News Flow, Web Attention and Extreme returns in the European Financial Crisis

Andreas Chouliaras*, Theoharry Grammatikos
Luxembourg School of Finance**

This version, 21/10/2013

Abstract

We examine the existence of stock market contagion effects among three groups of countries: the Euro-periphery countries (Portugal, Ireland, Italy, Greece, Spain), the Euro-core countries (Germany, France, the Netherlands, Finland, Belgium), and the major European Union -but not euro-countries (Sweden, UK, Poland, Czech Republic, Denmark), using daily stock market data from January 2004 till March 2013. Contagion effects for the tails of the marginal distributions are present for the Pre-crisis and the Euro-crisis periods within the Euro-periphery countries and from the Euro-periphery group to the Non-Euro and the Euro-core groups. We do not find a significant change in the contagion transmission mechanism when comparing the two periods, but for the Euro-crisis periods the extreme returns have a higher magnitude. Finally, we propose a connection between extreme stock market returns, the Web attention index and two News Flow indexes. Contagion effects are present for the Euro-crisis period, as the Euro-periphery Web Attention and News Flow variables significantly affect the probabilities of extreme bottom returns for the Euro-periphery, the Non-euro and the Euro-core groups. The effect is asymmetric in most of the cases since the Euro-periphery Web Attention and Pessimism Factors do not affect the probabilities of extreme top returns, with a few exceptions. The effects are positive, in other words more web attention and more bad news for the Euro-periphery in times of crisis is associated to higher probabilities of extreme bottom returns within and across groups. Granger-causality tests show that the News Pessimism and the News Relevance factor exhibit a two-way causality with the stock market movements while the Web Search Volume Index (SVI) one-way Granger-causes stock markets and extreme bottom returns in the three country groups.

JEL classification: G01, G14, G15, D83.

Keywords: Financial Crisis, Contagion, Web Attention, News Flow.

* Corresponding author. E-mail address: andreas.chouliaras@uni.lu
** Luxembourg School of Finance, 4, rue Albert Borschette, 1246 Luxembourg.

1. Introduction

The recent global financial crisis began as a crisis of the subprime mortgage loans in the United States of America in 2007. Since then, multiple waves of financial distress have hit the global financial markets. Since the beginning of 2010 the Euro area faces a severe financial crisis. What started off as a sovereign debt crisis from Greece soon transmitted to Portugal, Ireland, Cyprus and, at least partially, to Spain and Italy. Pretty soon it became clear for Europe that beneath the sovereign debt crisis surface there also existed a severe banking crisis. The evolution of this crisis and the transmission of financial distress from one country to another, with multiple asset classes being hurt (stock markets, bond yields, CDS spreads), make the case of financial contagion more pertinent than ever.[1]

A number of researchers investigate the recent eurozone financial crisis and its contagion effects. Contagion effects are found by studies that deal mainly with announcements of sovereign debt rating changes. For example, Afonso, Furceri, and Gomes (2012) use an event study approach, Arezki, Candelon, and Sy (2011) use a VAR framework with dummies to capture the effect of news announcements, and Missio and Watzka (2011) find contagion effects using dynamic conditional correlation models (DCC). Apart from rating announcements, Jiang, Konstantinidi, and Skiadopoulos (2012) use VAR models and find that news announcements affect but do not completely explain the magnitude of implied volatility spillovers. They interpret the inability of news and other events to fully explain the implied volatility spillovers as evidence of contagion. Mink and De Haan (2012) use an event study approach and find that news about Greece does not provoke abnormal returns on bank stock prices in 2010, while news about a Greek bailout does. Metiu (2012) employs a simultaneous equations model and examines the tails of bond yield distributions, an approach derived from the Extreme Value Theory and Value-at-Risk theories, and finds structural shift contagion effects for the crisis periods. Other papers, however, do not find contagion effects for the sovereign bond and the credit default swaps markets, for example, Caporin, Pelizzon, Ravazzolo, and Rigobon (2013) and Bhanot, Burns, Hunter, and Williams (2012).

In this paper, we investigate the transmission of the european financial crisis for three groups of countries: Two groups of eurozone countries, one for the Euro-core eurozone countries (Germany, France, Netherlands, Belgium, Finland) and one for the Euro-periphery eurozone countries (Portugal, Ireland, Italy, Greece, Spain), and finally a group for European Union (EU) but not euro countries (Sweden, UK, Poland, Czech Republic, Denmark). The main focus of our study is stock market spillovers between these groups. The correlations framework has been widely used by previous authors in related studies but there is no consensus in the research literature as to how to best define contagion when using that framework. As Bekaert et al. (2005) point out: "What is clear is that contagion is not simply revealed by increased correlation of market returns during a crisis period". Forbes and Rigobon (2001) claim that heteroskedasticity biases correlation tests for contagion. [2]

To avoid these problems we follow the approach of Bae, Karolyi, and Stulz (2003) who examine the coincidence of extreme return shocks across countries within a group and across groups. A number of authors have used this methodology. For example, Markwat, Kole, and van Dijk (2009), Lucey and

[1]The contagion literature is extensive (Allen and Gale (2000), Rigobon (2002), Kaminsky, Reinhart, and Vegh (2003), Pericoli and Sbracia (2003), Bekaert, Harvey, and Ng (2005), Forbes and Rigobon (2001), Aït-Sahalia, Cacho-Diaz, and Laeven (2010), Dungey, Fry, González-Hermosillo, and Martin (2005), Dungey et al. (2005), Corsetti, Pericoli, and Sbracia (2005)).

[2]By applying a correction they find no contagion for the 1997 Asian crisis. On the other hand, Corsetti et al. (2005) claim that the variance restrictions imposed by Forbes and Rigobon (2001) are "arbitrary and unrealistic". They find evidence for at least five countries facing contagion effects during the Hong Kong stock market crisis of 1997.

Sevic (2010), Christiansen and Ranaldo (2009), Gropp, Duca, and Vesala (2009). Another advantage of moving in line with multinomial logistic analysis as proposed by Bae et al. (2003) is that we can use control variables (covariates) in order to justify the characteristics of extreme returns. Furthermore, this approach allows us to study the crisis transmission across groups and also within groups. Since it is well accepted that the most vulnerable eurozone countries -the Euro-periphery group- were the most badly hit by the Euro-crisis, our main interest is to study for the crisis transmission from the Euro-periphery group to the other two groups (Euro-periphery vs. Euro-core, Euro-periphery vs. Non-euro) [3], [4]. We find that the Euro-periphery countries act as a source of contagion for the other Euro-periphery countries but also for the Euro-core and the Non-euro countries groups.

The original source of the markets' reactions and hence of the crisis spill-over must be traced to relevant information about the unfolding crisis. Previous researchers, as noted above, examine the impact of various official news announcements such as sovereign debt rating changes. Nevertheless, such specific news items give at best a partial and at worst a biased view of the impact of information on market prices since they do not reflect all available news and in many cases they are anticipated by market participants. The second objective of this paper, therefore, is to shed light on the impact of two main information sources on market prices and crisis transmission; the news flow imbedded in newspaper articles and newswires, and investors' attention about the crisis captured by their Web search activities. The study of news flow has attracted the researchers' interest rather recently with the advent of Data Mining and Sentiment Analysis techniques. Tetlock (2007) studies the interactions between a daily popular Wall Street Journal column and the Dow Jones returns by creating a media pessimism index using textual analysis based on the General Inquirer's Harvard IV-4 dictionary. Loughran and McDonald (2011) create their own dictionary with words that are better suited for a financial context, and associate the 10-K filing returns to the media tone, employing once more textual analysis based on their dictionary. One of the first works on the Web attention was by Varian and Choi (2009) who used the Google Trends Search Volume Index (SVI) to forecast economic indicators, such as car sales and unemployment claims[5]. Since then related work has started to appear in the finance literature. Da, Engelberg, and Gao (2011) examine the determinants of the SVI for stocks in the Russell 3000 and associate it with the returns in the index and the performance of IPO stocks. Da, Engelberg, and Gao (2010) use queries that may concern households, such as "recession", "unemployment", "bankruptcy" and create an investor sentiment index which they associate with the US market. Vlastakis and Markellos (2012) examine the effect of the SVI on the realized volatility (weekly sum of daily squared returns) of the 30 largest stocks traded in the NYSE and NASDAQ. Finally, Dzielinski (2012) study the S&P 500 returns along with the SVI results for the "economy" query.

We find that during the Euro-crisis a higher volume of Web search queries about the Euro-periphery crisis as measured via the Google Trends SVI is associated with higher probabilities of extreme bottom returns for all three country groups, while the probabilities of extreme top returns are not affected. As far as the News Flow is concerned, Pre-crisis no significant connection can be found between the

[3]Another study that uses this approach is Thomadakis (2012), but our main difference is that he considers the Eurozone countries as one group, thus not studying the within eurozone dynamics of the various subgroups, and he studies the interactions mainly with the USA, for the industrials sectors of the stock exchanges.

[4]One critique mentioned (e.g. by Ahlgren and Antell (2010)) is that Bae et al. (2003) arbitrarily pick the top and the bottom 5% from the sample of returns to examine the joint occurrence of extreme returns. This critique has indeed some merit but a choice of cutoff points is a necessary decision in order to proceed with this methodology and to study the tails of the marginal return distributions in order to see what happens in the presence of extreme returns. The results were found to be robust in the change of the percentiles.

[5]Google Trends can be found at: `http://www.google.com/trends/explore#cmpt=q`

news media Euro-periphery News Pessimism and Relevance factors and market returns. This however changes during the Euro-crisis period, with a higher Euro-periphery News Pessimism or Relevance being associated with higher probabilities of extreme bottom returns. Once more, in most of the cases the probabilities of extreme top returns are left unaffected by the News Pessimism factor.

The remainder of this paper is organized as follows. Section 2 presents the data. Section 3 presents the basic model and we explain how we study the crisis transmission within and across groups. Section 4 concerns the News Flow and the Web Attention, and Section 5 provides a set of robustness and alternative specifications. Section 6 is a conclusion.

2. The Data

The main area of study for this paper is the European Union area. Thus, we create three country groups: the Euro-periphery group contains the periphery eurozone countries (Portugal, Ireland, Italy, Greece, Spain), the second group, Euro-core, contains the core countries of the Eurozone (Germany, France, the Netherlands, Finland, Belgium), and the Non-euro group contains the major European Union (but not Euro) countries (Poland, Sweden, Czech Republic, UK, Denmark) [6]. We examine the period from 01/01/2004 till 13/03/2013 using daily financial data obtained from the Thomson Reuters Datastream. Our selection of countries for the Eurozone follows to a large extent the studies of Missio and Watzka (2011), Caporin et al. (2013), Bhanot et al. (2012), and Metiu (2012). Table 1 shows the summary statistics and correlation matrices of the percentage returns of the major stock market indices (Panel A and Panel B)[7]. We split our sample in three subperiods (Pre-crisis, US-crisis Euro-crisis) to be able to make comparisons between normal and abnormal times in the financial markets:

- the Pre-crisis period (from 1 January 2004 till 26 February 2007)
- the US-crisis period (27 February 2007, the date on which the Federal Home Loan Mortgage Corporation (Freddie Mac) announced that it will no longer buy the most risky subprime mortgages and mortgage-related securities, till the 7th of December, 2009) [8].
- the Euro-crisis period (8 December 2009, the day the Greek debt was downgraded by Fitch from A- to BBB+, with a negative outlook), till the end of our sample period (13 March 2013).

Insert Table 1 here

For the Pre-crisis period all groups of countries were having positive mean returns, consistent with the overall optimism in the financial markets. The best performing markets were firstly the Non-euro countries (+0.101%) followed by the Euro-periphery countries (+0.089%). Regarding the standard deviation we see that we have rather low values for all countries groups as this was a period of relative calmness for the financial markets. During the US-crisis period all country groups had a negative mean return. The Euro-periphery countries were the most badly hit with a mean return of

[6] We take the biggest five stock markets from each group using the market capitalisation ranking (as of 2011) from http://www.indexmundi.com/facts/indicators/CM.MKT.LCAP.CD/rankings

[7] All stock indices used are the Thomson Reuters Datastream indices created for each country

[8] We consider the 27th of February 2007 as the starting date of the US-crisis, as has been used in other papers as well, for example Grammatikos and Vermeulen (2012).

-0.074%, followed by the Euro-core countries which had a mean return of -0.049%, then the Non-euro with a -0.024%. Compared to the Pre-crisis period, the standard deviations have more than doubled, almost tripled for the Euro-periphery and the Euro-core groups. The descriptive statistics for the Euro-crisis period show that once more the Euro-periphery countries were the most severely affected from the financial crisis (-0.017%). The other two groups have positive mean returns for this period, indicating that they were able to better cope with the crisis. The standard deviations were lower than in the US-crisis period but still higher than the Pre-crisis period, especially for the Euro-periphery group. As far as the correlations are concerned the main remark is that they increased between the Pre-crisis and the crisis periods. The correlation between the Non-euro and the Euro-periphery group grew from the Pre-crisis value of 0.429 to 0.677 in the US-crisis period, then went down to 0.520 at the Euro-crisis period. The correlation between the Euro-periphery and the Euro-core followed a similar pattern, increasing from 0.579 Pre-crisis to 0.729 in the US-crisis, then declined to 0.646 in the Euro-crisis period, still much higher than the Pre-crisis period. The correlation between the Non-euro and the Euro-core group increased from 0.522 Pre-crisis to 0.727 in the US-crisis period, remaining at the elevated level of 0.734 in the Euro-crisis. Thus, for both crisis periods (US and Euro-crisis) the correlations are higher than what they were in the Pre-crisis period, and the most hit group is found to be the Euro-periphery group having negative mean returns in both crisis periods.

3. Extreme Returns

3.1. The Base Model

According to Bae et al. (2003) an extreme return, or exceedance, is one that lies below or above the lowest or the highest quantile of the marginal return distribution respectively. The original approach studies the coexceedances for the entire test period, taking as thresholds for extreme returns the 5th and the 95th percentiles. In our case, since we are mostly interested in the dynamics in the Pre-crisis and the Euro-crisis periods, we choose as thresholds the 10th and the 90th percentiles in order to have a sufficient number of observations, as in Boyson, Stahel, and Stulz (2010) (our findings are robust to the 5th and 95th percentiles). For each country we consider returns below the 10th percentile as extreme bottom returns and those above the 90th percentile as extreme top returns for this country.

This procedure is followed for all countries in all groups. Top extreme returns are treated separately from bottom extreme returns. The counts of joint occurrences of extreme returns, or coexceedances, within a group on a particular day are reported in Table 2. For each country we calculate the days for which it had an extreme (bottom or top) return separately. Then, the coexceedance count for each group and day is given as the sum of the exceedances for all countries that belong to that group for that specific day.

Insert Table 2 here

The left side of Table 2 presents bottom returns exceedances and the right side shows top returns exceedances. A coexceedance count of i units for bottom returns is the joint occurrence of i exceedances of bottom returns on a particular day for a specific group. By counting the total number of days with coexceedances of a given count and identifying which countries participate in those events and how often we have a good overview of the extreme returns for each country and group of countries.

We notice that out of the 10% lowest returns for all Euro-periphery countries the Greek stock market had the most days (106) on which it was the only country experiencing a bottom extreme return, followed by Ireland (56 days) and Portugal (37 days). A total of 54 days are reported for the Euro-periphery countries on which all of them experienced extreme bottom returns. On the other hand, from the top 10% distribution, on a total of 40 days all Euro-periphery countries experienced an extreme top return. For the Euro-core countries, 109 days are found on the bottom tails of each country that all five countries experienced an extreme return shock. There is a total of 91 days on which five Euro-core countries experienced top coexceedances. On 55 days all five Non-euro countries experienced bottom extreme returns, with the Czech Republic having the most days (84) as the only country experiencing an extreme bottom return. There were a total of 28 days on which all Non-euro countries had an extreme top return, with the Czceh Republic once more having the most days (95) with extreme top returns.

The methodology of Bae et al. (2003) can be applied to study two types of spill-over effects: within groups and across groups. In this paper we mainly focus on effects across groups.

3.2. Examining the presence of contagion effects

In order to capture contagion across countries within a group we consider a polychotomous variable, like Bae et al. (2003). In the theory of multinomial logistic regression models, if P_i is the probability of an event category i out of m possible categories, a multinomial distribution can be defined by

$$P_i = P(Y_t = i|x_j) = \frac{G(\beta_i' x_j)}{1 + \sum_{j=1}^{m-1} G(\beta_j' x_j)}, \tag{1}$$

where x is the vector of covariates and β_i the vector of coefficients associated with the covariates. The function $G(\beta_i' x)$ many times takes the form of an exponential function $exp(\beta_i' x)$, in which case Equation 1 represents a multinomial logistic (or multinomial logit) model. Such models are estimated using maximum likelihood, with the log-likelihood function for a sample of n observations given by

$$logL = \sum_{i=1}^{n} \sum_{j=1}^{m} I_{ij} log P_{ij}, \tag{2}$$

where I_{ij} is a binary variable that equals one if the ith observation falls in the jth category, and zero otherwise. Goodness-of-fit in that kind of models is measured using the $pseudo - R^2$ approach of McFadden (1974) where the unrestricted (full model) likelihood, L_Ω, and restricted (constants only) likelihood, L_ω, functions are compared:

$$pseudoR^2 = 1 - [logL_\omega / logL_\Omega]. \tag{3}$$

To capture the range of possible outcomes, and yet have a concrete model, we have a total of six categories: 0, 1, 2, 3, 4, and 5 coexceedances. For a model that has only constants, m-1, or five parameters, need to be estimated. But for every covariate added to the model, such as the daily exchange rate change, five additional parameters need to be estimated, one for each outcome. The top and the bottom coexceedances are estimated separately for extreme returns. Finally, we compute the probability of a coexceedance of a specific level, P_i, by evaluating the covariates at their unconditional values,

$$P^*_{ij} = \frac{exp(\beta'_i x^*_j)}{1+\sum_{j=1}^{m-1} exp(\beta'_j x^*_j)}, \tag{4}$$

where x_j* is the unconditional mean value of x_j.
The coefficients that are given by a multinomial logistic regression compare the probability of a given outcome with the base outcome (in our case the outcome 0 is the base outcome - i.e. the outcome where no country has an extreme return). As mentioned in Greene (2003), the coefficients of such a model are not easy to interpret. This is why in these models it is necessary to differentiate 1 in order to obtain the partial effects of the covariates on the probabilities

$$\delta_{ij} = \frac{\delta P_{ij}}{\delta \beta_i} = P_{ij}[x_j - \sum_{k=0}^{J} P_{ik}\beta_k] = P_{ij}[\beta_j - \bar{\beta}] \tag{5}$$

In multinomial logistic regressions the coefficients correspond to probabilities. Thus, these partial effects give us the marginal change in probability for a unit change in the independent covariate. In such models we are interested in seeing whether these marginal effects are statistically significant or not. These marginal effects may even have different signs than the corresponding coefficients, since the derivative $\frac{\delta P_{ij}}{\beta_{ik}}$ can have a different sign than the coefficient β_{jk}. [9] In our case, we have a variable Y_t that counts the number of coexceedances and takes the value i when extreme returns (top or bottom) occur for the same day in i stock market indices on day t. This variable is calculated separately for the Euro-core, the Euro-periphery, and the Non-euro groups. Then, in the multinomial logistic regression equation 4 P_i is equal to $P(Y_t = i|x_t)$ where $Y_t = 0, 1, 2, ...k$ is the coexceedance variable that is created for the Non-euro, and for each of the countries groups we defined (Euro-periphery vs. Euro-core etc.). So, we have k=5 for all three countries groups, where x_t is a vector of explanatory variables (covariates), on day t. In equation 4, the argument of the exponential part (representing the logistic function), is a function of the covariates (x_t) and the coefficients(the betas). This function is a linear expression of the arguments. Let's call it $g_i(t)$. We will use this function (which will take different forms) to study both the "within" and "across" groups contagion effects.

3.3. *Contagion within groups*

In this section, we study the three country groups to determine whether there exist contagion effects within them. Each of these groups has it's own set of covariates. In line with Bae et al. (2003) and Gropp et al. (2009), as independent variables incorporated in $g_i(t)$ we have the intercept, the conditional volatility of the group index at time t (h_t), the exchange rate change (per US dollars) in the group (e_t), the average short term (ST) interest rate level in the group (i_t) as a proxy for the

[9] To elaborate a little further on why it is crucial that marginal effects are calculated for such models, it is known that the coefficients of a multinomial logistic are obtained from comparing the probability of a given outcome with the base outcome. In our case, the outcome is 0, in other words, no coexceedances in the group. Thus, the estimated coefficient for covariate x_{13} for outcome 3, which is β_{13} and is the coefficient for the 1st covariate, calculated for the 3rd outcome, measures the probability of having an outcome equal to 3 (3 coexceedances in the group), instead of an outcome 0 (no coexceedances in the group), for a unit change in the covariate x_{13}. But in reality, there is also the possibility of having the outcome 2 instead of 0 for a unit change in covariate x_{13}. This is exactly why we need the marginal effects, to calculate the probabilities associated with a unitary covariate change in adjacent categories, and not taking as an alternative only the base outcome (0 in our study). This happens because the coefficients of a multinomial logistic regression model exhibit what is known as the "log odds ratio" property:

$$\ln \frac{P_{ij}}{P_{i0}} = \beta'_i x_j \tag{6}$$

interbank short term liquidity risk[10], and the average long term (LT) spread change (b_t) vis-*à*-vis Germany as a proxy for the sovereign risk[11].

We include exchange rate changes following Bae et al. (2003) who find that when currencies fall on average (which means that e_{it} rises) extreme returns are more common. Thus, the logistic regression $G(\beta_i' x) = exp(g_i(x_t))$ of equation 1 has the following form for $g_i(x_t)$:

$$g_i(x_t) = b_{0i} + b_{1i}h_{it} + b_{2i}e_{it} + b_{3i}i_{it} + b_{4i}b_{it} \tag{7}$$

where i=0, 1, 2, 3, 4, 5 for each country group, the coexceedance count for the group. Equation 7 represents the inter-group contagion formula for the three groups examined. For each group we calculate the equally weighted average group values, on a daily basis, of the conditional volatility (h_{it}), the exchange rate change (e_{it}), the short term (ST) interest rates levels (i_{it}), and the Long Term (LT) spread change vis-*à*-vis Germany (b_{it}).

We estimate these models for each group, and for three time periods:

- Entire Period (01/01/2004 - 13/03/2013)
- Pre-crisis (01/01/2004 - 26/02/2007)
- Euro-crisis (08/12/2009 - 13/03/2013)

It is worth noting that the coexceedances are calculated separately for each of the three periods. In other words, in each of the three periods the bottom and top extreme values correspond to the respective 10% and 90% threshold points of each period. Since we are mostly interested in the European countries, we focus on the Euro-crisis period and compare it to the Pre-crisis period. We first present the detailed findings for the Euro-periphery group for the entire period in Table 3.

Insert Table 3 here

The probability that no Euro-periphery country has a bottom-tail return is equal to 77.49%. This is calculated as the fraction of the number of 0 coexceedances divided by the total days $\frac{1859}{2399} = 0.774$. The coefficient β_{01} corresponds to the event Y=1, in other words the event where only one Euro-periphery country has an extreme return (an exceedance) on that day, and the probability of this event is calculated as $P_1 = \frac{\exp(\beta_{01})}{1+\sum_{i=0}^{k}\exp(\beta_{0k})}$. This probability is found to be equal to 10.3%. If currencies in the group fall on average (in which case e_{it} rises), the probability of extreme returns increases, since the signs of the exchange rate marginal effects are positive, and statistically significant at the 5% level for the first exceedance, and at the 1% level for the two, three, four and five bottom coexceedances. In their study Bae et al. (2003) measured returns in dollars and the fact that they came up with very similar results made them wonder whether the stock return contagion they measured was actually foreign exchange contagion. Thus, they also estimated their models in local currencies, but the results were similar to the dollar returns. But we estimate our models in local currencies from the beginning, so we do not face such an issue. The results for ST interest rates are mixed since two marginal effects are statistically significant, for the outcomes of one and five bottom coexceedances in the group,

[10] The interest rate series are the short term interest rates available in Datastream (3-month Interbank interest rates).

[11] Spreads are calculated as the difference between the yield of the 10 year government bond of country i's debt and the yield of the 10 year German government bond

but with contradictory signs. Regarding the LT spread change in the group vis-$\grave{a}$-vis Germany, we find negative and statistically significant marginal effects. For all coexceedances outcomes except for the second, the marginal effects are significant at the 1% level. The positive sign of the coefficient indicates that a 1% increase in the average Euro-periphery LT spread increases the probability of extreme bottom returns in the group. A change of 1% in the average LT spread of the Euro-periphery group, increases the probability of one bottom extreme return by 0.197%. To simplify the presentation in Table 4 we show a summary for the within groups results, for the entire period, the Pre-crisis and the Euro-crisis separately. The number of "+" (or "-") indicate the number of statistically significant (in the 1% or 5% levels) and positive (or negative) marginal effects.

Insert Table 4 here

Pre-crisis the effect of covariates on the probability of extreme returns is rather weak, while the role of covariates increases significantly in the crisis period in most of the cases. The results are even stronger when we take the extreme returns over the entire period which notably includes the US-crisis as well. The effect of volatility is somehow smaller in the Euro-crisis period compared to the Pre-crisis period, for the bottom tail. Exchange rates are not significant for the bottom coexceedances of the Non-euro group in the Pre-crisis period, with zero significant marginal effects, while they became significant in four of the five cases in the Euro-crisis period. Exchange rates have a positive coefficient for the bottom tail, and a negative coefficient for the top tail. This means that higher exchange rates (i.e. weaker currencies) lead to a higher probability of extreme bottom returns and a lower probability of extreme top returns. Average ST interest rates are not significant in most of the cases, while average LT spread changes become more significant in the Euro-crisis period, especially for the Euro-core and the Euro-periphery group as far as the bottom tails are concerned, and the Non-euro and Euro-core groups for the top tail returns. For the bottom tail, higher average group spreads (i.e. higher average group sovereign risk) lead to higher probabilities of extreme group bottom returns, while they decrease the probability of extreme top returns.

In summary, the findings so far indicate a much tighter co-movement of stock market returns within the three European country groups and a much tighter relationship of the fundamental factors (covariates) affecting the extreme stock market movements within each group during the Euro-crisis period compared to the Pre-crisis period.

3.4. Contagion across groups

Next we test for across groups effects. This deals with the question of whether the number of coexceedances in one group (e.g. the Euro-periphery group) can help predict the number of coexceedances in other groups (the Euro-core and the Non-euro groups). According to Bae et al. (2003), if a fraction of coexceedances in one group is unexplained by its own covariates, but can be explained by coexceedances in another area, this can be interpreted as evidence of contagion across groups [12].

[12] In general, the definition of contagion is far from being simple and commonly accepted. Pericoli and Sbracia (2003) provide five of the most widely accepted definitions of financial contagion. According to one of their definitions "Contagion is a significant increase in the probability of a crisis in one country, conditional on a crisis occurring in another country.". According to another of their definitions, "Contagion occurs when cross-country comovements of asset prices cannot be explained by fundamentals". Hence, these definitions are consistent with Bae et al. (2003).

To examine this question, we reestimate the models of Table 3 for all countries groups respectively, adding a covariate related to coexceedances (Y_{jt}^*) from another group. Our primary interest is to study for contagion effects from the Euro-periphery group to the Non-euro and the Euro-core groups. In this case the equations for the across groups examination take the following shape:

$$g_i(x_t) = b_0 + b_1 h_{it} + b_2 e_{it} + b_3 i_{it} + b_4 b_{it} + b_5 Y_{jt}^* \tag{8}$$

For example, to examine if the Euro-periphery group provokes contagion effects in the Euro-core group the dependent variable is the number of coexceedances in the Euro-core group and the first three covariates of the right hand side concern the Euro-core group, while the last covariate is related to the coexceedances of the Euro-periphery group on that day. The null hypothesis of no contagion effects can be rejected in case the coefficient of Y_{jt}^* is found to be statistically significant.

Table 5 provides the summary results for the across groups effects, for the three time periods we examine (the entire period, the Pre-crisis and the Euro-crisis periods) separately.

Insert Table 5 here

The evidence supports the hypothesis that there exist important effects from the Euro-periphery to the other two groups. Bottom (or top) coexceedances in the Euro-periphery group have a significant (and positive) impact on the coexceedances of the Euro-core and Non-Euro groups in most of the cases. In other words an increase in bottom (or top) coexceedances for the Euro-periphery group increases the probabilities of bottom (or top) coexceedances in the other groups as well. The results are stronger for the entire sample period (which also includes the US-crisis) but they are not perceptively different between the Pre-crisis and the Euro-crisis periods.

The findings indicate that extreme returns in the Euro-periphery group are indeed associated with extreme returns in the other two European groups but this relationship, interestingly and perhaps contrary to popular belief, does not seem to have intensified during the Euro-crisis period. Surprisingly, both for the Non-euro and the Euro-core (bottom coexceedances), the average marginal effects (when taking into account only the significant outcomes) are lower in the Euro-crisis compared to the Pre-crisis (0.038 vs 0.060 for the Non-euro and 0.024 vs 0.041 for the Euro-core). Of course, the absolute size of the contagion effects is stronger since the same 10% and 90% cutoff values are higher during the Euro-crisis period[13].

4. Information Flow and Extreme Returns

The statistical relationships between capital market prices in the Euro-periphery and the other European country groups provide useful indications about transmission of the Euro-crisis, but by and large these are second order effects. Coexceedances of extreme returns are correlated to various covariates as we have seen above but the ultimate source of their (and the covariates' variability)

[13]The 10% and 90% quantiles refer to deeper cutoff points before crisis and in the crisis period. The average 10% Pre-crisis and Euro-crisis cutoff percentage returns have the following values: -0.957% vs -1.261% for the Non-euro, -0.741% vs -1.902 for the Euro-periphery, -0.830% vs -1.450% for the Euro-core group. In other words, there is no significant change in the contagion mechanism, but the contagion effects became much deeper as far as the percentages of the extreme returns are concerned for the different groups.

must be the investors' consensus about various information flows related to the crisis. Although very difficult to observe, measure and summarize, one can conjecture that changes in relevant information flows is the root cause for these relationships to exist. Obtaining usable indicators of information flows and relating them directly to the (extreme) market returns is the focus of the rest of the paper. This approach should provide interesting contributions to the relatively new strand in the finance literature that examines directly how information is assimilated by and affects market prices.

In this context we analyze the content of news stories in the press or in the newswires and the Web Attention about the evolving Euro-crisis. As far as the news stories are concerned, we extract and analyze from Dow Jones Factiva [14] a total of 110.800 news articles covering the time period from January 1st, 2004 till March 13th, 2013. We collect news articles from seven sources: *Dow Jones Newswires, Thomson Reuters, Financial Times, The Wall Street Journal, The New York Times, The Telegraph, The Times.* We use these seven sources because first of all they returned the greatest number of news items for our queries via Factiva and Secondly, they are undoubtedly some of the most popular news sources worldwide. Dow Jones Newswires and Thomson Reuters give news items in newswires form, capturing news in real time. The Wall Street Journal and The New York Times are the main points of reference from the United States, and The Financial Times, The Telegraph and The Times are the main European news papers for the financial markets. We include both content from the print and the online editions (where available) from all our seven sources. For each country the relevant stories are obtained by a query searching for news with the name of the country plus one of the following terms each time: crisis, debt, economy, deficit, default. For example, for Greece the news were retrieved by searching for news stories containing any of the terms:

- "greek crisis"
- "greek debt"
- "greek economy"
- "greek deficit"
- "greek default"

The same applies to all five Euro-periphery countries. The importance of these search terms in the period examined is obvious and follows closely the search terms used in Google Trends (see below): A news item that contains the term Greek crisis is certainly linkedin to the crisis in Greece. "Greek debt" is relevant since the european crisis is also a debt crisis. The search term "greek economy" is included in order to capture the stories about the nations' economies. The "greek deficit" component is included since a lot of discussion is made around the deficits of the countries and the deficit is obviously one of the main factors to assess the financial performance of a nation. Finally, the "greek default" component captures the sovereign default risk debate, since the fear of countries defaulting elevated at various time points, with long term (LT) spreads reaching consecutive historical highs. These five search terms were also found to be the most relevant key words used in Google searches. The news sources and the total number of news items appear in Table 6:

Insert Table 6 here

We see from Table 6 that Pre-crisis most news are about Italy and Spain, but in the Euro-crisis period the most news concern Greece. The most news overall are obtained from the newswires (Dow

[14] Dow Jones Factiva can be found in `http://www.dowjones.com/factiva/index.asp`

Jones and Thomson Reuters - 48382 and 28189 out of a total 110800 items). One can easily notice that the amount of news surrounding each country increased dramatically from the Pre-crisis to the Euro-crisis periods. For example, there were a total of 899 articles regarding Greece in the Pre-crisis period, and this number jumped to 66959 for the Euro-crisis period.

As a next step, using textual analysis, based on the Loughran and McDonald (2011) dictionary [15], we calculate a "pessimism factor" as in Garcia (2013): for each day, the positive media content is defined as $G_t = \sum_i \frac{g_{it}}{w_{it}}$ and calculating the percentage of positive words over the total number of words in each article. The symbol g_{it} stands for the number of positive words in article i on day t, w_{it} for the total number of words in article i on day t. We do the same for the negative words, calculating the percentage of negative words over all articles on day t, and then summing once more for the whole day, obtaining the negative media content as $B_t = \sum_i \frac{b_{it}}{w_{it}}$, with b_{it} denoting the negative words in article i of day t. Thus, we obtain the "News Pessimism factor" on day t as the difference between the negative and the positive media measures, in other words

$$\textit{News Pessimism factor: } P_t = B_t - G_t \tag{9}$$

The "News Pessimism factor" is calculated for every Euro-periphery country, for every day. Then, the Euro-periphery "News Pessimism factor" (EPNP) is calculated as the average of the Euro-periphery News Pessimism factors on every day: $EPNP_t = \frac{\sum_{i=1}^{5} P_{it}}{5}$ where P_{it} is the "News Pessimism factor" for country i of the Euro-periphery countries (i takes values 1 to 5, one for each of the Euro-periphery countries).

A second metric we use is "News Relevance factor" (R_t): This is the total number of articles written in a day regarding one of the queries we are interested in:

$$\textit{News Relevance factor: } R_t = \sum_i I_{it} \tag{10}$$

with I_{it} being an binary variable taking the value 1 if article i belongs in the set of articles corresponding to each country's query for day t, and 0 if article i does not belong in the articles result set for this country. In a similar rein we then estimate a Euro-periphery News Relevance factor.

Finally, we attempt a connection between the financial data and the investor attention as measured by the search frequency of Google Trends via the Search Volume Index (SVI) [16]. Google Trends provides weekly (and for some frequent terms daily) time series that depict how much a key term (or terms) was searched for via the Google Search Engine for a certain period of time (see Da et al. (2011)). Google is by far the most popular search engine in the world, with an 88.8% market share as of June 2013 [17]. Thus, it is safe to assume that it captures the worlwide interest of the population as measured by the searches the individuals perform worldwide. Moreover, as mentioned in Da et al. (2011), when someone searches for something on Google (be it a stock, a bond or information about the crisis), he certainly is interested in it. Thus, Google Trends provide a direct measure of attention. Especially in crisis times, one could argue that the SVI can capture the uncertainty and the interest over topics and issues that trouble the markets and the nations and attract the investors' interest worldwide. We hasten to add, however, that the information about the SVI itself is not publicly

[15] The dictionary can be found at `http://www3.nd.edu/~mcdonald/Word_Lists.html`

[16] Google Trends can be found at: `http://www.google.com/trends/explore#cmpt=q`

[17] Source: `http://www.karmasnack.com/about/search-engine-market-share/`

available in real time. Investors can know about it only with a time delay. The Google Trends SVI is:

$$\textit{Web Search Volume Index: } SVI_{it} = k, k = 0, ..., 100 \tag{11}$$

SVI_{it} is a scaled time series taking a discrete value (0 to 100) for time t (0 meaning the query was not searched at all on time t, and a 100 when it was most searched for in the given time frame), based on the number of searches made via the Google Search Engine for a specific query and time period. Depending on the popularity of the query, Google Trends provides a time series of monthly (least searched), weekly, or daily (most searched) frequency. If a query is not searched enough for Google's threshold, no results are returned for this period [18]. Since we are mainly interested in the Euro-crisis period, and especially the Euro-periphery countries, we decided to proceed with the same sets of key search terms that we used in our News Flow analysis before, each one corresponding to a country. [19] [20]

The graphical illustrations of the Information Flow factors during the Euro-crisis period along with the bottom coexceedances for the Euro-periphery group appear in the following Figures:

Insert Figure 1 here

It is visually evident that there exists a correlation between the bottom coexceedances of the Europerphery group and the three Information Flow factors during the Euro-crisis. "Spikes" of bottom coexceedances (or "bottom coexceedances clustering") notably in the periods April-June 2010 and May-December 2011 seem to be related to the evolution of the Information Flow factors.

The summary statistics for the Information Flows for the Pre-crisis and the Euro-crisis subperiods appear in Table 7:

Insert Table 7 here

We notice that the Information Flow factors significantly increase from the Pre-crisis to the Euro-crisis period. More specifically, the mean value of the News Pessimism factor increased from 0.127% Pre-crisis to the value of 0.394% in the Euro-crisis period. As far as the News Relevance factor is concerned, it increased from an average of 1.285 Euro-periphery articles per day during the Pre-crisis period to an average of 22.046 articles per day during the Euro-crisis period. Regarding the

[18] The maximum time period for which daily data can be obtained by Google Trends is three months for each query. And every Google Trends time series returned is scaled with the maximum value for the specified timeperiod. For this reason, we scaled all three months time intervals for each country with a common scaling factor which was the day with the most searches in the entire time period, thus obtaining a homogenized scaling for the entire period.

[19] For each query, for example "Greek crisis", Google Trends provides 5 related search terms and their popularity. The choice of the 5 terms used in Google Trends and the News Flow was partially influenced by this popularity. Moreover the syntax of the Google queries was modified slightly also in function of their popularity. For example we used "Greece crisis" instead of "Greek crisis", because "Greece crisis" was much more searched for. Of course the two queries are referring to the same entities and thus it is safe to claim the two queries are equivalent.

[20] No daily data were available for these search queries for the Pre-crisis period, thus the SVI analysis is done only for the Euro-crisis period.

correlations, there exist significant changes when comparing the two subperiods. The News Pessimism factor-News Relevance factor correlation increased from 0.695 in the Pre-crisis to 0.939 in the Euro-crisis period. The correlations between the News Pessimism, the News Relevance factors and the group stock markets Pre-crisis were small but positive, but during the Euro-crisis period they become negative and much larger in magnitude. In other words, during the Euro-crisis a higher Euro-periphery News Pessimism and News Relevance is associated with lower stock returns for all three country groups (because of the negative sign). The correlations between the News Pessimism and the stock returns is higher (in absolute values) than the correlations with the News Relevance, which indicates that the News Pessimism provides a better fit for stock returns. The SVI is also negatively correlated with all three group stock indexes during the Euro-crisis period.

In this context, the model we previously employed is altered so that the dependent variable is the number of bottom (or top) coexceedances for one of the three country groups while the independent variable is the estimated News Relevance (or News Pessimism or Web SVI) factors for the Euro-periphery group. Notice that in this model we do not employ the other financial covariates as in Equation 8, because, as explained above, a priori they are all affected simultaneously by the same information flow which in our case is proxied by these factors [21]. The results for the three periods and for the three information flow metrics appear in Table 8.

Insert Table 8 here

The main findings of Table 8 are that Pre-crisis the effect of the periphery News Pessimism factor was virtually inexistent, but entering the Euro-crisis period, fourteen out of fifteen marginal effects affecting the probabilities of bottom coexceedances for our groups become significant. Thus, the media content becomes significant in the crisis period. We need to notice that the marginal effects are positive, which means that a higher News Pessimism in the media is associated with higher probabilities of extreme bottom returns in the groups' stock markets. Thus, the Euro-periphery News Pessimism factor does not only affect the probabilities of extreme bottom returns for the periphery group, but also for the Non-euro and the Euro-core groups. In other words, bad news about the periphery affect the probabilities of extreme bottom returns for the other two groups as well. The effect on the probabilities of top returns is insignificant for most of the cases. Only in one out of five outcomes of top coexceedances for the core group (for the entire period) and in three outcomes of the periphery group do we find significant marginal effects (for the entire period).

As far as the News Relevance factor is concerned (which measures the volume of relevant news for each group), the picture is more or less the same as with the News Pessimism factor, albeit the results are weaker. This can be attributed to the reasonable phenomenon of having more noise in a variable which measures the aggregate count of news, versus the more sophisticated News Pessimism variable which captures (via textual analysis) the actual pessimistic tone in the media content itself. This comparison favors the use of more sophisticated text analysis techniques such the one we use, which is based on the works of Loughran and McDonald (2011) and Garcia (2013). Once more, the marginal effects are positive, in other words higher values of the News Relevance factor (i.e. more news items related to our queries) are associated with higher probabilities of extreme returns.

[21] As a robustness check, we estimated the fundamental variables models along with the Information Flows variables. The results are weaker but we still get a number of significant coefficients for the Information Flows variables.

Finally, the Web Search Volume Index factor (based on the volume of Google Searches, for which daily data exist only for the Euro-crisis period), exhibits significant and positive marginal effects for twelve out of fifteen bottom coexceedances. In other words, more Web Attention for the Euro-periphery during the Euro-crisis is associated with higher probabilities of extreme returns for all three groups we studied. These results bring some useful implications for investors and policymakers: The effect of news, the tone of news and the Web Attention are closely related with the probabilities of extreme returns (especially in times of crisis). We find that this effect is not only contained within the borders of the specific group that the news or the Web Attention metrics capture, but it also spreads out across groups. For all three groups the probabilities of extreme bottom returns is affected in a positive way (i.e. more pessimistic news about the periphery are associated with higher probabilities of extreme returns for the other groups). Based on these results, one can argue in favor of "contagion" or "transmission" or "propagation" of News Pessimism and of Web Attention within and across groups during crisis times. Thus it might be useful for investors and policymakers to be aware of these dynamics and effects when making investment or policy decisions.

Until now we have established a contemporaneous association between Information Flows and the probabilities of extreme bottom returns during the Euro-crisis. But a question that naturally follows is how are these Information Flows associated among themselves and with stock market returns? Does investors' attention (Web Attention) precede extreme movements in the markets or are investors more interested after extreme market events? Moreover, does News Relevance and News Pessimism follow or precede the movements in the markets? And, lastly, does News Relevance and News Pessimism follow or precede the Web Attention factor? To study these questions, we perform a series of pairwise Granger causality tests (Granger (1969)) for each of the Euro-periphery Information Flows factors along with the average group stock market returns and separately with the bottom coexceedances. We calculate a series of Vector Autoregression (VAR) models and keep the optimal lag numbers by the Akaike information criterion (AIC, Akaike (1974)). A p-th order vector autoregression, or VAR(p), with exogenous variables x model can be written as

$$y_t = v + A_1 y_{t-1} + \ldots + A_p y_{t-p} + B_0 x_t + B_1 x_{t-1} + \ldots + B_s x_{t-s} + u_t \tag{12}$$

where y_t is a vector of K variables, each modeled as function of p lags of those variables and possibly a set of exogenous variables x_t. It is assumed that $E(u_t) = 0, E(u_t u_s') = 0$, $\forall$ t $\neq$ s.

For two stationary (i.e. without unit roots) variables X and Y, the Granger causality test is defined on the following equation:

$$Y_t = \alpha + \delta t + \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} + \beta_1 X_{t-1} + \ldots + \beta_q X_{t-q} + \epsilon_t \tag{13}$$

In such a setting, X Granger causes Y if *any* of $\beta's$ are found to be statistically significant. All of our time series were found to be stationary using the Augmented Dickey-Fuller (Dickey and Fuller (1979)) and the Phillips-Perron (Phillips and Perron (1988)) unit-root tests. The Granger-causality results appear in Table 9:

Insert Table 9 here

During the Pre-crisis period the Euro-periphery News Pessimism factor is found to be causing movements for the Euro-periphery News Relevance, the stock markets and the Bottom Coexceedances of the three groups. On the other hand at the Euro-crisis period the causality becomes two-sided, in other words the News Pessimism causes movements in the stock markets and the Bottom Coexceedances, but also the markets seem to be causing News Pessimism in the media. The causality relationships in the Euro-crisis period are much stronger than for the Pre-crisis period, and once more we see that the News Pessimism factor gives more significant results than the News Relevance factor. The Web Search Volume Index (SVI) is found to be causing movements in the stock markets, without the opposite causality effect being present (with the exception of the Non-Euro Bottom Coexceedances). Thus, the Web Attention seems to be preceding the subsequent movements in the stock markets. The causality relationship between the News Pessimism and the SVI is two sided, i.e. it is not possible to claim whether the Web Attention precedes or lags the News Pessimism in the news articles.

5. Robustness and alternative specifications

To verify the robustness of our results, as a first robustness check, instead of 10% and 90% extreme returns cutoffs, we used the 5% and 95% percentages. The results are robust in this change. Furthermore as a second robustness check, instead of the raw returns, we calculated extreme returns on the standardized residuals of a GARCH(1,1) model, accounting for the time-varying volatility effects, since in periods of high volatility, extreme returns are more probable. In order to calculate the volatility, we move in line with Christiansen and Ranaldo (2009), estimating a AR(1)-GARCH(1,1) model for each group's average returns:

$$Ret_t^{group} = c_0 + c_1 Ret_{t-1}^{group} + \epsilon_t \tag{14}$$

where ϵ_t $N(0, \sigma_t^2)$ and the variance follows a GARCH(1,1) process:

$$\sigma_t^2 = c_2 + c_3 \sigma_{t-1}^2 + c_4 \epsilon_{t-1}^2 \tag{15}$$

The volatilities are then obtained as the estimated $\hat{\sigma_t}$ from the AR(1)-GARCH(1,1) model. We notice that for the coexceedances filtered by a GARCH, the effect of volatility is not significant (for the raw returns coexceedances all volatility coefficients were found to be positive and statistically significant - in other words an increase in volatility increases the probability of extreme bottom returns). As far as the Google Trends SVI robustness checks are concerned, apart from the mean SVI, we also calculated the scaled SVI for the Euro-periphery using the most searched query as a common scaling factor for all country SVI time series. The results were found to be robust.

6. Conclusion

We examine the existence of stock market contagion effects among three groups of countries: the Euro-periphery countries (Portugal, Ireland, Italy, Greece, Spain), the Euro-core countries (Germany, France, the Netherlands, Finland, Belgium), and the major European Union -but not euro- countries (Sweden, UK, Poland, Czech Republic, Denmark), using daily stock market data from January 2004 till March 2013. Our analysis is split in two parts: the first part concerns extreme stock index returns, controlling for various fundamentals derived from financial market data (volatility, exchange rate change, short interest rates, long term spread change). We find contagion effects within the Euro-periphery group and from the Euro-periphery group to the Non-euro and the Euro-core groups for both the Pre-crisis and the Euro-crisis periods. The financial contagion transmission mechanism does not appear to have changed as a result of the crisis, although the magnitude of extreme returns have increased in the crisis period.

The second part is related to the relationship between extreme stock returns and three "Information flow metrics" estimated for the Euro-periphery group which as commonly accepted is at the origin of the Euro-crisis; the News Pessimism factor, the News Relevance factor and the Web Search Volume Index. We report evidence that the Euro-periphery information flow metrics significantly affect the probabilities of extreme returns in the three groups. The effect in the vast majority of cases is asymmetric, with the Euro-periphery SVI and News Pessimism factor affecting the probabilities of extreme bottom returns for the three groups. The effect is positive. In other words a higher SVI (i.e. more Web Attention in times of crisis) and a higher News Pessimism factor (i.e. more bad news) are associated with more extreme bottom returns. The implications of the overall findings are quite significant for investors who may want to diversify their portfolios and should be aware of the stock indices movement dynamics and of how extreme shocks propagate from one group of countries to the others, affecting their portfolios' overall risk. Furthermore, it would be useful for policy makers to be aware of these stock market dynamics, in order to assess policy decision making in times of extreme shocks (such as crisis times). Due to the high complexity of financial markets and the extremely high level of available information from the press and the web, agents can incorporate information extracted from textual analysis of news items and trends on the web that may be associated with the market movements. Last but not least, Granger-causality tests reveal a two-sided causality effect between the News Pessimism and the stock market extreme returns, while the Web Search Volume Index (SVI) is found to Granger-cause movements in the stock markets and Bottom Coexceedances in the three country groups.

Future research could move in the direction of higher frequency (intraday) financial markets dynamics. In addition, more sophisticated data mining and textual analysis techniques can be employed to further improve the quality of the Information Flow metrics. At the moment we have a contemporaneous model. This is designed for investors who want to assess within the trading day the probabilities of extreme returns given the evolution of the other asset classes and the Information Flow factors. An issue of interest would be to study in more detail the lead-lag effects affecting the probabilities of extreme returns with the view of possibly developing profitable trading rules.

References

Afonso, A., Furceri, D., Gomes, P., 2012. Sovereign credit ratings and financial markets linkages: application to european data. Journal of International Money and Finance 31, 606–638.

Ahlgren, N., Antell, J., 2010. Stock market linkages and financial contagion: A cobreaking analysis. The Quarterly Review of Economics and Finance 50, 157–166.

Akaike, H., 1974. A new look at the statistical model identification. Automatic Control, IEEE Transactions on 19, 716–723.

Allen, F., Gale, D., 2000. Financial contagion. Journal of Political Economy 108, 1–33.

Arezki, R., Candelon, B., Sy, A., 2011. Sovereign rating news and financial markets spillovers: Evidence from the european debt crisis. CESifo Working Paper Series 3411, CESifo Group Munich.

Aït-Sahalia, Y., Cacho-Diaz, J., Laeven, R. J., 2010. Modeling financial contagion using mutually exciting jump processes. Working Paper 15850, National Bureau of Economic Research.

Bae, Karolyi, Stulz, 2003. A new approach to measuring financial contagion. Review of Financial Studies 16, 717–763.

Bekaert, G., Harvey, C. R., Ng, A., 2005. Market integration and contagion. The Journal of Business 78, 39–70.

Bhanot, K., Burns, N., Hunter, D., Williams, M., 2012. Was there contagion in eurozone sovereign bond markets during the greek debt crisis? Working Paper WP 006FIN-73-2012, The University of Texas at San Antonio, College of Business.

Boyson, N. M., Stahel, C. W., Stulz, R. M., 2010. Hedge fund contagion and liquidity shocks. The Journal of Finance 65, 1789–1816.

Caporin, M., Pelizzon, L., Ravazzolo, F., Rigobon, R., 2013. Measuring sovereign contagion in europe. Tech. rep., National Bureau of Economic Research.

Christiansen, C., Ranaldo, A., 2009. Extreme coexceedances in new eu member states' stock markets. Journal of banking & finance 33, 1048–1057.

Corsetti, G., Pericoli, M., Sbracia, M., 2005. 'some contagion, some interdependence': More pitfalls in tests of financial contagion. Journal of International Money and Finance 24, 1177 – 1199.

Da, Z., Engelberg, J., Gao, P., 2010. The sum of all fears: investor sentiment and asset prices. SSRN eLibrary .

Da, Z., Engelberg, J., Gao, P., 2011. In search of attention. The Journal of Finance 66, 1461–1499.

Dickey, D. A., Fuller, W. A., 1979. Distribution of the estimators for autoregressive time series with a unit root. Journal of the American statistical association 74, 427–431.

Dungey, M., Fry, R., González-Hermosillo, B., Martin, V. L., 2005. Empirical modelling of contagion: a review of methodologies. Quantitative Finance 5, 9–24.

Dzielinski, M., 2012. Measuring economic uncertainty and its impact on the stock market. Finance Research Letters 9, 167–175.

Forbes, K. J., Rigobon, R., 2001. No contagion, only interdependence: Measuring stock market comovements. Journal of Finance 57, 2223–2261.

Garcia, D., 2013. Sentiment during recessions. The Journal of Finance 68, 1267–1300.

Grammatikos, T., Vermeulen, R., 2012. Transmission of the financial and sovereign debt crises to the emu: Stock prices, cds spreads and exchange rates. Journal of International Money and Finance 31, 517–533.

Granger, C. W., 1969. Investigating causal relations by econometric models and cross-spectral methods. Econometrica: Journal of the Econometric Society pp. 424–438.

Greene, W. H., 2003. Econometric analysis. Pearson Education India.

Gropp, R., Duca, M. L., Vesala, J., 2009. Cross-border bank contagion in europe. International Journal of Central Banking 5, 97–139.

Jiang, G. J., Konstantinidi, E., Skiadopoulos, G., 2012. Volatility spillovers and the effect of news announcements. Journal of Banking and Finance 3, 2260 – 2273.

Kaminsky, G. L., Reinhart, C., Vegh, C. A., 2003. The unholy trinity of financial contagion. Working Paper 10061, National Bureau of Economic Research.

Loughran, T., McDonald, B., 2011. When is a liability not a liability? textual analysis, dictionaries, and 10-ks. The Journal of Finance 66, 35–65.

Lucey, B., Sevic, A., 2010. Investigating the determinants of banking coexceedances in europe in the summer of 2008. Journal of International Financial Markets, Institutions and Money 20, 275–283.

Markwat, T., Kole, E., van Dijk, D., 2009. Contagion as a domino effect in global stock markets. Journal of Banking & Finance 33, 1996 – 2012, financial Globalisation, Risk Analysis and Risk Management.

McFadden, D., 1974. The measurement of urban travel demand. Journal of public economics 3, 303–328.

Metiu, N., 2012. Sovereign risk contagion in the eurozone. Economics Letters 117, 35–38.

Mink, M., De Haan, J., 2012. Contagion during the greek sovereign debt crisis. Journal of International Money and Finance .

Missio, S., Watzka, S., 2011. Financial contagion and the european debt crisis. CESifo Working Paper Series 3554, CESifo Group Munich.

Pericoli, M., Sbracia, M., 2003. A primer on financial contagion. Journal of Economic Surveys 17, 571–608.

Phillips, P. C., Perron, P., 1988. Testing for a unit root in time series regression. Biometrika 75, 335–346.

Rigobon, R., 2002. Contagion: How to Measure It?, University of Chicago Press, pp. 269–334.

Tetlock, P. C., 2007. Giving content to investor sentiment: The role of media in the stock market. The Journal of Finance 62, 1139–1168.

Thomadakis, A., 2012. Measuring financial contagion with extreme coexceedances. Tech. rep., School of Economics, University of Surrey.

Varian, H. R., Choi, H., 2009. Predicting the present with google trends. Google Research Blog 2.

Vlastakis, N., Markellos, R. N., 2012. Information demand and stock market volatility. Journal of Banking & Finance 36, 1808–1821.

Table 1: Descriptive Statistics and Correlations for Stock Indices

Panel A: Descriptive Statistics

	Pre-crisis 1/1/2004 - 26/2/2007			US-crisis 27/2/2007 - 7/12/2009			Euro-crisis 8/12/2009 - 13/3/2013		
	Non-euro	Euro-periphery	Euro-core	Non-euro	Euro-periphery	Euro-core	Non-euro	Euro-periphery	Euro-core
Mean (%)	0.101	0.089	0.081	−0.024	−0.074	−0.049	0.030	−0.017	0.029
Median (%)	0.113	0.104	0.115	0.012	0.012	0.017	0.034	0.017	0.047
Std. Dev. (%)	0.949	0.740	0.797	1.876	1.791	1.779	1.141	1.602	1.262
Minimum (%)	−6.020	−6.230	−9.926	−14.185	−13.341	−9.199	−6.805	−7.132	−7.787
Maximum (%)	8.505	4.816	6.203	15.236	10.509	16.062	6.406	12.474	8.240

Panel B: Correlations

	Pre-crisis 1/1/2004 - 26/2/2007			US-crisis 27/2/2007 - 7/12/2009			Euro-crisis 8/12/2009 - 13/3/2013		
	Non-euro	Euro-periphery	Euro-core	Non-euro	Euro-periphery	Euro-core	Non-euro	Euro-periphery	Euro-core
Non-euro	1.000			1.000			1.000		
Euro-periphery	0.429	1.000		0.677	1.000		0.520	1.000	
Euro-core	0.522	0.579	1.000	0.727	0.729	1.000	0.734	0.646	1.000

Note: European countries are split in three groups: the Euro-periphery countries (Portugal, Ireland, Italy, Greece, Spain), the Euro-core countries (Germany, France, Finland, the Netherlands, Belgium) and the European Union -non Euro- countries (Poland, Czech Republic, Sweden, UK, Denmark). For each of the groups the mean return and the standard deviations are calculated on the equally weighted mean portfolio of the stock market daily returns for each group.

Table 2: Summary statistics of bottom and top coexceedances of log returns for daily country groups stock indices, January 1st 2004 to March 13th 2013.

	Mean return (%) when $i = 5$	Number of bottom coexceedances						Number of top coexceedances						Mean return (%) when $i = 5$
		5	4	3	2	1	0	0	1	2	3	4	5	
Non-euro														
POL	−3.446	55	41	38	49	57	1847	1783	82	47	40	43	28	3.653
SWE	−3.727	55	45	54	54	32	1847	1783	42	56	61	53	28	4.025
CZE	−3.828	55	27	28	46	84	1847	1783	95	57	31	29	28	4.022
UK	−3.241	55	52	60	43	30	1847	1783	37	56	62	57	28	3.572
DEN	−3.370	55	47	48	42	48	1847	1783	61	42	55	54	28	3.382
Subtotal		55	53	76	117	251	1847	1783	317	129	83	59	28	
Euro-periphery														
POR	−3.253	54	66	44	39	37	1859	1817	61	45	34	60	40	3.008
IRE	−3.944	54	54	26	50	56	1859	1817	69	47	31	53	40	3.522
ITA	−3.636	54	69	45	52	20	1859	1817	20	62	56	62	40	3.678
GRE	−4.160	54	35	22	23	106	1859	1817	108	32	20	40	40	4.200
SPA	−3.503	54	64	49	44	29	1859	1817	29	56	54	61	40	3.652
Subtotal		54	72	62	104	248	1859	1817	287	121	65	69	40	
Euro-core														
GER	−2.855	109	44	32	19	36	1970	1938	46	25	31	47	91	2.530
FRA	−3.137	109	56	46	19	10	1970	1938	13	35	41	60	91	2.842
NL	−3.169	109	51	37	22	21	1970	1938	20	27	50	52	91	2.782
FIN	−3.303	109	38	24	19	50	1970	1938	48	26	26	49	91	3.240
BEL	−2.809	109	35	32	27	37	1970	1938	51	29	29	40	91	2.534
Subtotal		109	56	57	53	154	1970	1938	178	71	59	62	91	

Note: Exceedances for daily stock index top (bottom) log returns are the ones belonging to the highest (lowest) 10% of all daily returns. The coexceedances are defined as the joint occurrence of exceedances (extreme returns - bottom or top) across different country indexes on the same day. For example, out of a total sample of 2399 trading days, there are 104 occurrences of bottom-tail coexceedances for the Euro-periphery countries with 2 countries only, and in 23 of those days Greece is the one of the two countries with bottom-tail coexceedances.

Table 3: Within the Euro-periphery group bottom and top coexceedances of log returns for the entire period.

	(1)	(2)	(3)	(4)	(5)
	Margin / SE	Margin / SE	Margin / SE	Margin / SE	Margin / SE
Constant	-0.224***	-0.145***	-0.082***	-0.071***	-0.059***
	(0.015)	(0.011)	(0.010)	(0.010)	(0.009)
Volatility	0.085***	0.042***	0.018***	0.021***	0.014***
	(0.011)	(0.006)	(0.004)	(0.004)	(0.003)
Exchange Rate Change	0.024**	0.027***	0.015***	0.018***	0.010***
	(0.010)	(0.006)	(0.004)	(0.003)	(0.002)
ST Interest Rate	-0.012***	0.001	-0.000	-0.002	0.003***
	(0.004)	(0.002)	(0.002)	(0.001)	(0.001)
LT Spread Change	0.197***	0.017	0.113***	0.087***	0.065***
	(0.065)	(0.037)	(0.022)	(0.018)	(0.014)
Observations	2399	2399	2399	2399	2399
Baseline predicted probability	0.103	0.043	0.026	0.030	0.023
R^2	0.118				

Note: Columns (1) to (5) correspond to bottom coexceedances outcomes 1 to 5. In other words, column (1) presents the marginal effects in the case of one bottom coexceedance for the Euro-periphery group, and columns (2),(3),(4),(5) correspond to two, three, four and five bottom coexceedances for this group. The value of 0.197 for the Euro-periphery LT spread changes means that an increase of 1 percent in the average Euro-periphery long term spread (vis-$\grave{a}$-vis Germany) increases the probability of extreme bottom stock returns in the group by 0.197%, while the value of 0.024 for the average exchange rate change means that a one percent increase in the average Euro-periphery exchange rate increases the probability of one bottom Euro-periphery exceedance by 0.024%.
(***) : significance at 1% level
(**) : significance at 5% level
(*) : significance at 10% level

Table 4: Within groups Summary Results for bottom and top coexceedances.

	Bottom tail			Top tail		
Panel A: Entire Period						
	Non-euro	Euro-core	Euro-periphery	Non-euro	Euro-core	Euro-periphery
Volatility	+ + + + +	+ + + + +	+ + + + +	+ + ++	+ + + + +	+ + + + +
Exchange Rate Change	+ + + + +	+ + ++	+ + + + +	− − − − −	− − −−	− − − − −
ST Interest Rate	++	+			−	− − −−
LT Spread Change	+ + + + +	+ + ++	+ + ++	− − − − −	− − −	− − − − −
$Pseudo - R^2$	0.120	0.093	0.118	0.111	0.091	0.096
Panel B: Pre-crisis Period						
	Non-euro	Euro-core	Euro-periphery	Non-euro	Euro-core	Euro-periphery
Volatility	+ + +	++	+ + +	+ + +	++	+
Exchange Rate Change		−	−		+	+
ST Interest Rate	−−			−		+
LT Spread Change	+		+	−		
$Pseudo - R^2$	0.040	0.036	0.039	0.044	0.038	0.016
Panel C: Euro-crisis Period						
	Non-euro	Euro-core	Euro-periphery	Non-euro	Euro-core	Euro-periphery
Volatility	+	++	++	+ + ++	+ + ++	−
Exchange Rate Change	+ + ++	+ + +	+ + ++	− − −−	−−	− − −−
ST Interest Rate	+					
LT Spread Change	+	+ + ++	+ + +	− − −−	− − −−	−
$Pseudo - R^2$	0.194	0.181	0.176	0.172	0.167	0.119

Note: The number of "+" (or "-") indicate the number of statistically significant (in the 1% or 5% levels) and positive (or negative) marginal effects. For example, for the bottom tail returns and the entire period sample, all five volatility marginal effects are significant and positive for the Non-euro group, indicating that an increase in volatility increases the probability of extreme bottom returns in all five bottom coexceedances outcomes. For the top tail returns, the number of statistically significant marginal effects are five for the average exchange rate change in the Non-euro group, meaning that an increase in the average group's exchange rates (i.e. weaker group currencies on average) lead to lower probabilities of extreme top returns in all five top Non-euro coexceedances outcomes.

Table 5: Across groups Summary Statistical Results for bottom and top coexceedances.

	Bottom tail		Top tail	
Panel A: Entire Period				
	Non-Euro	Euro-core	Non-Euro	Euro-core
(from Euro-periphery)				
Volatility	+ + +	++	+ + +	+ + ++
Exchange Rate Change	+ + +		− − −−	
ST Interest Rate	+ + +	+ + +	+ + ++	+
LT Spread Change			−−	
Euro-periphery Coexceedances	+ + ++	+ + + + +	+ + ++	+ + + + +
$Pseudo-R^2$	0.299	0.373	0.228	0.321
Panel B: Pre-crisis Period				
	Non-Euro	Euro-core	Non-Euro	Euro-core
(from Euro-periphery)				
Volatility			++	++
Exchange Rate Change		−		+
ST Interest Rate			−	
LT Spread Change			−	
Euro-periphery Coexceedances	+ + +	+ + ++	+ + +	+ + ++
$Pseudo-R^2$	0.223	0.281	0.159	0.245
Panel C: Euro-crisis Period				
	Non-Euro	Euro-core	Non-Euro	Euro-core
(from Euro-periphery)				
Volatility			+ + +	+ + +
Exchange Rate Change	+ + +		− − −	
ST Interest Rate		+		
LT Spread Change		+	−	−−
Euro-periphery Coexceedances	+ + +	+ + ++	+ + +	+ + ++
$Pseudo-R^2$	0.290	0.373	0.262	0.342

Note: The number of "+" (or "-") indicate the number of statistically significant (in the 1% or 5% levels) and positive (or negative) marginal effects. For example, for the bottom tail returns and the entire period sample, three out of five volatility marginal effects are significant and positive for the Non-euro group, indicating that an increase in volatility increases the probability of extreme bottom returns for the Non-euro group in three out of five bottom Non-euro coexceedances outcomes. For the top tail returns, the number of statistically significant marginal effects are four for the average exchange rate change, and have a negative sign in all four cases, meaning that an increase in the average group's exchange rates (i.e. weaker group currencies on average) lead to lower probabilities of extreme top returns in four out of five top Non-euro coexceedances outcomes.

Table 6: Number of news stories per country and source.

Panel A: Entire Period						
Dow Jones Newswires	2167	3363	4839	31800	6213	48382
Thomson Reuters	1259	2203	3811	16600	4316	28189
Financial Times	423	799	1023	4772	1478	8495
The Wall Street Journal	318	514	831	4161	1168	6992
The New York Times	52	144	215	1154	275	1840
The Telegraph	409	359	1132	5017	1116	8033
The Times	139	1534	572	2975	924	6144
Total	4852	9314	12701	68147	15786	110800
Panel B: Pre-crisis						
Source	Portugal	Ireland	Italy	Greece	Spain	Total
Dow Jones Newswires	101	205	933	157	270	1666
Thomson Reuters	243	262	814	675	915	2909
Financial Times	29	58	170	37	194	488
The Wall Street Journal	9	27	95	8	34	173
The New York Times	1	12	20	2	7	42
The Telegraph	2	20	23	6	7	58
The Times	2	262	43	3	27	337
Total	389	885	2142	899	1479	5794
Panel C: Euro-crisis						
Dow Jones Newswires	1936	2752	3311	31400	5473	44872
Thomson Reuters	939	1480	2523	14800	2718	22460
Financial Times	382	576	783	4674	1050	7465
The Wall Street Journal	301	398	675	4110	1024	6508
The New York Times	51	111	186	1146	258	1752
The Telegraph	407	301	1101	5005	1086	7900
The Times	136	872	499	2965	847	5319
Total	4312	6894	9461	66959	13057	100683

Table 7: Euro-periphery Information Flow Descriptive Statistics

Panel A: Descriptive Statistics

	Pre-crisis 1/1/2004 - 26/2/2007			Euro-crisis 8/12/2009 - 13/3/2013		
(Euro-periphery)	Pessimism	Relevance	SVI	Pessimism	Relevance	SVI
Mean (%)	0.127	1.285	–	0.394	22.046	9.403
Median (%)	0	1	–	0.096	5	9
Std. Dev. (%)	0.043	2.172	–	0.968	55.908	4.711
Minimum (%)	−0.086	0	–	−0.235	0	0
Maximum (%)	1.495	40	–	12.374	649	38.4

Panel B: Correlations

	Pre-crisis 1/1/2004 - 26/2/2007					
	Pessimism	Relevance	SVI	Euro-periphery	Euro-core	Non-euro
Pessimism	1.000					
Relevance	0.695	1.000				
SVI	–	–	–			
Euro-periphery	0.027	0.012	–	1.000		
Euro-core	0.025	0.021	–	0.869	1.000	
Non-euro	0.031	0.031	–	0.791	0.805	1.000

	Euro-crisis 8/12/2009 - 13/3/2013					
	Pessimism	Relevance	SVI	Euro-periphery	Euro-core	Non-euro
Pessimism	1.000					
Relevance	0.939	1.000				
SVI	0.594	0.521	1.000			
Euro-periphery	−0.154	−0.042	−0.137	1.000		
Euro-core	−0.152	−0.042	−0.124	0.873	1.000	
Non-euro	−0.159	−0.047	−0.132	0.836	0.926	1.000

Note: Pessimism = Euro-periphery News Pessimism factor, Relevance = Euro-periphery News Relevance factor, SVI = Euro-periphery Web Search Volume Index

Table 8: The impact of Information Flows about the Euro-periphery, Across groups for bottom and top coexceedances.

	Bottom tail			Top tail		
Panel A: News Pessimism factor						
	Non-euro	Euro-core	Euro-Periphery	Non-euro	Euro-core	Euro-Periphery
Pre-crisis		−	+			
Euro-crisis	+++++	++++	+++++			
Entire period	+++	+++	+++++		+	+++
$Pseudo-R^2$	0.003/0.044/0.007	0.005/0.039/0.011	0.005/0.039/0.025	0.003/0.002/0.001	0.000/0.005/0.003	0.001/0.004/0.007
Panel B: News Relevance factor						
	Non-euro	Euro-core	Euro-Periphery	Non-euro	Euro-core	Euro-Periphery
Pre-crisis	+					+
Euro-crisis	++++	+++	+++	+	++	++
Entire period	+	++	+++++		++	+++++
$Pseudo-R^2$	0.005/0.021/0.002	0.002/0.016/0.005	0.001/0.019/0.015	0.001/0.005/0.001	0.003/0.009/0.005	0.005/0.006/0.011
Panel C: Web Search Volume Index						
	Non-euro	Euro-core	Euro-Periphery	Non-euro	Euro-core	Euro-Periphery
Euro-crisis	++++	++++	++++			
$Pseudo-R^2$	0.034	0.044	0.042	0.002	0.003	0.002

Note: The number of "+" (or "-") indicate the number of statistically significant (in the 1% or 5% levels) and positive (or negative) marginal effects. In other words, for the bottom tail returns and the entire period sample, three out of five periphery News Pessimism marginal effects were found to be significant and positive for the Non-euro group, indicating that an increase of one unit in the periphery News Pessimism factor increases the probability of extreme bottom returns for the Non-euro group in three out of five bottom Non-euro coexceedances outcomes.

Table 9: Granger Causality Tests between Euro-periphery Information Flows, Stock Market Returns and Bottom Coexceedances

	Pre-crisis 1/1/2004 - 26/2/2007			Euro-crisis 8/12/2009 - 13/3/2013		
	Pessimism	Relevance	SVI	Pessimism	Relevance	SVI
Pessimism	–					
Relevance	$\leftarrow^{**}$	–				
SVI	–	–		$\leftarrow^{***}, \rightarrow^{***}$	$\leftarrow^{***}, \rightarrow^{**}$	
Euro-periphery Stock Returns	$\leftarrow^{***}$	$\rightarrow^{*}$		$\leftarrow^{***}, \rightarrow^{***}$	$\leftarrow^{**}, \rightarrow^{***}$	$\leftarrow^{***}$
Euro-core Stock Returns	$\leftarrow^{***}$	$\leftarrow^{**}$		$\leftarrow^{***}, \rightarrow^{***}$	$\rightarrow^{***}$	$\leftarrow^{***}$
Non-euro Stock Returns	$\leftarrow^{**}$	$\leftarrow^{***}$		$\leftarrow^{***}, \rightarrow^{***}$	$\leftarrow^{**}, \rightarrow^{***}$	$\leftarrow^{**}$
Euro-periphery Coexceedances	$\leftarrow^{**}$	$\leftarrow^{**}, \rightarrow^{**}$		$\leftarrow^{***}, \rightarrow^{**}$	$\leftarrow^{*}, \rightarrow^{**}$	$\leftarrow^{***}$
Euro-core Coexceedances	$\leftarrow^{**}$	$\leftarrow^{**}$		$\leftarrow^{***}, \rightarrow^{***}$	$\leftarrow^{**}, \rightarrow^{***}$	$\leftarrow^{***}$
Non-euro Coexceedances	$\leftarrow^{*}$			$\leftarrow^{**}, \rightarrow^{***}$	$\rightarrow^{***}$	$\leftarrow^{**}, \rightarrow^{**}$

Note: Pessimism = Euro-periphery News Pessimism factor, Relevance = Euro-periphery News Relevance factor, SVI = Euro-periphery Web Search Volume Index. ***, **, * denote significance at 1%, 5%, 10% levels. "$\leftarrow$" means that the column element "Granger-causes" the row element, "$\rightarrow$" means that the row element "Granger-causes" the column element, "$\leftarrow, \rightarrow$" means that both elements "Granger-cause" each other.

Fig. 1. Euro-periphery bottom coexceedances count and Information Flow factors