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–04-2014
A Frequency-Specific Factorization to Identify Commonalities with an Application to the European Bond Markets

Simona Boffelli\textsuperscript{a}, Jan Novotn\textsuperscript{y}\textsuperscript{b}, Giovanni Urga\textsuperscript{c}

\textsuperscript{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.

\textsuperscript{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.

\textsuperscript{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

Abstract

We propose a frequency-specific framework to link the common features in the multivariate high-frequency price jumps with the low-frequency exogenous factors. We introduce measures of commonality and multiplicity based on high-frequency data and define the notions of co-arrivals and co-jumps to explore the contribution of individual assets. We employ this framework to study 10-year high-frequency European government bond yields as a function of macro-factors, macro-announcements and bond auctions. Both idiosyncratic and common jump arrivals are significant, with the idiosyncratic arrivals more sensitive to financial distress characterised by a low level of commonality in jump arrivals.

Keywords: Arrivals, Jumps, Co-arrivals, Co-jumps, European Government Yields, Bond auctions, Macro-announcements, Macfactors.

J.E.L. Classification Number: G12, C12, C32, H63


Jan Novotný acknowledges financial support from the Centre for Econometric Analysis and GAČR grant 14-27047S.
1. Introduction

The recent European sovereign debt crisis highlights the fact that bond markets can be as volatile as equity markets, thus diminishing the sense of safety—risk-free asset status—in particular when yields diverge across countries. The bond markets determine not only the cost of long-term borrowing for governments and the overall perception of countries’ fiscal stability, but government bonds belong to asset classes which are widely used by investors. When focused on the intraday level, the high-frequency properties of the bond market reveal the presence of jumps similar to those found with equities. Recent finance and econometrics literature suggests that price jumps embody specific risks (see, among others, Jarrow and Rosenfeld, 1984, and Pan, 2002), affect the pricing mechanisms of various financial instruments (see, e.g., Duffie et al., 2000, and Johannes, 2004) and provide a better explanation of credit and market risks (Carr and Wu, 2010).

Relevant literature reports a significant link between high-frequency government bond returns and news announcements: for instance, Andersen et al. (2007), de Goeij and Marquering (2006), and Beechey and Wright (2009) find a strong impact of the US-related news, especially that related to the real economy such as non-farm payroll news. There is evidence of a statistically significant link between bond markets and macroeconomic factors: for instance Ludvigson and Ng (2009) and Lustig et al. (2014) show the importance of industrial production in explaining bond returns, while Aizenman et al. (2013) examine the role of forward looking indicators given that bonds are inherently related to a country’s future performance. An analysis of the link between price jumps and macro-announcements appears rather recently in the literature: Lee (2012) and Boudt and Petitjean (2014) focus on equities, Dungey et al. (2009), and Jiang et al. (2011) consider bond markets, while Lahaye et al. (2011) cover different asset classes.

The contribution of this paper is threefold. First, we define the measures of commonality and multiplicity, which capture the degree of association for the price jump arrivals in the portfolio. We use these measures to test for price jumps that arrive randomly across the portfolio. Further, using frequency-specific factorization to treat idiosyncratic and common price jumps separately, we explicitly link the underlying high-frequency instantaneous intensities of the arrival processes to the low-frequency exogenous (macro-)factors.
Second, based on the co-feature framework introduced by Engle and Granger (1987) and Engle and Kozicki (1993), we propose the notion of co-arrivals defined as a linear combination of the arrival processes such that the aggregate number of arrivals is minimized. This concept is then used to define two additional co-arrival measures, the index and the grade, which assess the ratio of the eliminated total and common price jump arrivals due to the co-arrival, respectively. The notion of co-arrivals is then extended to co-jumps, where the magnitude of the price jumps is taken into account as well. This concept relates to the portfolio optimization as it allows identifying the portfolio with minimum contribution of price jumps such that it would not be exposed to shocks in the economy.

Third, we employ the commonality framework we propose to the high-frequency time series of European sovereign debt markets using the 10-year benchmark bonds for Belgium, France, Germany, Italy, the Netherlands and Spain over the period June 2007 to May 2012 including the Great Recession and the European sovereign debt crisis. The link between the properties of price jumps and the state of the economy is estimated through the evaluation of the dependence of price jumps on the economic indicators: unemployment, industrial production and economic sentiment, observed at monthly frequency, and the aggregate monthly surprise carried by macro-announcements and government bond auctions.

This paper provides a novel framework to the common price jumps literature as for instance in Jacod and Todorov (2009) and Bollerslev et al. (2008), who investigate co-jumps in high-frequency equities, or Dungey et al. (2009) who evaluate the price jumps across the term structure of the high-frequency US treasury bonds. In addition, our theoretical framework is specifically built to deal with the commonality of rare events—price jump arrivals—across a large portfolio of time series and link it to the real economy indicators measured at different frequencies. Thus, it provides an alternative approach to the MIXed DAta Sampling (MIDAS) literature, as proposed in a series of papers by Ghysels et al. (2004), Ghysels et al. (2005), Ghysels et al. (2006), Ghysels et al. (2007) and Engle et al. (2013), which links data at daily frequency with data sampled at lower frequency such as monthly or quarterly.

The remainder of the paper is as follows. In Section 2, we introduce the frequency-specific framework, define the measures of commonality and multiplicity with their formal
link to low-frequency exogenous factors and provide the foundations of co-arrivals and co-jumps. In Section 3, we describe the data used for the empirical application and summarize the testing procedures for price jump identification. In Section 4, we report the empirical results and some extensions. Section 5 concludes.

2. Theoretical Framework

We model the vector of $N$ log-prices as an $N$-dimensional vector $Y' = \left( Y^{(1)}, \ldots, Y^{(N)} \right)$, where the vector $Y = \{ Y_t \}_{0 \leq t \leq T}$ is defined on the $N$-dimensional probability space $\left( \Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P} \right)$ over the time interval $[0, T]$. The vector of log-prices is a semi-martingale, $\mathcal{F}_t$-adapted and its continuous-time dynamics can be specified by the stochastic differential equation

$$dY_t = \mu_t dt + \sigma_t dB_t + dJ_t,$$

where $\mu_t$, $\sigma_t^{(j,j')}$, $dB_t$ and $dJ_t$ are well behaved. In particular, $\mu_t$ is $(N \times 1)$ vector of the drift processes with $\mu_t^{(j)}$, $j = 1, \ldots, N$, each $\mathcal{F}_t$-adapted, locally bounded and predictable processes. Matrix elements $\sigma_t^{(j,j')}$, with $j, j' = 1, \ldots, N$, are $\mathcal{F}_t$-adapted, càdlàg and almost surely bounded away from zero. The vector $dB_t$ contains $N$ independent $\mathcal{F}_t$-adapted standard Brownian motions $dB_t^{(j)}$, with $j = 1, \ldots, N$. The term $dJ_t$ represents the $(N \times 1)$-dimensional vector of pure jump Lévy processes composed of a linear combination of the finite activity processes.

In what follows, we provide a frequency-specific factorization to link the high-frequency jump process of financial assets to the low frequency domain of the macroeconomic indicators.

2.1. Frequency-Specific Factorization

The jump term $dJ_t$ in (1), capturing the jump dynamics of $N$ mutually correlated financial assets, can be factorized as follows

$$dJ_t : \underbrace{U_t dJ_t}_{(N \times 1)} \rightarrow \underbrace{U_t}_{(N \times M)} \underbrace{dJ_t}_{(M \times 1)} ,$$

where the matrix of jump magnitudes $U_t$ contains $\mathcal{F}_t$-adapted random processes $U_t^{(i,m)}$ driven by an unspecified distribution $G^{(i,m)}_U$ with mean $\mu_U^{(i,m)}$ and standard deviation
σ_{U}^{(j,m)}$, with $j = 1, \ldots, N$ and $m = 1, \ldots, M$, where $M$ denotes the number of independent jump processes affecting the set of $N$ assets. The $M$-dimensional vector $dJ_t$ drives the arrivals of jumps, where every $dJ_t^{(m)}$ is either zero or one.

The distinction between idiosyncratic and common jumps is introduced by restricting the matrix $U_t$ to be

$$U_t = \begin{pmatrix} U_t^I & U_t^C \end{pmatrix},$$

(3)

where $U_t^I \equiv \text{diag}(U_t^{I(1)}, \ldots, U_t^{I(N)})$ and the $j$-th diagonal term $U_t^{I(j)}$ is almost surely non-zero and finite if there are idiosyncratic arrivals for the $j$-th asset. On the other hand, the matrix process $U_t^C$ has at least two almost surely non-zero elements in each column corresponding to common arrivals among assets.

The corresponding vector of arrivals is decomposed as

$$dJ_t' = \begin{pmatrix} dX_t^{(1)} , \ldots, dX_t^{(N)} \end{pmatrix} \begin{pmatrix} dX_t^{(N+1)}, \ldots, dX_t^{(M)} \end{pmatrix},$$

(4)

with assumption of $M > N$. $dX_t^{(1)}, \ldots, dX_t^{(N)}$ part of the vector $dJ_t$ corresponds to the idiosyncratic shocks. The remaining $dX_t^{(N+1)}, \ldots, dX_t^{(M)}$ components correspond to the common shocks as they are observed at an empirically relevant frequency.$^1$

Every $X_t^{(m)}$, with $j = N + 1, \ldots, M$, drives a particular common pattern among a sub-set of time series. Further, we assume that at every time $t$, every column of the matrix process $U_t^C$ has at least two fixed components with almost surely non-zero value. This ensures that $U_t^C$ corresponds to the common jumps and is not contributing to the idiosyncratic shocks.

In the next section, we assume a specific form of the price jump arrival process, which allows us to link the number of price jump arrivals to the exogenous low-frequency factors and establish the mixed-frequency framework.

---

$^1$Note that we are using a continuous time framework to describe stylized facts as they are observed at a high-frequency domain, the presence of common jumps at a rate which exceeds a natural rate implied by the given sampling frequency. We assume perfect synchronicity in the jump arrivals, as in Aït-Sahalia et al. (2009) and Bollerslev et al. (2013). However, the framework is broad enough to encompass alternative multidimensional jump specifications with finite activity, as suggested in Barndorff-Nielsen and Shephard (2004, 2006), Andersen et al. (2012), Lee and Mykland (2008).
2.2. Linking the High-Frequency Price Jump Arrivals to Low-Frequency Factors

We explicitly link the high-frequency price jump arrival processes to the low-frequency domain of the economic activity indicators. We consider the $M$-dimensional arrival vector $J_t$ at time $t$ from equation (4) to be composed of mutually independent, doubly stochastic Poisson processes given as

$$X^{(m)}_t = \int_0^t dX^{(m)}_t,$$

with $j = 1, \ldots, M$, each driven by the instantaneous stochastic intensity $d\Lambda^{(m)}_t = E \left[ dX^{(m)}_t | \mathcal{I}_{t-} \right]$, conditional on the information set $\mathcal{I}_{t-}$, which denotes the information available up to time $t-$. The integrated stochastic intensity is then given as $\Lambda^{(m)}_t = \int_0^t d\Lambda^{(m)}_t$. The $M$-dimensional intensity process $\Lambda_t$ can in general be mutually correlated; however, the price jump arrivals are drawn independently. The choice of the doubly stochastic Poisson processes allows us to explicitly link the high-frequency process with the low-frequency process driving the information set $\mathcal{I}_t$.

To establish the formal mixed-frequency link between the intensity process and the underlying macro-factors, let us focus on $d\Lambda^{(m)}_t$ corresponding to the arrival process $X^{(m)}_t$. We assume a specific functional form of the intensity process as suggested by Lee (2012). The continuous-time extension of such an instantaneous intensity process at time $t$ conditional on the information set $\mathcal{I}_{t-}$ takes the form

$$d\Lambda^{(m)}_t = \frac{1}{\alpha_0 + \alpha_1 Z_t} dt,$$

where $Z_t$ indicates a predictor variable of the price jump arrivals at time $t$, which incorporates all the information available up to time $t$, i.e., based on the information set $\mathcal{I}_{t-}$. Following Lee (2012), we consider the predictor variable to be a binary variable capturing the information about the presence of shock, revealed to markets at time $t$. We denote the random times when the shocks arrive as $t_1, \ldots, t_L \in [0, T]$, with $L$ being the random integer. The instantaneous intensity with the predictor variable $Z_t$ has functional properties similar to Dirac $\delta$-function, $\frac{1}{\alpha_0 + \alpha_1 Z_t} \sim \sum_{l=1}^L \delta (t - t_l)$.

The definition of the intensity process can therefore be rewritten as

$$\int_0^T \Lambda^{(m)}_t dt = \int_0^T \frac{1}{\alpha_0 + \alpha_1 Z_t} dt + \sum_{l=1}^L \int_0^T \frac{1}{\alpha_0 + \alpha_1} \delta (t - t_l) dt.$$

Thus, the expected number of price jump arrivals of the process $X^{(m)}_t$ in the time interval $[0, T]$ is
where $1_{Z_t}$ is an indicator function taking value of 1 if a signal is present at time $t \in [0, T]$, and 0 otherwise, $L$ is a random integer, $c_0 = T/\alpha_0$ and $c_1 = 1/(\alpha_0 + \alpha_1)$.\footnote{We may extend the predictor variable $Z_t$ as a $P$-dimensional vector. The corresponding formula (6) would be a combination of sums and Dirac $\delta$-functions capturing all cross-terms between the signals for different $Z_t^{(p)}$’s.} Such an intensity results in the implied instantaneous probability for an occurrence of a price jump at time $t_l$ as $P \left( \text{jump} \in dt_l \right) = d\Lambda_{t_l} > 0$ as $dt \to 0$, while for all other times $t \neq t_l$, the implied instantaneous probability results in $P \left( \text{jump} \in dt \right) \to 0$ as $dt \to 0$. Such a set-up corresponds to the doubly stochastic Poisson process and is based on the stochastic intensity employed by Lee (2012) to explain the drivers for the price jump arrivals in high-frequency DJIA equities. In addition, mutual correlations between the different components $d\Lambda_{t_l}^{(m)}$ come in a straightforward manner through the dependence on the set of shared predictors.

Equations (6) and (7) explicitly link the high-frequency price jump arrival process as a function of the predictor variables and the low-frequency number of price jump arrivals for a given period. The number of price jump arrivals depends on the number of shocks arriving to the economy. It is in fact impossible to control for all these shocks explicitly as there is huge diversity of information arriving to the markets. Instead, we model the number of shocks for a given period as a function of the prevailing low-frequency exogenous factors.

We model the link between the realized number of high-frequency price jumps and the low-frequency exogenous factors as follows: We split the time interval $[0, T]$ into $K$ equidistant sub-periods $[T_{k-1}, T_k]$, with $0 = T_0 < \cdots < T_K = T$. The total number of shocks during each sub-period, $L_k$, is then assumed to be a function of the low-frequency exogenous factors, $f \left( \Upsilon_k^{(1)}, \ldots, \Upsilon_k^{(S)} \right)$, where $\Upsilon_k^{(s)}$, with $s = 1, \ldots, S$, is the value of the s-th exogenous factor during the period $[T_{k-1}, T_k]$. Further, we consider each $[T_{k-1}, T_k]$ to correspond to a calendar month and $f \left( \Upsilon_k^{(1)}, \ldots, \Upsilon_k^{(S)} \right)$ to be a linear function. Under such setting, the equation (7) gives the following specification linking the two mixed-frequency domains
\[ N_{[T_{k-1}, T_k]}^{(m)} = \beta_0 + \sum_{s=1}^{S} \beta_s \gamma_k^{(s)} + \varepsilon_k, \]  
\[ (8) \]

where \( N_{[T_{k-1}, T_k]}^{(m)} \) represents a total number of price jump arrivals from the process \( X_t^{(m)} \).

In the next section, we define the measure of commonality and the measure of multiplicity, which allow us to assess the degree of association between high-frequency price jump arrivals for a set of \( N \) assets. We then use the mixed-frequency framework to link the two measures to low-frequency market factors.

2.3. Measures: Commonality and Multiplicity

We deal with the \( N \)-dimensional portfolio aiming to describe the commonality in the price jump arrival processes for a sub-period \([T_{k-1}, T_k] \subset [0, T] \). For that reason, we introduce two measures.

The first one is the measure of commonality defined as the ratio of common price jump arrivals to all the arrivals from the portfolio perspective, which asserts the probability of a price jump arrival observed at any asset to be accompanied by a jump at other asset(s):

\[ Q_{[T_{k-1}, T_k]} = \frac{N_{[T_{k-1}, T_k]}^{(2+)}}{N_{[T_{k-1}, T_k]}}, \]
\[ (9) \]

where \( N_{[T_{k-1}, T_k]} \) corresponds to the number of all unique price jump arrival times, and \( N_{[T_{k-1}, T_k]}^{(2+)} \) is the number of all unique common price jump arrival times.

In particular, \( Q_{[T_{k-1}, T_k]} \) measures the ratio of the aggregate intensity of the common Poisson processes to the overall aggregate intensity of all Poisson processes in the \( N \)-dimensional process \( Y_t \). The measure of commonality takes values \( Q_{[T_{k-1}, T_k]} \in [0, 1] \), where \( Q_{[T_{k-1}, T_k]} = 0 \) denotes the case when all arrivals are idiosyncratic, while \( Q_{[T_{k-1}, T_k]} = 1 \) corresponds to the case when idiosyncratic arrivals are completely absent. In the Appendix C, we show the link to Pearson correlation coefficient for \( N = 2 \).

The second measure is the measure of multiplicity, which estimates the average multiplicity of every arrival across portfolio and thus asserts how many assets will jump on average, given that we observe a jump at any of the asset:
The Measure of Multiplicity. For an \( N \)-dimensional process \( Y_t \) specified by (1) in the time interval \( [T_{k-1}, T_k] \), the measure of multiplicity is defined as

\[
Q^\mu_{[T_{k-1}, T_k]} = \frac{\sum_{j=1}^N N J^{(j)}_{[T_{k-1}, T_k]}}{N_{[T_{k-1}, T_k]}},
\]

where \( N J^{(j)}_{[T_{k-1}, T_k]} \) corresponds to the number of all unique price jump arrival times for the asset \( j \).

For a portfolio of \( N \) assets, the measure of multiplicity takes values \( Q^\mu_{[T_{k-1}, T_k]} \in [1, N] \), where \( Q^\mu_{[T_{k-1}, T_k]} = 1 \) occurs only in the case of idiosyncratic arrivals, while \( Q^\mu_{[T_{k-1}, T_k]} = N \) for the portfolio wide arrivals.

Following the set-up for linking the number of price jump arrivals and the low-frequency factors, we proceed as follows: We split the time interval \([0, T]\) into \( K \) sub-periods \([T_{k-1}, T_k]\) and for each period, we calculate \( Q_{[T_{k-1}, T_k]} \) and \( Q^\mu_{[T_{k-1}, T_k]} \). Each of the two measures is based on the particular combinations of the number of price jump arrivals, and therefore following the specification (8), we link the measures, \( Q_{[T_{k-1}, T_k]} \) and \( Q^\mu_{[T_{k-1}, T_k]} \), to the exogenous low-frequency factors, \((\Upsilon^{(1)}_k, \ldots, \Upsilon^{(S)}_k)\). In particular, we consider a linear approximation

\[
Q_{[T_{k-1}, T_k]} = \beta_0^q + \sum_{s=1}^S \beta_s^q \Upsilon^{(s)}_k + \varepsilon^q_k,
\]

\[
Q^\mu_{[T_{k-1}, T_k]} = \beta_0^{\mu} + \sum_{s=1}^S \beta_s^{\mu} \Upsilon^{(s)}_k + \varepsilon^{\mu}_k.
\]

The measures of commonality and multiplicity capture the common properties of price jump arrivals, where in a first step the high-frequency information extracted from the time series is aggregated to a monthly level. In a second step, this information is regressed on low-frequency factors. The two measures provide an overall picture of jumps commonality without identifying the contribution of each asset. For that reason, we follow the co-feature framework introduced by Engle and Kozicki (1993), and define the concept of co-arrivals as a linear combination of price jump arrivals which eliminates them in the aggregate index.
2.4. Co-arrivals

Let us consider a map

\[ Y_t^{(j)} \to j_t^{(j)} = \begin{cases} 
1 & \text{if price jump is present} \\
0 & \text{otherwise} 
\end{cases}, \quad j = 1, \ldots, N \]  

(13)

which maps the \( N \)-dimensional log-price process \( Y_t \) into an \( N \)-dimensional indicator process \( j_t \).

The Co-arrivals. For \( N \)-dimensional log-price process \( Y_t \) from (1) mapped into the vector \( j_t \) using the mapping (13) in the time interval \([T_{k-1}, T_k] \subset [0, T]\), the co-arrival is defined as the non-zero linear combination \( w' = (w^{(1)}, \ldots, w^{(N)}) \), such that \( N_{T_{k-1}, T_k}^{(w)} \) is minimized, where \( N_{T_{k-1}, T_k}^{(w)} = \sum_c 1 (w' j_c \neq 0) \) and sum runs over all non-zero elements of the indicator time series \( j_t \). The vector of weights, \( w \), is the co-arrival vector.

The notion of co-arrivals is thus defined as the solution to the minimization problem

\[ w = \arg \min_{\tilde{w}} \sum_c 1 \left( \sum_{n=1}^{N} \tilde{w}^{(n)} j_c^{(n)} \neq 0 \right). \]  

(14)

The interesting case to explore occurs when all price jump arrivals are eliminated, i.e. \( N_{T_{k-1}, T_k}^{(w)} = 0 \). Such a situation can be tested using the following rank test

\[ H_0 : \text{rank} (J) = N , \quad H_A : \text{rank} (J) < N , \]  

(15)

where the matrix \( J \) is composed of the realization of the indicator process \( j_t \) over the time interval \([T_{k-1}, T_k]\), keeping only the non-zero elements. Due to the finite activity of the arrival processes, there is almost surely a finite number of non-zero columns in the matrix \( J \).

The notion of co-arrivals allows further understanding of the link between idiosyncratic and common price jump arrivals and of how complex the common structure in the price jump arrival is. In particular, the presence of idiosyncratic price jump arrivals for each asset implies that the solution \( N_{T_{k-1}, T_k}^{(w)} \) cannot be reached. The same conclusion is obtained when there are too many different common price jump arrivals, i.e., the rank of the matrix \( U_t^C \) in (4) is full. In addition, the non-zero elements of the co-arrival vector
suggest which assets tend to have close price jump arrival properties, in terms of being both the smallest and the closest at the same time.

Based on the notion of co-arrivals, we define two additional measures for every co-arrival vector: the index and the grade. We proceed in parallel to (9) and (10): We split the time interval $[0, T]$ into $K$ sub-periods $[T_{k-1}, T_k]$ and calculate the two measures.

The first measure, the index, evaluates how successful the co-arrival vector is in eliminating the jump arrivals from the composite index.

The Index. For an $N$-dimensional process $Y_t$ specified by (1) in the time interval $[T_{k-1}, T_k]$ and the co-arrival vector $w$, the index is given by

$$i_{w,[T_{k-1},T_k]} = \frac{N[T_{k-1},T_k] - N^{(w)}[T_{k-1},T_k]}{N[T_{k-1},T_k]},$$

(16)

where $N^{(w)}[T_{k-1},T_k]$ corresponds to the number of all unique price jump arrival times where $w$ is a solution to problem (14).

The index can be interpreted as a measure of the ratio of the aggregate intensities of the stochastic Poisson processes, which are removed from the composite index, to the overall aggregate intensities of all Poisson processes involved. The index takes values $i_{w,[T_{k-1},T_k]} \in [0, 1]$, with $i_{w,[T_{k-1},T_k]} = 0$ corresponding to the case of no elimination of arrivals captured by the co-arrival vector, while the case of $i_{w,[T_{k-1},T_k]} = 1$ specifies the full elimination of arrivals by the co-arrival vector $w$. If the null of (15) is rejected, then $i_{w,[T_{k-1},T_k]} < 1$. This measure considers both idiosyncratic and common arrivals, as defined in (2).

The second measure, the grade, quantifies the elimination of the common jump arrivals from the composite index.

The Grade. For an $N$-dimensional process $Y_t$ specified by (1) in the time interval $[T_{k-1}, T_k]$ and the co-arrival vector $w$, the grade is given by

$$g_{w,[T_{k-1},T_k]} = \frac{N^{(2+)}[T_{k-1},T_k] - N^{(w;2+)}[T_{k-1},T_k]}{N^{(2+)}[T_{k-1},T_k]},$$

(17)

where $N^{(w;2+)}[T_{k-1},T_k]$ corresponds to the number of all unique common price jump arrival times where $w$ is a solution to (14).
The grade can be interpreted as a ratio of the aggregate intensity of the common Poisson processes, which are not present in the composite index, to the overall aggregate intensity of all common Poisson processes in the $N$-dimensional process $Y_t$. The grade focuses solely on the common arrivals in (2) and ignores the idiosyncratic arrivals. The grade takes values $g_{w,[T_{k-1},T_k]} \in [0,1]$, with $g_{w,[T_{k-1},T_k]} = 0$ corresponding to the case where none of the co-arrivals was eliminated by the vector $w$, while $g_{w,[T_{k-1},T_k]} = 1$ indicates that all co-arrivals were eliminated in the composite index. Both measures therefore provide additional information on how many (co-)arrivals were eliminated due to the co-arrival vector $w$.

So far, we have exploited the commonality in the price jump arrivals. We now also consider the magnitude of price jumps and introduce the notion of co-jumps.

2.5. Co-Jumps

Consider a statistic, $\hat{G}$-statistic, which for a given univariate price process satisfying the assumptions imposed in (1), allows us to test the null of no price jumps over any time interval $[T_{k-1},T_k] \subset [0,T]$. In particular, we say that asset $Y_t^{(j)}$ has a price jump(s) in the time interval $[T_{k-1},T_k] \subset [0,T]$ if the null hypothesis of the $\hat{G}$-statistic is rejected. The $N$-dimensional process $Y_t$ is closed under the linear combination in terms of its properties and therefore the $\hat{G}$-statistic can be applied for any linear combination of $Y_t$ as well.

Given the properties of the $\hat{G}$-statistic, we define co-jumps as follows.

The Co-jumps. For $N$-dimensional log-price process $Y_t$ from (1) in the time interval $[T_{k-1},T_k] \subset [0,T]$, the co-jump is defined as the non-zero linear combination $w' = (w^{(1)}, \ldots, w^{(N)})$, such that the null hypothesis of the $\hat{G}$-statistic for $w'Y_t$, denoted as the $\hat{G}^{(w)}$-statistic, cannot be rejected. The vector $w$ is called the co-jump vector.

In the case of complex common price jump patterns or the presence of idiosyncratic price jumps for each component of $Y_t$, the null may be rejected for every $w$. In such a case, we proceed to identify co-jumps as a solution to an optimization problem analogous to (14), where the objective function is the $p$-value of the $\hat{G}^{(w)}$-statistic to be minimized as the null hypothesis of the test is no price jumps.

---

3Example of such a statistic was introduced by Barndorff-Nielsen and Shephard (2006) and Huang and Tauchen (2005).
For a given set of assets, the notion of co-jumps thus complements the commonality of price jump arrivals and provides additional information about the role of the variance-covariance process as well as the magnitude of price jumps. In addition, it provides a portfolio perspective and gives information on how diverse the multivariate price jump process is in the portfolio. It also suggests whether the price jumps can be diversified out of the portfolio, given the covariance process.

In the next section, we describe the data used in this paper together with the price jump tests and use our theoretical framework to identify commonalities in the European government bond yields observed at high-frequency and link them to the economy indicators.

3. Empirical Results

3.1. Data

In this section, we describe the main characteristics of government bond yields used in this study. We then review the macro-factors, news announcements and bond auctions employed as the explanatory variables. The details of the data selection and the overview statistics for each set of variables can be found in the Internet Appendix.

3.1.1. Yields

We consider 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.\footnote{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 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). We remove outliers by applying a filter which is very close to the one proposed in Brownlees and Gallo (2006) and further elaborated by Barndorff-Nielsen et al. (2011, p. 156).
3.1.2. Macro-factors

We employ two real economy indicators: unemployment and industrial production, and a forward looking indicator, the economic sentiment (ES). Our choice is motivated by the existing literature, for example Mody (2009) and Aizenman et al. (2013); industrial production is often found to be particularly relevant to describe asset behaviour in a number of studies, for example Schwert (1989), Ludvigson and Ng (2009) and Lustig et al. (2014). Data come from Eurostat.

3.1.3. 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 $\sigma$ and actual value of the release denoted as $A$. Data were collected from Bloomberg. In our application, we adopt the standard surprise measure $\zeta$ defined as

$$\zeta = \frac{(A - E)}{\sigma}. \quad (18)$$

The Internet Appendix reports the full description of the macro-announcements.

3.1.4. 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

---

5Economic sentiment is provided by Eurostat and is a weighted index comprising five sectoral confidence indicators: industrial confidence, services confidence, consumer confidence, construction confidence, and retail trade confidence indicators.
Figure 1: Number of monthly arrivals.

Note: The figure reports the total number of arrivals per month for each country.

they not only correspond to the maturity of the bonds analysed but they even represent
the most relevant ones.

In this paper, in order to illustrate the methodology introduced in Section 2, first, we
identify price jump arrivals using the Lee and Mykland (2008) test (LM henceforth), where
the volatility is adjusted by its intraday pattern as suggested by Andersen and Bollerslev
(1998). 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
details are presented in the Internet Appendix.

3.2. Frequency-specific Factorization

First, we empirically validate the frequency-specific framework (3) and (4) and show
that idiosyncratic and common price jump arrivals have different drivers. Figure 1 reports
the number of arrivals per month for the individual countries. It shows that in the after-
math of the Lehman Brothers collapse, the number of jump arrivals significantly increases
for all countries except Germany. Then, following the tranquil period from June 2009
to April 2010, the overall number of arrivals increases again as the European debt crisis
emerged with the Greek bailout in May 2010 in particular for Italy and Spain, the two
peripheral countries in the sample.

Figure 2 depicts the proportion of idiosyncratic arrivals, defined in (4), for each month
and country. For instance, for Italy we notice that in June 2007, 20% of jump arrivals in
the Italian bonds are idiosyncratic, while the remaining 80% occurred jointly with arrivals in other countries. For Belgian bonds, the highest ratio of idiosyncratic arrivals is achieved in February 2012, with 82.7% of arrivals being idiosyncratic. It is notable that for each country there is at least one month without any idiosyncratic jump.

Another important feature that Figure 2 reveals is that the idiosyncratic and common jump arrivals have different dynamics. In particular, for the case of Germany, we see a prolonged period with no or few idiosyncratic arrivals, persisting up to the end of 2010. The situation changes dramatically with the deepening of the European debt crisis, when the proportion of idiosyncratic arrivals increases substantially. All other countries experience changes in dynamics as well, though with a pronounced country-specific pattern: for instance Italy in mid-2009, and at the peak of the European debt crisis around the end of 2011, with the rising uncertainty in correspondence to the political downturn. In addition, we note that the increasing share of idiosyncratic arrivals around the end of 2011 and start of 2012 occurs in correspondence to downgrades from rating agencies; for instance, S&P downgraded Belgium on the 25th November 2011, Italy on the 19th September 2011 and then again on the 12th January 2012, Spain on the 12th October 2011 and on the 12th January 2012, and France on the 13th January 2012.

Then, we test specification (8), where we focus on the difference between the idiosyncratic and common arrivals. For every asset \( j \), the common arrivals are defined in this case as all price jump arrivals from \( dX_t^{(N+1)}, \ldots, dX_t^{(M)} \) present at asset \( j \).

We estimate the following specification

\[
N_k^{(j,\omega)} = \alpha_0^{(j,\omega)} + \sum_{s=1}^S \alpha_s^{(j,\omega)} \Upsilon_k^{(j, s)} + \varepsilon_k^{(j,\omega)}, \quad j = 1, \ldots, 6, \omega = \{\Sigma, I, C\},
\]  \hspace{1cm} (19)

where \( j \) indexes the countries and \( \omega = \{\Sigma, I, C\} \), where \( \Sigma \) refers to the total number of arrivals for country \( j \), \( I \) is the number of idiosyncratic arrivals for country \( j \), and \( C \) the number of common arrivals for country \( j \) with any other country, respectively; \( N_k^{(j,\omega)} \) is the number of \( \omega \)-type of price jump arrivals for month \( k \) for country \( j \), and \( \Upsilon_k^{(j, s)} \) is the set of \( S \) country-specific covariates for every month \( k \). We estimate the system (19) using the SURE estimation method to take into account the contemporaneous correlation across countries.
Figure 2: Proportion of idiosyncratic arrivals per month for each country.
In particular, we consider two possible sources of price jump arrivals: the overall macroeconomic factors and the amount of surprising information which the economy absorbs during the particular time window. We use three different specifications:

\[ N_k^{(j;ω)} = \alpha_0^{(j;ω)} + \alpha_1^{(j;ω)} dU_k^{(j)} + \alpha_2^{(j;ω)} dI_k^{(j)} + \alpha_3^{(j;ω)} dE_k^{(j)} + \varepsilon_k^{(j;ω)} , \]  
\[ N_k^{(j;ω)} = \beta_0^{(j;ω)} + \beta_1^{(j;ω)} L_k^{(US)} + \beta_2^{(j;ω)} L_k^{(EA)} + \beta_3^{(j;ω)} L_k^{(j)} + \beta_4^{(j;ω)} L_{\bar{y},k} + \varepsilon_k^{(j;ω)} , \]  
\[ N_k^{(j;ω)} = \gamma_0^{(j;ω)} + \gamma_1^{(j;ω)} Z_k^{(US)} + \gamma_2^{(j;ω)} Z_k^{(EA)} + \gamma_3^{(j;ω)} Z_k^{(j)} + \gamma_4^{(j;ω)} Z_{\bar{y},k} + \varepsilon_k^{(j;ω)} . \]  

Specification (20) contains three country-specific covariates: the unemployment \( UE \), industrial production \( IP \) and economic sentiment \( ES \), all of them expressed as a monthly percentage change. This specification captures the effect of the prevailing macroeconomic environment on the intensity of jump arrivals. Specification (21) contains four covariates: \( L^{(US)} \), \( L^{(EA)} \) and \( L^{(j)} \) being the number of macro-announcements with large surprise originating from the US, the Euro area and the \( j \)-th country, respectively; \( L_{\bar{y}}^{(j)} \) represents the number of government bond auctions held in the \( j \)-th country with large surprise. An announcement is considered to carry a large surprise to the market if \(|ζ| > σ(ζ)| \), where \( ζ \) is specified in (18), while in the case of auctions, \( ζ \) is defined as the difference in the average yield between current and previous 10-year auctions. This specification captures the effects of large surprises on the integrated jump arrival intensity. Specification (22) contains four covariates: \( Z^{(US)} \), \( Z^{(EA)} \) and \( Z^{(j)} \) being the number of surprises from macro-announcements originating from the US, the Euro area, respectively, and the \( j \)-th country; \( Z_{\bar{y}}^{(j)} \) represents the number of surprises from auctions held in the \( j \)-th country. We define the number of surprises as

\[ Z_k^{(j)} = \sum_{t \in k} \mid S_t^{(j)} \mid , \]  

where the sum covers all individual surprises \( S^{(j)} \) for all macro-announcements or auctions originating from the country/region \( j \) in a given month \( k \). This specification captures the effects of all surprising announcements on the integrated intensity of jump arrivals.

Table 1 reports the estimates of the models (20)-(22). The contribution of the three groups of covariates varies for idiosyncratic and common arrivals, indicating the pres-
ence of different mechanisms underlying the drivers of arrival processes. This constitutes further evidence of the need to distinguish the two arrival processes, as our proposed frequency-specific factorization approach allows. For instance, results in Table 1 show that macroeconomic announcements from US are more associated with common arrivals than idiosyncratic ones. In addition, the higher statistical significance of the loading coefficients associated to large surprises, $L_{US}$, $L_{EU}$ and $L^{(j)}$ relative to those associated to the simple release of macro-announcements, $Z_{US}$, $Z_{EU}$ and $Z^{(j)}$, suggests that it is the surprise brought to the market rather than the pure release of an announcement causing price jumps and co-jumps. As far as government bond auctions are concerned, results in Table 1 indicate that although auctions held in distressed countries, like Italy and Spain, determine idiosyncratic jumps in these countries they do no impact on the other countries.\footnote{We may also estimate a model where all covariates in (20)-(22) enter the specification simultaneously. However, we do not do this because the limited amount of observations available.} Finally, we see that improved macroeconomic conditions, in terms of employment and industrial production, determine a lower number of jumps as witnessed by the negative coefficients associated to these two macroeconomic factors.

Thus far, we have provided empirical evidence that frequency-specific factorization (2) is a useful approach to mimic the arrival processes. In particular, the empirical evidence suggests that the idiosyncratic and common price jump arrivals tend to have different dynamics and are caused by different factors. In the next section, we calculate the measure of commonality and the measure of multiplicity and link them to the exogenous macro-factors and surprises.

3.3. Commonality

In Figure 3, we depict the measure of commonality and the measure of multiplicity as defined in (9) and (10), respectively. The upper display of Figure 3 plots the measure of commonality $Q$ on a monthly basis. $Q$ takes values between 0.14 and 0.68. In September 2009 it reaches 0.68, which means that 68% of the arrival times were common, or, conversely, only 32% of jump arrival times in that month contains purely idiosyncratic arrivals. Further, we see a significant structural break in mid-2009, after which the measure of commonality reaches its highest value. The structural break in the measure of
Table 1: Modelling monthly arrivals.

<table>
<thead>
<tr>
<th>Country</th>
<th>$N_k^{(1, 1)}$</th>
<th>$N_k^{(2, 1)}$</th>
<th>$N_k^{(3, 1)}$</th>
<th>$N_k^{(1, 2)}$</th>
<th>$N_k^{(2, 2)}$</th>
<th>$N_k^{(3, 2)}$</th>
<th>$N_k^{(1, 3)}$</th>
<th>$N_k^{(2, 3)}$</th>
<th>$N_k^{(3, 3)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_0$</td>
<td>3.21a</td>
<td>2.52a</td>
<td>2.59a</td>
<td>2.79a</td>
<td>1.37a</td>
<td>2.59a</td>
<td>2.93a</td>
<td>1.97a</td>
<td>2.44a</td>
</tr>
<tr>
<td>$E$</td>
<td>-3.57a</td>
<td>-5.22a</td>
<td>-1.97a</td>
<td>0.53</td>
<td>0.02</td>
<td>0.72b</td>
<td>-2.31a</td>
<td>-3.66a</td>
<td>-1.44a</td>
</tr>
<tr>
<td>$IP$</td>
<td>-0.25</td>
<td>-0.24</td>
<td>-0.18</td>
<td>-1.00b</td>
<td>-1.28c</td>
<td>-0.81b</td>
<td>-0.71c</td>
<td>-0.58</td>
<td>-0.47</td>
</tr>
<tr>
<td>$ES$</td>
<td>-0.92a</td>
<td>-1.02a</td>
<td>-0.78a</td>
<td>0.05</td>
<td>-0.04</td>
<td>0.09</td>
<td>-0.30</td>
<td>-0.17</td>
<td>-0.38c</td>
</tr>
<tr>
<td>$R^2$</td>
<td>52.35%</td>
<td>47.97%</td>
<td>31.59%</td>
<td>13.20%</td>
<td>6.53%</td>
<td>14.97%</td>
<td>43.75%</td>
<td>26.30%</td>
<td>27.19%</td>
</tr>
<tr>
<td>Panel B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_0$</td>
<td>2.37a</td>
<td>1.14b</td>
<td>2.24a</td>
<td>2.43a</td>
<td>0.69</td>
<td>2.25a</td>
<td>3.00a</td>
<td>1.66a</td>
<td>2.61a</td>
</tr>
<tr>
<td>$L_{US}$</td>
<td>0.09 b</td>
<td>0.09</td>
<td>0.07b</td>
<td>0.08b</td>
<td>0.12c</td>
<td>0.08b</td>
<td>0.02</td>
<td>-0.02</td>
<td>0.05</td>
</tr>
<tr>
<td>$L_{EA}$</td>
<td>0.04</td>
<td>0.12</td>
<td>-0.02</td>
<td>0.00</td>
<td>0.04</td>
<td>0.00</td>
<td>-0.01</td>
<td>0.10</td>
<td>-0.06</td>
</tr>
<tr>
<td>$L_y^{(1)}$</td>
<td>0.02</td>
<td>0.00</td>
<td>0.02</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>-0.04</td>
<td>-0.01</td>
<td>-0.02</td>
</tr>
<tr>
<td>$L_y^{(2)}$</td>
<td>0.28</td>
<td>0.58b</td>
<td>-0.03</td>
<td>-0.01</td>
<td>0.09</td>
<td>-0.01</td>
<td>0.10</td>
<td>0.43</td>
<td>-0.12</td>
</tr>
<tr>
<td>$R^2$</td>
<td>9.48%</td>
<td>11.20%</td>
<td>10.72%</td>
<td>12.02%</td>
<td>6.47%</td>
<td>12.59%</td>
<td>1.70%</td>
<td>8.05%</td>
<td>12.94%</td>
</tr>
<tr>
<td>Panel C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_0$</td>
<td>2.07a</td>
<td>0.52</td>
<td>2.19a</td>
<td>2.04a</td>
<td>0.39</td>
<td>1.98a</td>
<td>2.66a</td>
<td>1.01</td>
<td>2.35a</td>
</tr>
<tr>
<td>$Z_{US}$</td>
<td>0.03a</td>
<td>0.03a</td>
<td>0.02a</td>
<td>0.03a</td>
<td>0.03b</td>
<td>0.02a</td>
<td>0.01</td>
<td>0.01</td>
<td>0.02b</td>
</tr>
<tr>
<td>$Z_{EA}$</td>
<td>0.01</td>
<td>0.04a</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
</tr>
<tr>
<td>$Z_y^{(1)}$</td>
<td>-0.01</td>
<td>0.00</td>
<td>-0.01</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>-0.02</td>
<td>-0.02</td>
<td>0.00</td>
</tr>
<tr>
<td>$Z_y^{(2)}$</td>
<td>0.54a</td>
<td>1.05a</td>
<td>0.04</td>
<td>-0.01</td>
<td>0.13</td>
<td>-0.14</td>
<td>0.35b</td>
<td>0.91a</td>
<td>-0.04</td>
</tr>
<tr>
<td>$R^2$</td>
<td>20.98%</td>
<td>25.13%</td>
<td>15.63%</td>
<td>17.54%</td>
<td>9.06%</td>
<td>15.95%</td>
<td>10.91%</td>
<td>18.76%</td>
<td>14.91%</td>
</tr>
</tbody>
</table>

Note: $N_k^{(1, 1)}$ refers to the dependent variable denoting the total number of arrivals for a given country $j$ and month $k$; $N_k^{(2, 1)}$ is the dependent variable denoting the total number of idiosyncratic arrivals for a given country $j$ and month $k$; and $N_k^{(3, 1)}$ is the dependent variable denoting the total number of common arrivals for a given country $j$ and month $k$. All the variables are introduced in (19). In Panel A, we report specification in (20), which employs the rate of change for Employment ($E$), Industrial Production ($IP$), and Economic Sentiment ($ES$) for the $j$-th country, respectively. In Panel B, we report specification in (21), which employs the number of macro-announcements per month which brought to the market a large surprise, defined as $|\zeta| > SD(\zeta)$, where $\zeta$ is specified in (18): $L_{US}$, $L_{US}^{(1)}$, $L_{US}^{(2)}$, and $L_y^{(1)}$ denoting the number of macro-announcements with large surprise originating from the US, the Euro area, and the country $j$. $L_y^{(j)}$ denotes the number of surprising auctions held in the corresponding $j$-th country. In Panel C, we report specification in (22), which employs the total amount of surprise from macro-economic announcements per month, defined as $Z_y^{(j)} = \sum_{i \in k} \zeta_{i}^{(j)}$, with the sum running over all individual surprises $\zeta^{(j)}$ for all macro-announcements originating from the country/region $j$. The significance levels are defined as follows: $a$ for 1%, $b$ for 5%, and $c$ for 10%.
commonality further supports the hypothesis of the frequency-specific factorization because the portfolio-wide properties of the idiosyncratic and common jump arrivals vary differently over time. In addition, the figure depicts the shaded region which corresponds to [1-st, 99-th] centiles of the empirical distribution of $Q$ based on the 10,000 realizations of a Monte Carlo simulation, where the same number of price jump arrivals arrive independently. For each month, the observed $Q$ is out of the shaded region and therefore we may assess that $Q$ is taking a value significantly different from the Monte Carlo simulation based on the randomly arriving price jumps. In particular, $Q$ is significantly higher and thus the bonds tend to have significantly common price jump arrival processes.

The lower display of Figure 3 captures the measure of multiplicity $Q^\mu$ on a monthly basis, taking values between 1.21 and 3.33. First, the measure of multiplicity is highly correlated to the measure of commonality, with correlation coefficient equal to 0.95. The high correlation between the two measures suggests the presence of two significant clusters, as can also be seen in Figure 2, the first cluster is formed by Germany and the second by all remaining countries. Second, we notice an increase in both measures of commonality and multiplicity in the middle of 2009, which can be explained by the drop of idiosyncratic arrivals for all countries in the middle of 2009. In addition, both measures suggest that the overall correlation in the jump arrivals decreases during financial distress, as we can see for the cases of both the subprime crisis in the late 2008 and early 2009 and the most severe phase of the European crisis in 2011. For example, we observe that the average multiplicity of a jump arrival during the financial crisis is around 1.5, while this value raises in the aftermath of financial distress. These results are in contrast to the overall evidence reported in the literature, such as Forbes and Rigobon (2002), and Bekaert et al. (2005, 2009, 2011), who found that, during financial distress, the correlation among financial time series increases. The figure also contains the shaded region corresponding to [1-st, 99-th] centiles of the empirical distribution of $Q^\mu$ using the same Monte Carlo simulation set-up as for $Q$. The results also suggest that for each month, $Q^\mu$ is taking values significantly higher than the value based on the randomly arriving price jumps, and thus the price jump arrivals tend to overlap significantly.

To further understand the behaviour of the measure of commonality and the measure of multiplicity, we estimate (11) and (12), respectively. Given the high correlation between
Figure 3: The measures of commonality, $Q_t$, and multiplicity, $Q^\mu_t$, on monthly basis.

(a) The measure of commonality.

(b) The measure of multiplicity.

Note: The figure reports (a) the measure of commonality and (b) the measure of multiplicity defined in (9) and (10), respectively, for every month. In addition, the shaded region corresponds to [1-st, 99-th] centiles of the empirical distribution of $Q$ and $Q^\mu$, respectively, based on the 10,000 realizations of a Monte Carlo simulation, where the same number of price jump arrivals arrives independently.

the measure of commonality and the measure of multiplicity, we only report results for the measure of commonality, the ones for multiplicity are available upon request. In particular, we are interested in explaining the role of the arriving information and therefore we employ the following linear specification

$$Q_k = \alpha_0 + \tilde{\alpha}_0 DU_k + \sum_{q=1}^G \alpha_q Z_{k,t}^{(q)} + \sum_{j=1}^K \omega_j \bar{Y}_{j,k,t}^j + \varepsilon_t,$$

(24)

where $DU_t$ is a dummy step function, the set of covariates $Z_{k,t}^{(q)}$ represents the aggregate number of surprises, which hit the economy for a particular type of news announcement $q$ with $G$ being total number of different types of news announcements, $Z_{j,k,t}^{(j)}$ is the aggregate amount of surprise originating from a 10-years bond auction from a country $j$, and $\varepsilon_t$ is the error term assumed to be i.i.d. and following $N(0,\sigma^2)$, quantities $Z_{k,t}^{(q)}$ and $Z_{j,k,t}^{(j)}$ are defined in (23).

Finally, the identified structural break variable $DU_k$ is defined as

$$DU_k = \begin{cases} 
1 & \text{if } k \geq \text{July 2009} \\
0 & \text{otherwise}
\end{cases}.$$

We have used OLS with the Huber-White sandwich estimator to control for hetero-
skedastic errors to estimate the equation (24). The results of the estimated specification (24) reads

\[
Q_k = 0.000 + 0.149 \text{CPI}^{US} + 0.144 \text{GDP}_{Fin}^{US} - 0.080 \text{UniOfMichigan}^{US} \\
- 0.028 \text{PhilFEDIndex}^{US} + 0.054 \text{PPI}^{US} - 0.284 \text{MonthlyBulletin}^{EA} \\
+ 0.149 \text{GDP}^{Spain}_{Preliminary} + 0.117 \text{Unempl}^{Spain} + 0.039 \text{Unempl}^{Greece} \\
- 0.347 \text{Avg.Yield}^{Italy} + 0.352 \text{Avg.Yield}^{Portugal} + 0.168 DU_t
\]

where the \( R^2 = 0.700 \) and F-statistics being 9.115 with p-value of 0.000, values in square brackets denote the standard errors, and \(* * *\), \(* *\), and \(*\) refer to 1%, 5%, and 10% significance level, respectively.

The model suggests a large influence of the news announcements originating from the US on the commonality of jump arrivals. The Euro area Monthly Bulletin, providing information about monetary policy decisions together with a detailed analysis of the current economic situation and risks to price stability is the only Euro area-specific announcement variable significantly explaining the measure of commonality. In addition, the Spanish GDP and unemployment announcements together with Greek unemployment announcements are the only country-specific macro-economic news announcements which enter into the model. The relevance of the unemployment announcements for these two countries is interesting given that Greece and Spain are the two European countries which experienced that most severe increase of unemployment rate during the period considered in this paper, with Greece moving from 8.2% in June 2007 to 23.8% in June 2012 and Spain from 8.0% to 24.5%, respectively.

With respect to auctions, the change in the average yield at which Italy and Portugal succeeded in selling their 10-year bonds significantly explain commonality. For instance, Italy with its huge public debt and the need to refinance it justifies the attention paid by markets to the Italian government bond auctions. We also test several alternative specifications and aggregations across the types of news announcements, but no significant changes are noticed. In particular, country-specific macro-factors have no power to explain the measure of commonality. The full set of results of the alternative specifications are
not reported but are available upon request.

In the next section, we analyse the commonality in jumps using the notion of co-arrivals.

3.4. Co-arrivals

First, we test for the presence of co-arrivals by evaluating the null hypothesis in (15). For every month, the $J$ matrix is of full rank and therefore co-arrivals, corresponding to $N_k^{(w)} = 0$, do not exist. We therefore proceed to search for the co-arrival vectors $w$, which minimize the objective function (14).

Figure 4 reports the non-zero components of the global co-arrival vectors estimated for each month. Throughout the section, we use the following symbols: $\Diamond$, $\Box$, and $\triangle$, to denote the country/group of countries forming the non-zero co-arrival vector for the given month. For instance, for January 2008, we report the symbol $\Diamond$ for France and the Netherlands, meaning that the linear combination of these two countries formed the co-arrival vector. For July 2008, we note that the symbol $\Diamond$ is present for France and Germany, and the symbol $\Box$ for Spain and Germany: this implies the existence of two different co-arrival vectors. We also notice that there are few cases where three different co-arrival vectors exist, namely March 2010 and November 2010.

We note that up to the beginning of 2011, Germany was present in nearly all the co-arrival vectors, which can be explained by both a low number of arrivals per month for Germany, and by a large share of common arrivals with other countries. During the period of subprime crisis between mid-2008 and mid-2009, the co-arrival vectors consist of one country, Germany, suggesting a strong misalignment with respect to other countries since Germany does not show any idiosyncratic arrivals over that period. As the European debt crisis deepened at the beginning of 2011, the Netherlands is present in all co-arrival vectors. In particular, in the first half of 2011, we see a strong alignment between France and the Netherlands, which gives rise to co-arrivals for 10 consecutive months. Finally, Italian yields are very jumpy and have a large proportion of idiosyncratic arrivals, contributing to the co-arrivals in two instances only. A similar pattern is identifiable for the other two countries in the sample with a large number of idiosyncratic jumps, Spain and Belgium.

Next, for every identified co-arrival vector, we calculate the index and the grade as defined in (16) and (17), respectively. The upper display of Figure 5 depicts the index for
Figure 4: Components of the co-arrival vectors monthly.

Note: The figure reports the non-zero components of the global co-arrival vectors estimated for every month. The same symbol denotes the country/countries forming the non-zero co-arrival vector for the given month. For instance, for January 2008 the symbol ◇ for France and Netherlands indicates that two countries form the co-arrival vector.

every estimated co-arrival vector, where each point denotes the proportion of the overall amount of Poisson processes, which were removed from the portfolio built using the weights corresponding to a given co-arrival vector. In particular, for June 2007, the index takes a value of 0.911 meaning that 91.1% of the mass of all Poisson processes were removed from the portfolio given as a linear combination with weights corresponding to the co-arrival vector. Conversely, we can say that we were not able to remove 8.9% of all Poisson processes using that linear combination. The least effective linear combination with respect to the removal of Poisson processes was found for July 2009, where we did not remove 22.6% of Poisson processes. When we look at Figure 4, we see that for that particular month there are two different linear combinations: Belgium and the Netherlands, and Spain and Germany. The figure thus confirms that both linear combinations have the same efficiency in removing the Poisson processes. In addition, during the aftermath of the Lehman Brothers collapse, we observe a period where the index systematically increases suggesting that co-arrivals are able to remove a large amount of Poisson processes from the aggregate index.

The lower display in Figure 5 depicts the grade for every identified co-arrival vector, where each point denotes the proportion of the overall number of common Poisson processes, removed from the portfolio with weights corresponding to a given co-arrival vector. In particular, for June 2007 the grade takes value 0.680, which means that 68% of all
common Poisson processes were removed from the aggregate portfolio. The grade suggests that the co-arrival vectors were able to remove between 50% in October 2011 to a maximum of 94.1% in June 2011 of all common Poisson processes. Further, we see that different co-arrival vectors at a given month do not have necessarily the same efficiency with respect to removing the common Poisson processes. To illustrate this point, consider that July 2008 depicts the case where two different co-arrival vectors exist involving France and Germany, and Spain and Germany; the grade takes two different values, 0.72 and 0.68, respectively. This implies that a co-arrival portfolio built using the combination of France and Germany yields contains less co-arrivals with respect to a co-arrival portfolio of Spain and Germany yields. On the other hand, July 2009 depicts the case where two different co-arrival vectors have the same grade equal to 0.72.

The presence of the structural break is confirmed by the grade of the co-arrival vectors. In particular, the grade suggests the presence of two prevailing regimes with the turning point in August 2009, one month behind the structural break defined by $DU_k$. The average value of the grade is 0.75 in the first regime, and rises to 0.82 afterwards. This suggests a change in the market behaviour; in particular, that co-arrivals are able to remove more common jumps since the structural break.

In the next section, we analyse commonality in price jumps using both their arrival times and magnitudes.

### 3.5. Co-jumps

First, we test for the presence of co-jumps as the existence of a linear combination $u'Y_t$ such that the $G_t^{(w)}$-statistic cannot reject the null hypothesis for each month $k$, where each
Note: The figure reports the $G_n^{(w)}$-statistic estimated for each month. The 5% critical value to reject the null hypothesis of no jumps is equal to $-2.11$ and lies well below the plotted values.

month was sampled into 5-minute steps. In particular, we search for $w$ such that the $p$-value is maximized.

Figure 6 depicts the $G_n^{(w)}$-statistic for the identified co-jump vector implied by the $p$-value minimization. For example, in June 2007 the $G_n^{(w)}$-statistic takes value $$. The $G_n^{(w)}$-statistic ranges between $-3.77 \cdot 10^{-4}$ and $-2.97 \cdot 10^{-6}$. Thus, we can conclude that for every identified co-jump vector at each month $k$, we cannot reject the null of no co-jumps, given that the $G_n^{(w)}$-statistic in all instances is well below the critical value of $-2.11$, for the 5% critical value.

Figure 7 reports the individual components of the (weak) global co-jump vectors, which were used to depict the $G_n^{(w)}$-statistic in Figure 6. For example, in June 2007, the vector of weights reads $w_{IT} = 0.505$, $w_{FR} = 0.319$, $w_{ES} = 0.613$, $w_{BE} = 0.312$, $w_{NL} = 0.198$, and $w_{DE} = 0.359$. This means that the six countries contribute to the co-jump portfolio in that month as follows: Italy by 21.9%, France by 13.8%, Spain by 26.6%, Belgium by 13.5%, the Netherlands by 8.6% and Germany by 15.6%. The maximum individual component in absolute value is for Italy in January 2012, with $w_{IT} = 0.868$ in correspondence with a severe political crisis which lead to the new government with the specific task to restore the sustainability of the Italian public debt. On the other hand, the minimum contribution to the co-jump occurs for Germany in November 2011, with $w_{DE} = 0.018$, when peripheral countries experienced the highest increase in their government bond spreads accompanied
Note: The figure reports the country components of the co-jump vectors for each month.

by increasing concerns on their economic and financial conditions. Let us note that the
coop-jump optimal portfolio is estimated ex-post given the knowledge of realization of the
price process.

Figure 7 provides clear evidence that in most cases, all countries contribute to the
coo-jump vectors, as the values of the individual components are different from zero. The
components corresponding to Germany change sign between December 2010 and May
2011. This suggests either a change in the structure of the jump process, and/or a
change in the covariance structure of Germany with respect to all other countries. In
addition, the worsening of the European debt crisis explains the higher variance in the
absolute values of the global co-jump components. It is worth to note that we estimate
price jumps for each month and each country independently and that we reject the null
of no price jumps for every case.

3.5.1. The Role of the Diversification Effect

We now compare the $G_n^{(w)}$-statistic of the composite index of the co-jumps with the
one for the equally weighted portfolio. This allows us to shed some light on whether the

---

7In order to understand the change in sign for Germany, we report in the Internet Appendix the results
of the daily pair-wise correlation between yields using the consistent and efficient estimator using the
high-frequency covariance proposed by Aït-Sahalia et al. (2010). 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 corresponding to the period of change in the global co-jump vectors
as reported in Figure 7. 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.
Figure 8: $G_n^{(w)}$-statistic for “1/N” portfolio.

Note: The figure reports the $\log_{10}$ transform of the $G_n^{(w)}$-statistic for equally weighted, “1/N”, portfolio. $\theta$ indicates the 5% critical value to reject the null hypothesis of no jumps.

The presence of co-jumps can be explained by the diversification effect, as originally proposed by Bollerslev et al. (2008). The diversification effect corresponds to the case in which for any given sampling and any given portfolio of $N$ uncorrelated time series, the ratio of the magnitude of the idiosyncratic jumps with respect to the prevailing integrated variance in the equally weighted portfolio decreases with the size of the portfolio at the rate of $1/\sqrt{N}$. To this purpose, we construct an equally weighted composite index of yields and analyse the $G_n^{(w)}$-statistic for that. Figure 8 plots the $G_n^{(w)}$-statistic for such an index evaluated for every month. The value of the $G_n^{(w)}$-statistic ranges between $-1.112.11$ in April 2009 and $-7.52$ on July 2011. Since the 95% critical value to reject the null hypothesis of no jumps is $-2.11$, we reject the null hypothesis for every month and therefore any of the months that do not show the global co-jumps. For the sake of completeness, for all time series the weights are set to be equal to $w^{(j)} = w = 1/\sqrt{6}$, for all $j = 1, \ldots, N$. In addition, the minimum value of the “1/N” index is above any individual country specific $G_n^{(w)}$-statistic. The disappearance of jumps in the co-jumps cannot be thus entirely assigned to the diversification effect; the fact that there is a significant structure in price jumps—including both arrival times and magnitudes—across the countries.

In the Internet Appendix, we report the $G_n^{(w)}$-statistic for the three market portfolios based on the zero-coupon bond approximation and show that co-jumps are not present in all three cases.

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. We have performed a robustness check by estimating the measure of commonality and measure of multiplicity calculated at daily and weekly frequencies. First, there is evidence of that both measures are coarse, 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 as they are depicted in Figures 1. 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, as reported in Figure 2. 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. The Internet Appendix reports the details of the weekly and daily measures.

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.

4. Conclusions

In this paper, we proposed a novel frequency-specific framework to link the common features in price jumps with the low-frequency exogenous factors. We employed the framework to study the European government 10-year bond yields for Belgium, France, Germany, Italy, the Netherlands and Spain from 1 June 2007 to 31 May 2012 sampled at 5-minute frequency as a function of the monthly real economy indicators. In order to analyse the drivers of commonality, we introduced the measures of commonality and multiplicity and linked them to relevant macro-factors (unemployment, industrial production and economic sentiment), observed at monthly frequency, and to the aggregate monthly surprise carried by macro-announcements and government bond auctions. We explored the characteristics of bonds via the notion of co-arrivals, which identifies the structure of the vertical clusters in the common arrival of jumps, and co-jumps, which captures the presence of common jumps.
We provided evidence of statistically significant differences between idiosyncratic and common jump arrivals, showing that the idiosyncratic arrivals are more sensitive to financial distress. In particular, we found that the commonality feature of the jump arrivals is explained by news announcements from the US, the European Monthly Bulletin, the Spanish GDP and unemployment, and Greek unemployment. In both the subprime crisis and the European debt crisis in 2011, 10-year European yields show a low level of commonality as well as a low level of correlation in jump arrivals. This finding is in contrast to the evidence from asset returns literature (see, e.g., Bekaert et al., 2005), of a persistently higher correlation during distress periods as compared to more tranquil times. In particular, the measure of commonality during the financial crisis of 2008/2009 is of half value compared to the values in the aftermath of the crisis. This means that the probability of observing a common jump during the crisis is half the probability immediately afterwards. For instance, the measure of multiplicity indicated that during the 2008 crisis, if a jump occurred, then up to two countries were affected by the same jump, while in the aftermath of the 2008 crisis, if a jump occurred, more than three countries were affected. In correspondence with the European debt crisis, commonality decreased to the levels seen around the 2008/2009 crisis.

During the subprime crisis of 2008, the overall number of jump arrivals increased, which was not observed during the European debt crisis. Further, for the European debt crisis, we observed a significant change in the structure of common jumps in yields, providing clear evidence that Euro area was hit by country specific risks. Finally, from December 2010 to May 2011, the behaviour of German yields showed a completely different pattern compared to the other countries. In this period, we observed a significant change in correlation between German yields and yields from any other country in the sample. As the German bonds witnessed a sharpe increase in bid-ask spread, our findings serve as supportive evidence of the increase of the risk-awareness of investors. They tend to favour the German bonds serving as a safe haven. Therefore, the markets experienced a higher number of investors who were demanding to buy the German bonds while the supply side shrank.

The findings in this paper suggest additional developments. First, it will be interesting to extend the theoretical framework to identify commonalities in a fully high-frequency
framework, developing a multivariate counterpart of Lee (2012). Second, the empirical application can be extended by considering other maturities. This is part of an ongoing research agenda.

**References**


