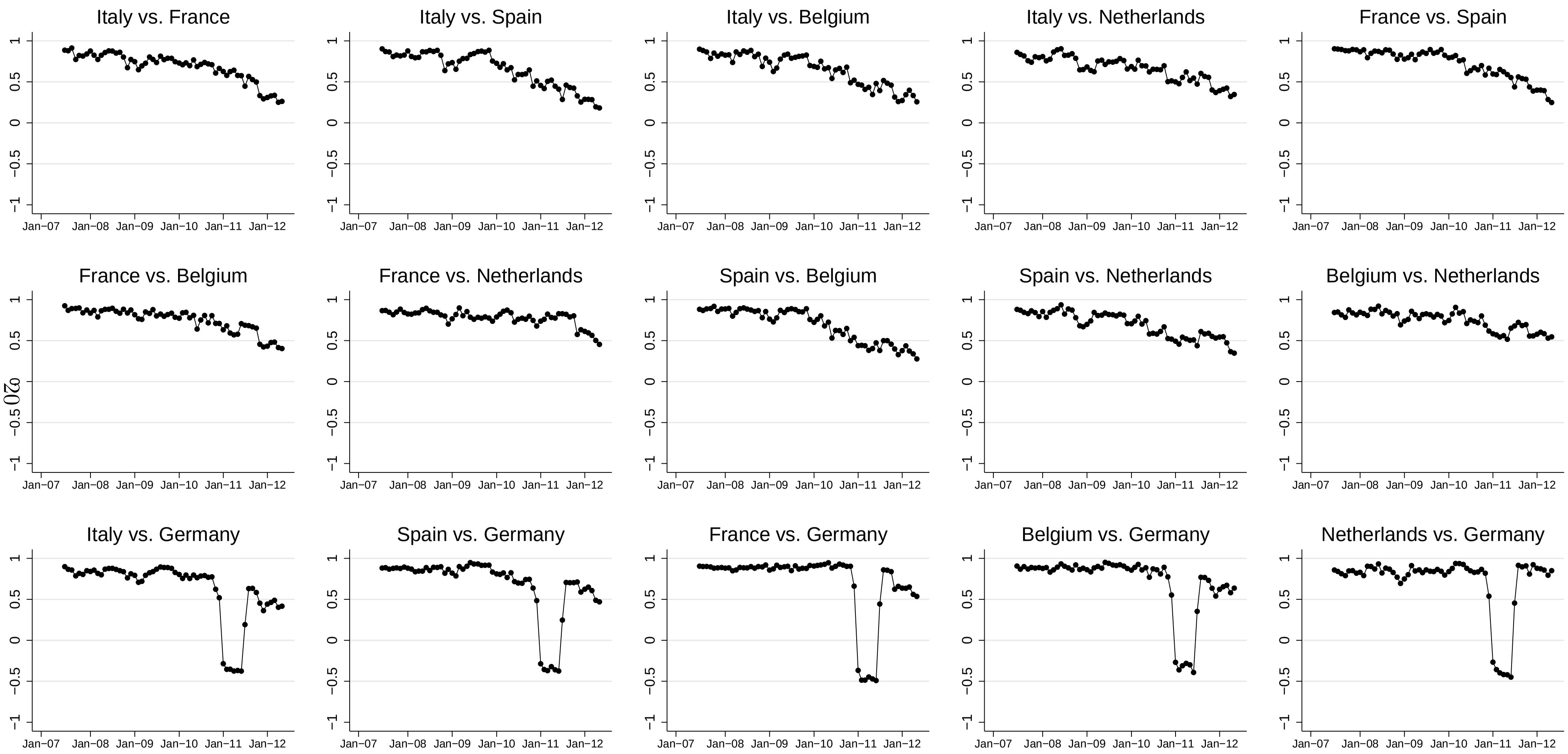
The change in correlation between Germany and all other countries between December 2010 and May 2011 suggests that the risk-awareness of investors increased due to the distress and they started to sell all bonds except German ones, which were perceived as a safe haven.

F.0.4. Temporal Aggregation

Throughout the paper, our analysis is based on monthly time series in order to link the commonalities to the macro-economic factors measured at monthly frequency. In this section, we contrast our previous results using measure of commonality and measure of multiplicity calculated at daily and weekly frequencies.

Figure F.2 depicts the daily measure of commonality (upper panel) and the daily measure of multiplicity (lower panel). First, there is evidence of that both measures are coarse,

Figure F.1: Pair-wise correlation in yields.



Note: The figure reports the average daily correlations between all combination of countries for every month using the procedure by Aït-Sahalia et al. (2010).

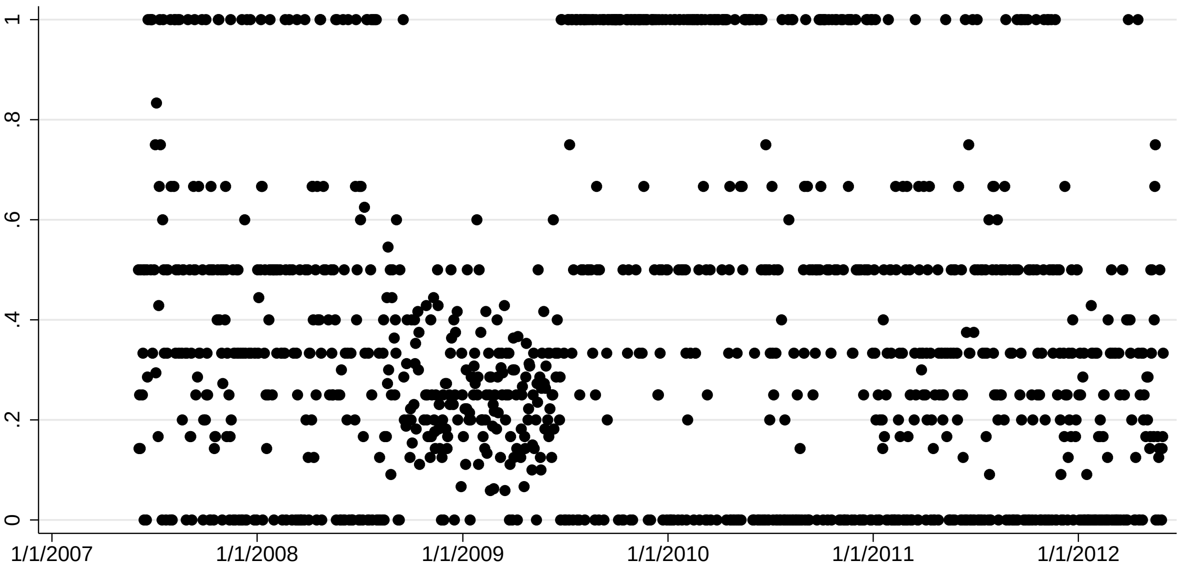
and this supports the claim that aggregating data at daily frequency is not appropriate due to the low number of arrivals. Second, the coarseness disappears during two periods in our sample: *i)* the aftermath of the Lehman Brothers collapse, and *ii)* the deepening of the European debt crisis in early 2011. The disappearance of the coarseness during the former period can be explained by the increase in the number of arrivals. In addition, both measures reach lower values in these two periods, which can be explained by the rise of the idiosyncratic arrivals for all countries except Germany. The later period, however, shows the absence of a significant increase of the absolute number of jump arrivals; this indicates a structural change in the bond market.

Further, Figure F.3 depicts the measure of commonality and the measure of multiplicity on a weekly frequency. In this case, coarseness nearly disappeared, while there is a confirmation of the different market behaviour during the aftermath of the Lehman Brothers collapse and the deepening of the European debt crisis in early 2011.

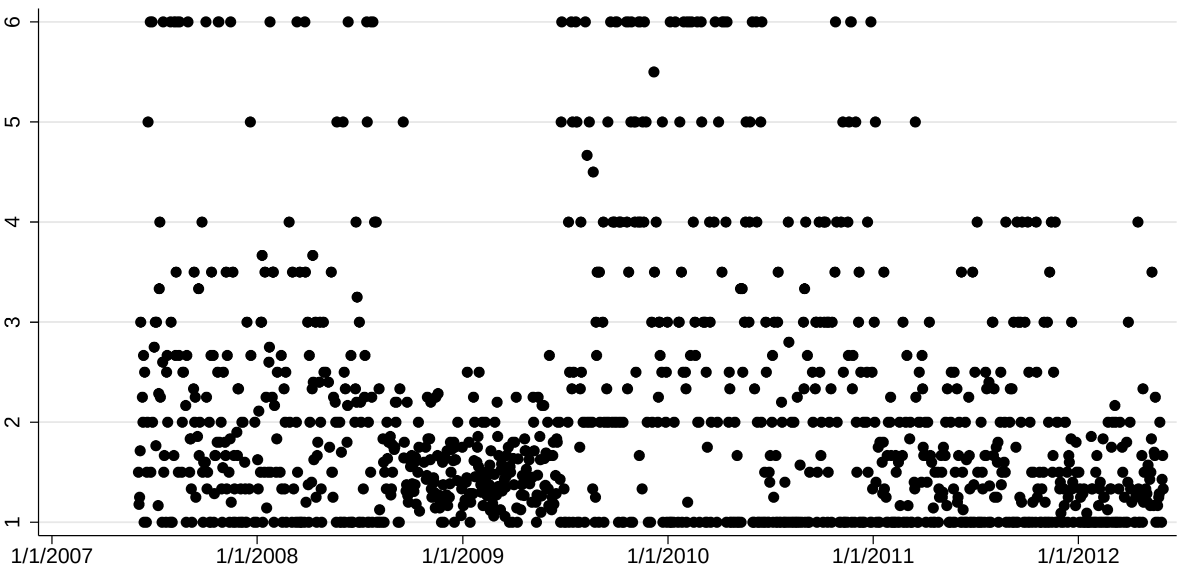
The main conclusion from the temporal aggregation exercise is that the monthly frequency used throughout the paper is the optimal aggregation to evaluate the link between high-frequency price jump arrivals and low frequency macro-variables.

Figure F.2: The measure of commonality and the measure of multiplicity on daily basis.

(a) The measure of commonality.



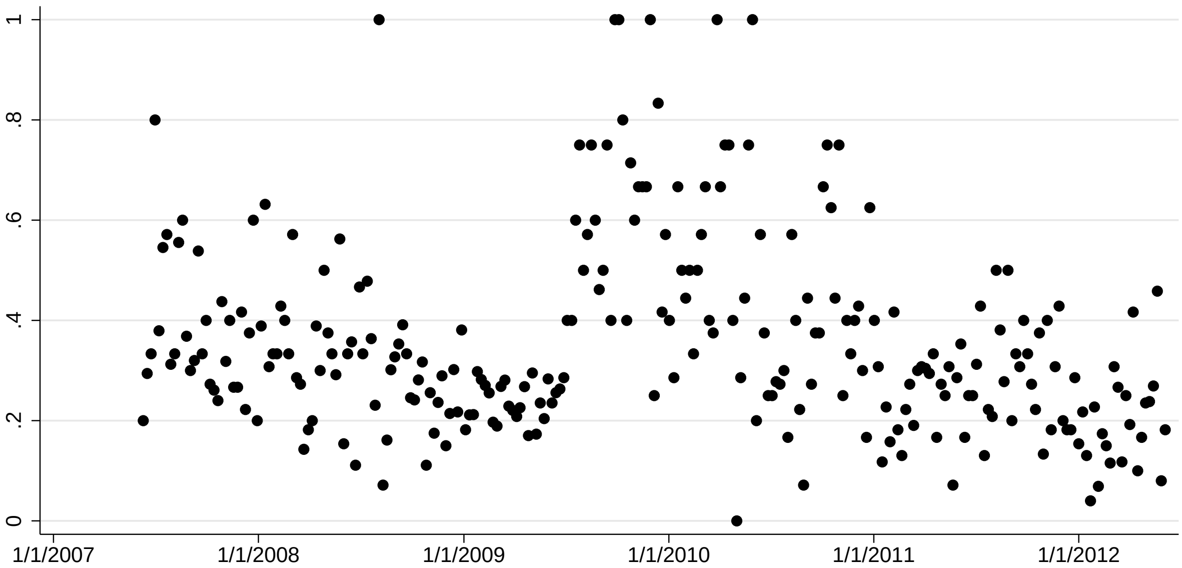
(b) The measure of multiplicity.



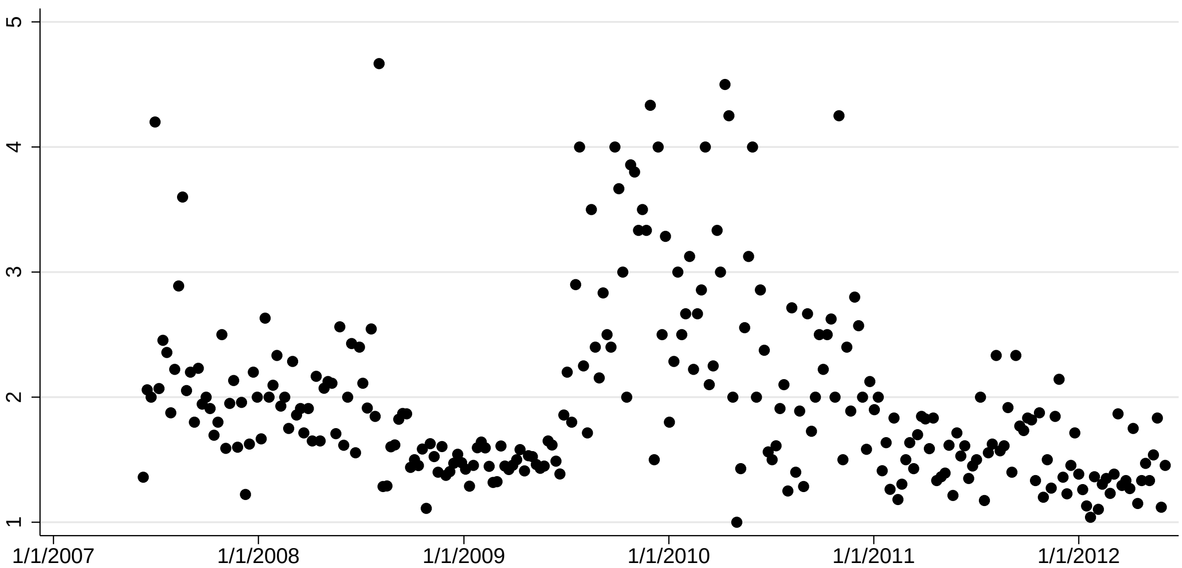
Note: The figure reports the measure of commonality and multiplicity for each day.

Figure F.3: The measure of commonality and the measure of multiplicity on weekly basis.

(a) The measure of commonality.



(b) The measure of multiplicity.



Note: The figure reports the measure of commonality and multiplicity for each week.

References

Agresti, A., 1996. An introduction to categorical data analysis, Wiley series in probability and mathematical statistics. Vol. 135. Hoboken, NJ: Wiley-Interscience.

Aït-Sahalia, Y., 2004. Disentangling diffusion from jumps. Journal of Financial Economics 74 (3), 487–528.

Aït-Sahalia, Y., Fan, J., Xiu, D., 2010. High-frequency covariance estimates with noisy and asynchronous financial data. Journal of the American Statistical Association 105 (492), 1504–1517.

Aït-Sahalia, Y., Jacod, J., 2007. Volatility estimators for discretely sampled Lévy processes. The Annals of Statistics 35 (1), 355–392.

Aït-Sahalia, Y., Jacod, J., 2009. Testing for jumps in a discretely observed process. The Annals of Statistics 37 (1), 184–222.

Aizenman, J., Hutchison, M., Jinjarak, Y., 2013. What is the risk of european sovereign debt defaults? fiscal space, cds spreads and market pricing of risk. Journal of International Money and Finance 34, 37–59.

Andersen, T., Bollerslev, T., Dobrev, D., 2007. No-arbitrage semi-martingale restrictions for continuous-time volatility models subject to leverage effects, jumps and iid noise: Theory and testable distributional implications. Journal of Econometrics 138 (1), 125–180.

Andersen, T. G., Bollerslev, T., 1998. Deutsche mark–dollar volatility: intraday activity patterns, macroeconomic announcements, and longer run dependencies. Journal of Finance 53 (1), 219–265.

Andersen, T. G., Dobrev, D., Schaumburg, E., 2012. Jump-robust volatility estimation using nearest neighbor truncation. Journal of Econometrics 169 (1), 75–93.

Attinasi, M. G., Checherita, C., Nickel, C., 2011. What explains the surge in euro area sovereign spreads during the financial crisis of 2007-2009? In: Kolb, R. W. (Ed.), Sovereign Debt: From Safety to Default. Vol. 605. John Wiley & Sons.

Barndorff-Nielsen, O., Shephard, N., 2004. Power and bipower variation with stochastic volatility and jumps. Journal of Financial Econometrics 2 (1), 1–37.

Barndorff-Nielsen, O., Shephard, N., 2006. Econometrics of testing for jumps in financial economics using bipower variation. Journal of Financial Econometrics 4 (1), 1–30.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., Shephard, N., 2011. Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. Journal of Econometrics 162 (2), 149–169.

Bikbov, R., Chernov, M., 2010. No-arbitrage macroeconomic determinants of the yield curve. Journal of Econometrics 159 (1), 166–182.

Boudt, K., Croux, C., Laurent, S., 2011. Robust estimation of intraweek periodicity in volatility and jump detection. Journal of Empirical Finance 18 (2), 353–367.

Brownlees, C. T., Gallo, G. M., 2006. Financial econometric analysis at ultra-high frequency: Data handling concerns. Computational Statistics & Data Analysis 51 (4), 2232–2245.

Christensen, K., Oomen, R., Podolskij, M., 2010. Realised quantile-based estimation of the integrated variance. Journal of Econometrics 159 (1), 74–98.

Corsi, F., Pirino, D., Renò, R., 2010. Threshold bipower variation and the impact of jumps on volatility forecasting. Journal of Econometrics 159 (2), 276–288.

Jiang, G. J., Oomen, R. C., 2008. Testing for jumps when asset prices are observed with noise: a swap variance approach. Journal of Econometrics 144 (2), 352–370.

Lee, S., Hannig, J., 2010. Detecting jumps from lévy jump diffusion processes. Journal of Financial Economics 96 (2), 271–290.

Lee, S., Mykland, P., 2008. Jumps in financial markets: A new nonparametric test and jump dynamics. Review of Financial Studies 21 (6), 2535–2563.

Ludvigson, S. C., Ng, S., 2009. Macro factors in bond risk premia. Review of Financial Studies 22 (12), 5027–5067.

Lustig, H., Roussanov, N., Verdelhan, A., 2014. Countercyclical currency risk premia. Tech. Rep. 3.

Mancini, C., 2006. Estimating the integrated volatility in stochastic volatility models with lévy type jumps. Preprint.

Mody, A., 2009. From Bear Stearns to Anglo Irish: how eurozone sovereign spreads related to financial sector vulnerability. Vol. 9. International Monetary Fund.

Schwert, G. W., 1989. Why does stock market volatility change over time? Journal of Finance 44 (5), 1115–1153.