The Impact of Textual Sentiment on Sovereign Bond Yield Spreads: Evidence from the Eurozone Crisis

Sha Liu[*]
University of Southampton, UK

This study examines the relation between textual sentiment (media pessimism), the concentration/volume of news, and sovereign bond yield spreads, specifically in Greece, Ireland, Italy, Portugal and Spain during the European sovereign debt crisis from 2009 to 2012. The findings suggest that higher media pessimism and greater concentration/volume of news collectively communicate additional value-relevant information that has not been quantified by traditional determinants of yield spreads. If higher media pessimism is coupled with greater concentration/volume of news and other factors remain unchanged, yield spreads would move upwards, causing prices to fall. Media pessimism and the number of news stories respectively and collectively help predict the widening of yield spreads. Higher media pessimism level is strongly associated with more news stories being reported, suggesting that "no news is good news."(JEL: E43,G12,G14,G15)

Keywords : textual sentiment, media pessimism, information supply, sovereign bond yield spreads, European sovereign debt crisis

I. Introduction

Since the introduction of the Euro, a number of studies have investigated the determinants of sovereign bond yield spreads in the Eurozone. Traditionally, sovereign creditworthiness, sovereign bond

liquidity risk, and general risk aversion have been identified as the main factors that have a critical impact on the interest rates paid by governments relative to the benchmark government bond (see Codogno et al., 2003; Geyer et al., 2004; and Bernoth et al., 2006). According to the Efficient Market Hypothesis (EMH), asset prices should reflect all available information. Some information is quantitative and direct, such as public financial information, certain macroeconomic variables, and credit ratings of a country. Some quantitative variables are potential proxies of information, such as the credit default swap (CDS) premium (the seller of the CDS compensates the buyer in the event of a debt default or other credit events), liquidity measures (how quickly the bonds can be traded in the market to prevent a loss or make the required profit) and general risk aversion (how much investors would like to pay over the riskless asset). These variables may reflect the information that is hard to measure directly but has been processed by market participants.

In the emerging textual analysis literature, researchers argued that texts are an equally important source or proxy of news and information in comparison to quantitative measures. Verbal information can be complex, but sentiment (negativity/positivity) is now understood to be articulated in many forms of human discourse, news reports, blogs and other forms of written communication (Kearney and Liu, 2014). The studies on equities did find clear evidence of the effects of textual sentiment on asset price returns and trading volumes (see Tetlock, 2007; and Tetlock et al., 2008). This research examines to what extent the well-documented role of textual sentiment in equity markets applies to debt markets, by studying the long-term sovereign bond yield spreads of Greece, Ireland, Italy, Portugal and Spain (GIIPS) relative to Germany during the European sovereign debt crisis from 2009 to 2012. The countries of GIIPS have been threatened by a possible default and have seen the dramatic widening of their sovereign bond yield spreads as investors continually reassessed their default and other risks. During the Eurozone crisis period, there were a large volume of news stories reporting on credit rating adjustments, analyzing the causes of the crisis, determining appropriate fiscal policies, and discussing possible rescue packages and impacts on investor confidence. It is reasonable to argue that the content and complexity of news stories may not be completely absorbed into the aforementioned risk factors. Does textual sentiment have any additional impact on yield spreads? In the meantime, the concentration or volume of news may also be relevant to the movement

of yield spreads, as a greater volume of news indicates more information supply and possibly higher investor attention and different trading behaviors as compared with those relatively peaceful periods. As such, this research examines textual sentiment together with the concentration or amount of news. A total of 18,610 country-specific news stories are collected, from which the country-specific sentiment is extracted based on Loughran and McDonald's (2011a) finance negative word list.

This research confirms that, besides traditionally recognized factors, sentiment in news stories also conveys potential pricing-relevant information of sovereign bonds in the Eurozone. Particularly, media pessimism (i.e. negative textual sentiment) and the concentration/volume of news collectively play an important role in determining yield spreads. The concentration/volume of news is measured by the number of news stories or the number of words in a day. If the interaction term (the extent of media pessimism coupled with the concentration/volume of news) increases by 1 unit and other factors remain unchanged, yield spreads would move upward by approximately 18 to 32 basis points, putting downward pressure on prices. This effect is more powerful if the number of news stories is used as the measure of the concentration/volume of news. Media pessimism does not directly determine sovereign bond yield spreads, though. Logistic regressions reveal that the rise of media pessimism and the concentration/volume of news relative to their five-day moving averages respectively and collectively increase the likelihood that yield spreads widen relative to the previous day. These results are consistent with the story that investors of sovereign bonds respond to the heightening of media pessimism accompanied by an increased concentration/volume of news. From the other way around, the study reveals that greater yield spreads and/or general risk aversion have an obvious predictive ability over the rise of media pessimism relative to its five-day moving average, but their effects are much weaker than that of the number of news stories. Higher media pessimism is very likely to coexist with more news stories being reported, suggesting that "no news is good news." Fewer news stories in a day indicates less negative news content.

This study has made the following contributions to the previous literature: 1) it is the first to investigate the role of textual sentiment in debt markets and link the sovereign bond yield spreads literature with the textual sentiment literature. By and large, the direct effects of textual sentiment on sovereign bonds seem to be weaker than the effects on stock price returns; and 2) this study links textual sentiment and the

amount of information supply with market reactions. The existing studies of textual sentiment generally do not make any assumptions in relation to investors' cognitive ability or emotional biases. On the one hand, not much is known about the relationship between textual sentiment and investor sentiment (irrational behaviors). Textual sentiment can be associated with or even cause investor sentiment. This research provides some evidence of investor behaviors by finding that media pessimism alone does not determine sovereign bond yield spreads, but the market responds to higher media pessimism coupled with greater concentration/volume of news (the interaction term). This could reflect investor sentiment because media pessimism itself is a noisy measure of information, while the predictability of the interaction term over traditional determinants of yield spreads (proxies of credit risk, liquidity risk and general risk aversion) may suggest market "anomaly". On the other hand, from the rational point of view, textual sentiment could be a relatively objective measure of some material information that is difficult for conventional quantitative measures to proxy. From this perspective, the findings of this study suggest that higher media pessimism and greater amount of news collectively communicate additional value-relevant information that has not been quantified by traditional determinants of yield spreads, and a better empirical model of yield spreads should incorporate textual sentiment. On the whole, this paper suggests a novel perspective of studying the determinants of bond yield spreads or interest rates and arouses further discussions around the role of textual sentiment in asset pricing.

The rest of this paper is organized as follows. Section II presents a review of both the literature of sovereign bond yield spreads in the Eurozone and the literature of textual sentiment analysis in equity markets. Section III describes the data, including news stories, sentiment, yield spreads and other variables, and provides preliminary analysis of the data. Section IV reports the results of examining the relation between textual sentiment, the concentration/volume of news and yield spreads by using panel GMM (the General Method of Moment) regressions and logistic regressions. Section V summarizes the main findings and draws together the conclusions.

II. Related literatures

There is a relatively large body of literature exploring the determinants

of sovereign bond yield spreads in the EU. The yield spreads studied are usually the country's 10-year, 5-year, or 2-year government bond yield differences relative to Germany. Researchers have identified three main factors that determine yield spreads: sovereign credit risk, sovereign bond liquidity risk and general risk aversion

The role of sovereign credit risk and general risk aversion for either the pre-crisis or crisis period (after Sept 2008) has been confirmed by several studies. Studies that find significant effects of credit risk include Codogno et al. (2003), Heppke-Falk and Huefner (2004), Bernoth et al. (2006), Gomez-Puig (2006), Faini (2006), Beber et al. (2009), and Barbosa and Costa (2010), and Schwarz (2014). A number of papers, such as Geyer et al. (2004), Bernoth et al. (2006), Sgherri and Zoli (2009), Attinasi et al. (2009), Barrios et al. (2009), Haugh et al. (2009), Manganelli and Wolswijk (2009), and Gerlach et al. (2010) conclude that general risk aversion plays a significant role in driving sovereign bond yield spreads.

In relation to the role of liquidity, the results are more complicated and controversial. Geyer et al. (2004) find no significant effects from liquidity related variables. Codogno et al. (2003), Bernoth et al. (2006), Gomez-Puig (2006), Jankowitsch et al. (2006), Beber et al. (2009), Ejsing and Sihvonen (2009), Attinasi et al. (2009), Barrios et al. (2009), Haugh et al. (2009), Favero et al. (2010), Gerlach et al. (2010), and Schwarz (2014) find that liquidity effects are quantitatively limited and only relevant for some countries. Pagano and Von Thadden (2004), Jankowitsch et al. (2006), and Gomez-Puig (2006) find that liquidity matters and is related to the bond market's size. Gomez-Puig (2006) indicates that after the introduction of the euro, liquidity is a highly relevant factor in determining sovereign yield spreads, but Bernoth et al. (2006) find that the effect is no longer apparent following the EMU (Economic and Monetary Union). Moreover, Beber et al. (2009) conclude that credit risks often explain the largest part of bond spreads. In the case of heightened market uncertainty, however, liquidity considerations gain in importance.

Since the global economic and financial crisis in 2008, the most recent literature focuses on the impact of the crisis. Empirical findings from several studies (see Barrios et al., 2009; Ejsing and Sihvonen, 2009; Mody, 2009; and Barbosa and Costa, 2010) suggest that the importance of credit risk and liquidity risk has increased in the crisis period. Moreover, the importance of credit risk has increased to a greater extent than liquidity. Mody (2009), Sgherri and Zoli (2009), and

Schuknecht et al. (2011) also find that sovereign credit risk appears to be the major driver of the changes in sovereign yield spreads in the crisis period. Mody (2009) highlights that at the beginning of the financial market turmoil in the second half of 2007 and early 2008, spreads were largely determined by common risk factors. After March 2008, the different conditions of the countries' national financial sectors became determinants of the sovereign bond spreads. Attinasi et al. (2009) conclude from their results that the poor condition of the financial sector and the governments' announcement of bank rescue packages led to the transfer of credit risk from the financial sector to the public sector, and the size of the rescue package does not statistically affect the yield spreads.

No previous studies investigate the effects of textual sentiment on sovereign yield spreads. There are only a few relevant studies examining the effect of credit-related announcements. For example, Afonso et al. (2012) employ an event study analysis to investigate the reaction of government yield spreads before and after the announcements of major credit rating agencies for EU countries. They find that changes in both the credit rating notations and the outlook (with some differences across rating agency) lead to a significant response of government rating bond yield spreads, and the effect is particularly important for the case of negative announcements.

A. Textual sentiment analysis

There is an increasing body of literature that studies the role of textual sentiment in equity markets. Three text sources have been studied: corporate disclosures/filings, news stories and internet messages. The news stories category, which is most relevant to this paper, includes Tetlock (2007), Tetlock et al. (2008), Engelberg (2008), Sinha (2010), Garcia (2013), Ahmad et al. (2014), and Ferguson et al. (2014). Tetlock (2007) and Garcia (2013) study general economic and financial news on Wall Street Journal and New York Times, respectively. Ferguson et al. (2014) study firm-specific UK news media articles from The Financial Times, The Times, The Guardian and Mirror, and so on. Tetlock et al. (2008), Engelberg (2008), Sinha (2010) and Ahmad et al. (2014) employ much wider news sources, and they all research into firm-specific news.

A critical part of textual sentiment analysis is the computational content analysis method. The most frequently used methods in the literature are the dictionary-based approach, which uses a 'mapping' algorithm in which a computer program reads the text and classifies the

words (or phrases) in the text into different categories based on pre-defined dictionary categories (Li, 2010). Hitherto the most appropriate dictionary categories are Loughran and McDonald's (2011a) finance word lists, because they are developed specifically for the domain of business and finance.

The findings that are most relevant to this paper are summarized as follows. Textual sentiment has important effects on stock prices and returns. The literature has found strong evidence of the immediate effects of sentiment (Kearney and Liu, 2014). Tetlock (2007), Tetlock et al. (2008) and Garcia (2013) all concluded that negative sentiment or a large increase in negative sentiment causes downward pressure on market prices immediately. Ahmad et al. (2014) also find that an increase in negative media sentiment (or media pessimism) leads to lower next-day returns. In regard to the implications for market efficiency, Tetlock et al. (2008) conclude that negative words in news articles capture otherwise hard-to-quantify aspects of firms' fundamentals, and further state that the stock market is relatively efficient with respect to firms' hard-to-quantify fundamentals, although results of Tetlock (2007) show that the hypothesis that pessimism represents negative fundamental information not yet incorporated into prices receives very little support from the data. Using rolling regressions, Ahmad et al. (2014) show that the effect of sentiment on individual stocks is episodic in nature. These episodes are dominated by periods during which the effects of media pessimism are transitory rather than permanent. By and large, it has been agreed that sentiment in texts does convey some incremental information over quantitative financial information, and has at least some potential predictive power on market movements.

III. Data

A. *Data descriptions and variable construction*

The relevant variables include textual sentiment measures, sovereign bond yield spreads, measures of sovereign credit risk, liquidity, and general risk aversion. The description of news data is mentioned first.

News

Compared with general economic and finance news, news that focuses

on sovereign debt, fiscal conditions, credit ratings and rescue actions should be more relevant to sovereign bond yield spreads. Thus, sentiment is to be extracted from this category of news. Because of the scarcity of such news in normal economic periods, the European sovereign debt crisis between January 2009 and December 2012 is chosen as the sample period, because it experienced much more qualified news stories. Greece (GR), Ireland (IE), Italy (IT), Portugal (PT) and Spain (ES) are chosen as the countries being studied because they experienced the most difficult situations and had the greatest volume of news during this period.

News stories are downloaded from LexisNexis, a popular news database. For each country the search criteria is such that: 1) the headline contains the country name, and 2) there are at least five mentions of 'sovereign' or 'debt' in the news article. The sources chosen are 'All English News,' which consists of newspapers, magazines, newswires and press releases, web-based publications, blogs, and so on. Subject terms 'Public Finance' or 'Public Debt,' as well as the option 'Strong References Only' are selected to make sure that the news stories retrieved are highly relevant to public debt and government solvency. Finally, I select 'Duplication Options – On – High similarity' to remove the articles that are highly similar to stories published earlier.

Table 1 presents the number of news stories collected for each country, grouped by year. Greece had an increasing amount of news year by year, as its condition worsened. Ireland had the largest volume of news in 2010, when there was a marked increase in Irish two-year bond yields in April and at the year end when the EU-IMF rescue deal was reached. Italy had relatively fewer stories from 2009 to early 2011, but there was a surge of news in late 2011 accompanying the growing concerns of levels of debt and the resignation of Berlusconi IV Cabinet following the market pressure on government bond prices. Both Spain and Portugal had increasing news volumes from 2009 to 2011 as well, but they had only fewer than 100 news stories in 2009, when there were fewer problematic issues to report. Spain saw the greatest amount of news in 2012, as it gradually became the main concern of the Eurozone and a bailout package was reported in June to be under consideration. In October 2012, the Troika[1] was in negotiations with Spain to establish

---

1. The European Commission with the European Central Bank and the International

TABLE 1.   Number of news stories

	2009	2010	2011	2012
Greece	352	1,382	1,815	2,159
Ireland	450	1,058	753	474
Italy	62	190	2,021	877
Portugal	48	714	1,169	353
Spain	81	897	1,198	2,557

   Note: This table presents the number of news stories that report each country's sovereign creditworthiness, fiscal conditions, debt crisis, etc.

an economic recovery program. A total of 18,610 articles are collected.

   Skeptical readers may question whether or not a number of the news stories in the sample recount the past movement of sovereign bonds. If this was the case, it may cause estimation problems due to endogeneity. Dozens of news stories for each country are examined and few such news stories are found. Therefore, this issue is not a serious concern in the sample.

Sentiment Measures

After the news stories are downloaded, the dictionary-based approach is employed to extract sentiment. A custom content analysis program that counts the number of words in news stories in a specified period (e.g., by minute, hourly, daily, or yearly) that fall within Loughran and McDonald's (2011a) finance negative word list[2] is used. This word list is chosen because it is more appropriate than general negative word lists, and has already become popular in more recent studies (see Loughran and McDonald, 2011b; Garcia, 2013; Jegadeesh and Wu, 2013; and Loughran and McDonald, 2013; Ahmad et al., 2014).

   For each country, the daily value of the basic sentiment score (BSC) is the fraction of negative words relative to the total number of words in a given day. All articles on weekends are moved to Friday and the very few articles published on holidays are omitted. There are many missing values due to no news coverage on those days. Two sentiment

Monetary Fund

   2.   This word list contains 2349 negative words that occur in 5% or more of the 10-Ks from 1994 to 2008 (Loughran and McDonald, 2011a).

TABLE 2.   Descriptive statistics of sentiment measures

Country	Measure	Obs	Mean	Variance	Skewness	Kurtosis	J-B P-val	Q10 P-val
Greece	NEG_Z	737	0.0000	1.0000	–0.0770	3.6912	0.00	0.00
	NEG_DTR	1,039	0.0021	1.1675	0.1737	4.0007	0.00	0.00
Ireland	NEG_Z	744	0.0003	1.0014	0.1718	2.8911	0.13	0.00
	NEG_DTR	1,039	–0.0023	1.2557	0.2755	2.8885	0.00	0.00
Italy	NEG_Z	515	0.0001	1.0008	0.9835	11.6267	0.00	0.01
	NEG_DTR	1,039	0.0032	1.2146	0.8080	6.6149	0.00	0.00
Portugal	NEG_Z	580	0.0001	1.0008	0.6478	5.2601	0.00	0.00
	NEG_DTR	1,039	0.0001	1.2284	0.9974	6.3835	0.00	0.00
Spain	NEG_Z	699	0.0001	1.0008	0.2523	5.1563	0.00	0.00
	NEG_DTR	1,039	0.0000	1.1716	0.6244	4.6987	0.00	0.00

Note: This table provides descriptive statistics of two sentiment measures, including mean, variance, skewness, kurtosis, the Jarque-Bera (J-B) test of normality, and the Ljung-Box Q test for autocorrelation.

measures are then constructed. The first is the standardized negative sentiment score (the Z score) for country i on day t, as shown in equation (1). This measure still contains missing values. The second measure is constructed through two steps. First, let the value of BSC in a day be 0 if there is no news coverage on that day, and then calculate the five-day moving average of BSC (from day t–4 to day t). Second, the detrended sentiment measure (equation (2)) is computed as the difference of BSC and the five-day moving average. The choice of 5 days is arbitrary. This measure is consecutive by construction.

$$NEG\_Z_{i,\,t} = \frac{BSC_{i,t} - Sample\,mean\,of\,BSC_i}{Sample\,standard\,deviation\,of\,BSC_i} \tag{1}$$
$$i = GR, IF, IT, PT \text{ or } ES$$

$$NEG\_DTR_{i,t} = BSC_{i,t} - (1/5)* \sum_{T=t-4}^{t} BSC_{i,t} \tag{2}$$
$$i = GR, IE, IT, PT \text{ or } ES$$

The descriptive statistics of the sentiment series are reported in table 2. For each country, the standardized sentiment scores have a zero mean and unit variance by construction, and the mean of the detrended sentiment series is also close to zero. Ireland has a slightly higher variance of the detrended sentiment than the other countries. The J-B column denotes the p-value of tests against the null hypothesis that the sentiment series is normally distributed. The results clearly show that none of the sentiment series are normally distributed except for the standardized sentiment scores of Ireland. The Q10 column displays the p-value of tests for autocorrelation up to the 10th lag. The results indicate that every sentiment series is serially correlated.

Yield spread (the dependent variable)

The measure of sovereign bond yield spreads for each country is as follows:

$$YS_{i,t} = yield_{i,t} - yield_{BD,t}, \text{ where } i = GR, IE, IT, PT \text{ or } ES \tag{3}$$
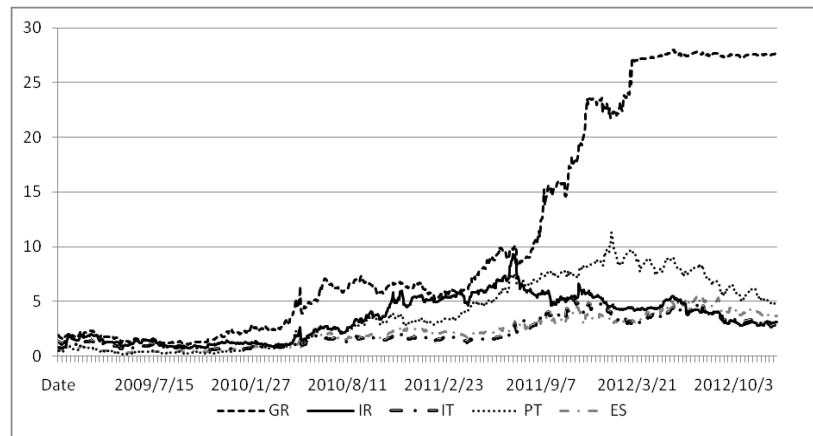
FIGURE 1. Sovereign 10 +− year bond yield spreads (percent) relative to Germany

Note: Data source: Datastream and author's calculation

The sovereign bond yield spread is a country's 10+ − year sovereign bond index[3] ($yield_{i,t}$) minus the Germany (*BD*) counterpart ($yield_{BD,t}$). It is a daily time series for each country. Figure 1 plots the yield spreads for each country during the sample period. Greece began a sharp rise since 2010, until it reached an astonishing 27.975 percent in June 2012. Ireland reached the highest point at 9.332 percent in July 2011, and then gradually lowered to around 3–4 percent at the end of 2012. Italy began an obvious rise from the second half of 2011, moving from less than 2 percent to the highest point at 5.312 percent in November 2011. Both Portugal and Spain began a steady rise from early 2010. They reached the highest point at 11.346 percent in January 2012 and 5.641 percent in July 2012, respectively.

Measures of sovereign credit risk

The first sovereign credit measure is the *CDS* premium. *CDS* is a financial swap agreement in which the seller of the *CDS* compensates the buyer in the event of a debt default or other credit event. The more

---

3. The indeces used are FTSE global government bond sub-indices, and the index value is the average yield of all bonds in the sub-index portfolio.

likely the buyer is to default, the higher the premium that the seller asks for. Therefore, the *CDS* premium can be considered as a proxy for credit risk. The measure used in this research, denoted as *CDS*, is each country's sovereign 10–year bond *CDS* premium minus Germany's counterpart.[4]

The credit-related macroeconomic and fiscal variables that have appeared in the literature are usually available at lower frequencies. To coordinate with other daily variables, only variables that are available in the Economists Intelligence Unit or Eurostat databases at quarterly frequency (the highest frequency available) are selected. These variables are the current account balance (percent of *GDP*) and industrial production (percent change per annum), denoted as *CAB* and IND respectively.[5] The two quarterly variables are transformed to daily variables by assuming that the values are the same within each quarter. They are also stated as the difference relative to Germany.

Measure of liquidity

Similar to Sgherri and Zoli (2009), the proxy for liquidity (*LIQ*) used in this research is the market value of long-term sovereign bond. Specifically, it is the market value[6] (in US dollars) of the Bank of America Merrill Lynch (BOFA ML) Government 10+Y Bond Index (sub-indices for each country), stated as the difference against Germany.

General risk aversion

In line with the literature, the common risk factor (i.e., general risk aversion) is measured by the difference between the US corporate bond yield and Treasury bond yield. In particular it is the difference between the yield of BOFA ML US High-Yield 1–10Y corporate bond index and the yield of BOFA ML US 1–10Y Treasury bond index, denoted as *RISK*.

---

4. Divided by 100

5. Another quarterly variable, the government consolidated gross debt (percent of GDP), is found to be non-stationary and not suitable for the first-difference transformation; hence, it is excluded.

6. Divided by 10,000

TABLE 3.   Panel data unit-root tests

A.

Variable	LLC	Breitung	IPS	HT
YS	−0.90	2.01	1.96	1.00
NEG_Z	N/A	N/A	−34.23***	N/A
NEG_DTR	−61.21***	−48.36***	−55.14***	0.04***
CDS	1.40	1.06	1.45	1.00
CAB	0.13	−3.60***	−1.98**	0.99***
IND	−1.55*	−3.21***	−3.83***	0.99***
LIQ	0.71	−0.01	1.06	1.00*
RISK	−10.07***	4.49	−11.26***	0.99***

B.

D(YS)	−53.08***	−44.08***	−52.83***	0.07***
D(CDS)	−46.40***	−46.18***	−50.94***	0.15***
D(LIQ)	−52.97***	−47.31***	−55.01***	0.05***
D(RISK)	−38.36***	−15.17***	−44.43***	0.36***
D(NEG_Z)	N/A	N/A	−43.24***	N/A
D(NEG_DTR)	−91.73***	−60.61***	−68.07***	−0.42***

Note: This table reports the statistics and the significance level (1 percent ***, 5 percent **, 10 percent *) of panel unit root tests on each variable. The Levin–Lin–Chu (LLC) test, the Breitung test, the Im–Pesaran–Shin (IPS) test and Harris–Tzavalis (HT) test are undertaken. The IPS statistics listed are the *Z*-t-tilde-bar statistics. The LLC, Breitung and HT tests require strongly-balanced panels, hence they are not applicable to *NEG_Z*. Variables shown in panel (b) are the first-differenced variables.

B. Stationarity tests

Pooled regressions are implemented in the study because the sample period is relatively short for individual countries, and there are many days with no news coverage. Panel unit-root tests are taken on each data series to examine if they are stationary. Four tests are employed: the Levin-Lin-Chu (2002) (*LLC*) test, the Breitung (2000) test, the Harris-Tsavalis (1999) (*HT*) test and the Im-Pesaran-Shin (2003) test. Each test has its benefits and drawbacks, and no one particular test is dominant. Table 3 panel (a) reports the statistics and the significance level (1 percent ***, 5 percent **, 10 percent *). *YS*, *CDS* and *LIQ* are found to be non-stationary for any of the tests. Both sentiment measures (*NEG_Z* and *NEG_DTR*) are stationary at the 1 percent level of significance. Results are mixed for general risk aversion (*RISK*) and the two macroeconomic variables, *CAB* and *IND*. Given the results, *YS*, *CDS*, *LIQ* and *RISK* are then transformed to the first-differenced

variables. To be consistent, the sentiment measures are also transformed. Panel (b) of table 3 shows that the 1 percent level. *CAB* and *IND* are left unchanged, because if taking first difference, the majority of their daily values would become 0, rendering these controlling variables useless.

IV. Textual sentiment, amount of information and yield spreads

A. Hypotheses

The main hypotheses tested in this study are as follows. They are related to the effects of textual sentiment and the concetration / volume of news on sovereign bond yield spreads.

$H_1^0:$ Negative textual sentiment does not significantly affect sovereign bond yield spreads.

$H_1^a:$ Negative textual sentiment is significantly directly related to sovereign bond yield spreads.

$H_2^0:$ The concentration/volume of news does not significantly affect sovereign bond yield spreads.

$H_2^a:$ The concentration/volume of news is significantly directly related to sovereign bond yield spreads.

$H_3^0:$ The interaction of negative textual sentiment and the concentration/volume of news does not significantly affect sovereign bond yield spreads.

$H_3^a:$ The interaction of negative textual sentiment and the concentration/volume of news is significantly directly related to sovereign bond yield spreads.

The study employs both linear *GMM* regressions and logistic regressions (nonlinear) to test the above hypotheses. It also examines the reversal direction: what factors help predict the movement of media pessimism? In this regard, yield spreads, credit risk measures, the

liquidity measure, general risk aversion, as well as the number of news
stories and the number of words are considered.

B. The impact of textual sentiment on yield spreads

The most direct relationship that one could postulate is that textual
sentiment is another determinant of sovereign bond yield spreads
besides sovereign credit risk, liquidity risk and general risk aversion.
Other possible control variables include the number of news stories
(*N_stories*) and the number of words (*N_words*)[7], because they reflect
the concentration of news or the amount of information, and possibly
the attention of investors worldwide. Therefore, the following equation
is examined:

$$D(YS)_{i,t} = \alpha_{10} + \alpha_{11} \cdot D(YS)_{i,t-1} + \alpha_{12} \cdot D(NEG)_{i,t} +$$
$$\alpha_{13} \cdot D(News\_Amount) + \alpha_{14} \cdot CR_{i,t} + \qquad (4)$$
$$\alpha_{15} \cdot D(LIQ)_{i,t} + \alpha_{16} \cdot D(RISK)_t + \varepsilon_{i,t}$$

where the dependent variable is the first difference of the yield spreads
measure, α_{10} is the constant, *NEG* denotes one of the sentiment
measures, the amount of news (*News_Amount*) is proxied by either the
number of news stories or the number of words, and *CR* denotes one of
the credit risk measures: *D(CDS)*, *CAB* or *IND*. The other variables have
been defined in the previous section.

The estimation method used for equation (4) and equation (5) below
is the two-step feasible *GMM* (Hayashi, 2000, Chapter 3), with
heteroskedasticity and autocorrelation robust standard errors. The
country dummy variables are used as the excluded instruments for the
first lag of yield spreads. The reasons to choose *GMM*, rather than
general least squares (*GLS*) are that the presence of the lagged
dependent variable on the right hand side of equation (4) and (5) gives
rise to autocorrelation, and yield spreads may predict negative sentiment
in news stories, leading to endogeneity issue in equation (4) and (5).
The *GLS* estimators are inappropriate for these reasons.

In table 4, the exact media pessimism measure (*NEG_Z* or
NEG_DTR) used is indicated at the second to last line of the table. The
results show that whatever measure is used, media pessimism does not

----

7.   Divided by 10,000

directly determine yield spreads. The number of news stories has no significant effect on yield spreads either. Replacing the number of news stories with the number of words (not reported), the results are qualitatively similar. The respective results for the two sentiment measures may serve as a robustness check. Consistent with the literature, liquidity measures and general risk aversion are found to have strongly significant impacts on yield spreads, all at the 1 percent level. The credit risk measures do not show any significant impact on yield spreads, however.

The next analysis examines whether media pessimism, together with the concentration/volume of news, has any impact on yield spreads. The number of news stories or the number of words per day reflects the amount of information supply in the market, and possibly indicates market participants' attention on the specific issues of an individual country. One can briefly see from figure 1 and table 1 that the overall trend of yield spread movements is related to the number of news stories reported. The amount of news could to some extent affect market participants' assessment of risk and trading behaviors, especially when the news is rather negative. As a result, the interaction of media pessimism and the amount of news may affect investors' required returns on the sovereign bonds. The following model is assessed:

$$D(YS)_{i,t} = \alpha_{20} + \alpha_{21} \cdot D(YS)_{i,t-1} + \alpha_{22} \cdot D(NEG * News\_Amount)_{i,t} +$$

$$(5)$$

$$\alpha_{23} \cdot CR_{i,t} + \alpha_{24} \cdot D(LIQ)_{i,t} + \alpha_{25} \cdot D(RISK)_t + \varepsilon_{i,t}$$

where each variable is defined as previously. These equations are also estimated by the two-step feasible *GMM*, with heteroskedasticity and autocorrelation robust standard errors. Table 5 displays the estimated coefficients and their corresponding t-statistics. Aside for the interaction of *NEG_Z* and *N_words* (regressions (3) and (4)), the other three measures of sentiment and information interaction are positively and significantly associated with sovereign bond yield spreads, at least at the 5 percent level. The interaction between *NEG_DTR* and the detrended *N_stories* has the strongest explanatory power on yield spreads, at the 1 percent level. If the interaction measures (the extent of media pessimism coupled with the concentration/volume of news) increases by 1 unit and other factors remain unchanged, yield spreads would rise by approximately 18 to 32 basis points. There is strong evidence that when

TABLE 4.   Does media pessimism directly determine yield spreads?

Dependent variable: $D(YS)_t$

	(1)	(2)	(3)	(4)	(5)	(6)
$D(YS)_{t-1}$	0.7899	0.7407	0.9044	0.8323	0.9123	0.9908
	(2.67***)	(2.11**)	(2.92***)	(2.61***)	(2.49***)	(2.90***)
$D(NEG)_t$	0.0023	0.0023	0.0022	0.0020	0.0021	0.0021
	(0.72)	(0.72)	(0.63)	(1.41)	(1.37)	(1.30)
$D(CDS)_t$	0.0113			0.0114		
	(1.59)			(1.60)		
CAB_t		−0.0008			−0.0003	
		(−0.97)			(−0.60)	
IND_t			0.0000			0.0002
			(0.07)			(1.01)
$D(LIQ)_t$	−0.1337	−0.1356	−0.1336	−0.1027	−0.1030	−0.1024
	(−8.18***)	(−7.97***)	(−7.90***)	(−10.11***)	(−9.66***)	(−9.65***)
$D(RISK)_t$	0.2559	0.2726	0.2654	0.1477	0.1512	0.1491
	(4.44***)	(4.57***)	(4.27***)	(5.44***)	(5.04***)	(5.06***)
$D(N\_stories)$	−0.0008	−0.0007	−0.0008	−0.0006	−0.0006	−0.0006
	(−0.74)	(−0.62)	(−0.70)	(−0.53)	(−0.46)	(−0.47)
Constant	−0.0004	−0.0068	−0.0004	0.0012	−0.0012	0.0019
	(−0.12)	(−0.95)	(−0.11)	(0.62)	(−0.31)	(0.96)
NEG measure	NEG_Z	NEG_Z	NEG_Z	NEG_DTR	NEG_DTR	NEG_DTR
Observations	2,586	2,586	2,586	5,190	5,190	5,190

(*Continued*)

TABLE 4. (Continued)

Note: 1) This table reports the estimated coefficients and their respective z-statistics (in the brackets). The equations are estimated by the two-step feasible GMM, and the standard errors are robust to heteroskedasticity and autocorrelation. The excluded instruments are the country dummy variables. The symbols *, **, and *** indicate significance level of 10 percent, 5 percent and 1 percent, respectively. 2) *NEG* is the media pessimism measure. The two specific measures used are *NEG_Z* and *NEG_DTR*. The former is the standard negative sentiment score, and the latter is the detrended sentiment score. 3) In regressions (4) to (6), the *N_stories* measure is detrended against its five-day moving average.

TABLE 5.   The interaction of media pessimism and the concentration/volume of news

Dependent variable: $D(YS)_t$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$D(YS)_{t-1}$	0.7840	0.7342	0.7868	0.7424	0.8311	0.9134	0.8297	0.9121
	(2.68***)	(2.11**)	(2.66***)	(2.11**)	(2.62***)	(2.50***)	(2.61***)	(2.49***)
$D(NEG*N\_stories)_t$	0.0032	0.0031			0.0028	0.0028		
	(2.28**)	(2.16**)			(2.80***)	(2.71***)		
$D(NEG*N\_words)_t$			−0.0044	−0.0041			0.0018	0.0018
			(−1.14)	(−1.04)			(2.20**)	(2.04**)
$D(CDS)_t$	0.0117		0.0113		0.0115		0.0116	
	(1.61)		(1.60)		(1.61)		(1.63*)	
CAB_t		−0.0008		−0.0008		−0.0003		−0.0003
		(−1.02)		(−0.96)		(−0.6)		(−0.61)
$D(LIQ)_t$	−0.1299	−0.1323	−0.1342	−0.1360	−0.1018	−0.1021	−0.1021	−0.1025
	(−8.33***)	(−8.15***)	(−8.16***)	(−7.98***)	(−10.16***)	(−9.68***)	(−10.15***)	(−9.68***)
$D(RISK)_t$	0.2541	0.2715	0.2584	0.2748	0.1477	0.1512	0.1477	0.1512
	(4.50***)	(4.65***)	(4.51***)	(4.61***)	(5.47***)	(5.07***)	(5.46***)	(5.06***)
Constant	−0.0002	−0.0069	−0.0005	−0.0068	0.0012	−0.0012	0.0012	−0.0013
	(−0.06)	(−0.97)	(−0.14)	(−0.95)	(0.63)	(−0.31)	(0.63)	(−0.32)
NEG measure	NEG_Z	NEG_Z	NEG_Z	NEG_Z	NEG_DTR	NEG_DTR	NEG_DTR	NEG_DTR
Observations	2,586	2,586	2,586	2,586	5,190	5,190	5,190	5,190

(*Continued*)

TABLE 5. (Continued)

Note: 1) This table reports the estimated coefficients and their respective *z*-statistics (in the brackets). The equations are estimated by the 2-step feasible GMM, and the standard errors are robust to heteroskedasticity and autocorrelation. The excluded instruments are the country dummy variables. The symbols *, **, and *** indicate significance level of 10 percent, 5 percent and 1 percent, respectively. 2) *NEG* is the media pessimism measure. The two specific measures used are *NEG_Z* and *NEG_DTR*. The former is the standard negative sentiment score, and the latter is the detrended sentiment score. 3) In regressions (5) to (8), the *N_stories* and *N_words* measures are detrended against their five-day moving average.

textual sentiment is more negative and there is a greater amount of news, sovereign bond yield spreads are higher. Both the liquidity measures and general risk aversion consistently and significantly affect yield spreads at the 1 percent level. Regarding the sovereign credit risk measures, the *CDS* premium at best marginally explains the yield spreads, while current account balance and industrial production have no significant effects.[8]

Skeptical readers may point out that news stories in the sample are from print media (e.g. newspapers, magazines), newswires and press releases, web-based publications (i.e. web news), and blogs, but the sample does not include real-time television news and internet messages. This neglects the role of television and other potential information sources (e.g. Twitter) which are faster in spreading news than the aforementioned sources, especially on the key event days. For instance, the bailout agreement of Greece may be immediately broadcasted via television at the press conference and spread on Twitter, and the market reacts instantly. To circumvent this concern, I exclude all key event days from the estimation period and replicate the regressions in table 4 and table 5. The key events identified are ratings adjustments, announcements of bailout package or changes, and announcements of austerity measures.[9] The results show that excluding key event days does not lead to any material changes in estimations. Textual sentiment does not directly determine sovereign bond yield spreads (not reported), but media pessimism, coupled with the concentration/volume of news, has significant impact on yield spreads (table 6).

Does textual sentiment about sovereign debt, fiscal conditions, and credit ratings have anything to do with general consumer sentiment? The correlation coefficients between each country's consumer confidence index (monthly indicators transformed to daily variables by assuming that the values are the same within each month) and the two measures of textual sentiment are calculated. The unreported results reveal that there is very low correlation between textual sentiment and

---

8. Estimation results where industrial production is employed as the credit risk measure are not presented due to space limit. The results are qualitatively the same and available upon request.

9. I refer to the following three lists to determine the key events during the Eurozone Crisis: http://cfa.is/1dFD6D7, http://bit.ly/1aaV4Lm, http://bit.ly/JutlxT. A total of 129 key event days are determined for the five countries.

general consumer sentiment. If general consumer sentiment is incorporated as an additional independent variables in equation (5), it does not materially affect any results presented in table 5.[10]

Logistic regressions

In the following step, logistic regressions are employed to further analyze the relation between media pessimism, the concentration/volume of news, and sovereign yield spreads. The idea is to test if the direction of sentiment and the concentration of information movement during a certain period help identify the direction of yield spreads movement relative to the previous trading day. Logistic regressions do not assume a linear relationship between the dependent and independent variables.

The dependent variable (*YSUP*) is set to equal 1 if the yield spread today rises relative to yesterday; otherwise it is set to be 0. On the right hand side, a dichotomous variable *NEG_up* is created, which has value 1 if the BSC score is above its five-day moving average (from day *t*–4 to day *t*)[11], and value 0 otherwise. Using the same method, *Stories_up* and *Words_up* are created to indicate the direction of movements for the number of news stories and the number of words. The interaction of *NEG_up* and *Stories_up* (*Words_up*) indicates that both media pessimism and the number of news stories (the number of words) increase relative to their five-day moving averages. The controlling variables include the *CDS* premium, liquidity risk measure and general risk aversion as usual.

Table 7 reports the heteroskedasticity robust results. Specific country effects are controlled by adding country dummies. Overall, the likelihood that yield spreads widen relative to the previous day increases by 21 percent[12] if media pessimism is above its five-day moving average; this is significant at the 1 or 5 percent level (column (1), (3)

---

10. These results are not reported because this analysis can only be treated as supplemental due to the nonstationarity nature of the general consumer sentiment series. If we take the first difference, the majority of the daily values would become 0, rendering this controlling variable useless. These results are available upon request.

11. Equivalent to NEG_DTRt >0

12. The odds ratio is $\exp(0.187) = 1.21$. The rest of the odds ratio calculations are omitted.

TABLE 6. The interaction of media pessimism and the concentration/volume of news–excluding key–event days

Dependent variable: $D(YS)_t$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$D(YS)_{t-1}$	0.8511	0.8199	0.8752	0.8656	0.9109	0.9639	0.9119	0.9683
	(2.54***)	(2.19**)	(2.57***)	(2.25**)	(2.34**)	(2.36**)	(2.33**)	(2.34**)
$D(NEG*N\_stories)_t$	0.0037	0.0037			0.0027	0.0028		
	(2.16**)	(2.16**)			(2.47***)	(2.49***)		
$D(NEG*N\_words)_t$			−0.0047	−0.0046			0.0020	0.0020
			(−1.00)	(−0.97)			(2.12**)	(2.13**)
$D(CDS)_t$	0.0088		0.0086		0.0085		0.0085	
	(1.09)		(1.06)		(1.03)		(1.03)	
CAB_t		−0.0006		−0.0005		−0.0002		−0.0002
		(−0.7)		(−0.57)		(−0.45)		(−0.41)
$D(LIQ)_t$	−0.1266	−0.1282	−0.1306	−0.1316	−0.0986	−0.0989	−0.0988	−0.0991
	(−7.83***)	(−7.69***)	(−7.69***)	(−7.52***)	(−9.49***)	(−9.12***)	(−9.48***)	(−9.11***)
$D(RISK)_t$	0.2469	0.2575	0.2494	0.2587	0.1372	0.1395	0.1378	0.1399
	(4.10***)	(4.19***)	(4.08***)	(4.12***)	(4.76***)	(4.50***)	(4.78***)	(4.50***)
Constant	−0.0008	−0.0056	−0.0014	−0.0055	0.0009	−0.0010	0.0010	−0.0008
	(−0.20)	(−0.74)	(−0.38)	(−0.72)	(0.41)	(−0.23)	(0.43)	(−0.19)
NEG measure	NEG_Z	NEG_Z	NEG_Z	NEG_Z	NEG_DTR	NEG_DTR	NEG_DTR	NEG_DTR
Observations	2,481	2,481	2,481	2,481	5,061	5,061	5,061	5,061

(*Continued*)

TABLE 6. (Continued)

Note: 1) This table reports the estimated coefficients and their respective z-statistics (in the brackets). The sample excludes the days where there are key events, including ratings adjustment, bailout agreement, and announcements of the austerity package and details. The equations are estimated by the 2-step feasible GMM, and the standard errors are robust to heteroskedasticity and autocorrelation. The excluded instruments are the country dummy variables. The symbols *, **, and *** indicate significance level of 10 percent, 5 percent and 1 percent, respectively. 2) *NEG* is the media pessimism measure. The two specific measures used are *NEG_Z* and *NEG_DTR*. The former is the standard negative sentiment score, and the latter is the detrended sentiment score. 3) In regressions (5) to (8), the *N_stories* and *N_words* measures are detrended against their five-day moving average.

TABLE 7. Logit regressions: does textual sentiment help predict the increase of yield spreads?

Dependent variable: $YSUP_t$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
NEG_up_t	0.1865 (3.08***)		0.1394 (2.06**)			0.1705 (2.62***)	
$Stories\_up_t$		0.1704 (2.76***)	0.1045 (1.51)				
$Words\_up_t$					0.1064 (1.71*)	0.0443 (0.66)	
$NEG\_up_t * Stories\_up_t$				0.1711 (2.50***)			
$NEG\_up_t * Words\_up_t$							0.0954 (1.34)
$D(CDS)_t$	0.0238 (0.73)	0.0244 (0.75)	0.0237 (0.73)	0.0246 (0.76)	0.0262 (0.78)	0.0240 (0.74)	0.0251 (0.77)
$D(LIQ)_t$	-2.5636 (-10.42***)	-2.5637 (-10.41***)	-2.5645 (-10.41***)	-2.5605 (-10.41***)	-2.5087 (-10.36***)	-2.5561 (-10.41***)	-2.5632 (-10.42***)
$D(RISK)_t$	4.8138 (11.48***)	4.8286 (11.51***)	4.8122 (11.48***)	4.8340 (11.53***)	4.5579 (10.63***)	4.8083 (11.46***)	4.8393 (11.54***)
Constant	0.1028 (1.41)	0.1224 (1.71)	0.0848 (1.15)	0.1432 (2.05**)	0.1465 (2.06**)	0.0936 (1.26)	0.1650 (2.38**)
Observations	5,195	5,195	5,195	5,195	5,210	5,195	5,195
Pseudo R^2	0.11	0.11	0.11	0.11	0.11	0.11	0.11
Country dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes

(*Continued*)

TABLE 7. (Continued)

Note: 1) This table reports the estimated coefficients and their respective z-statistics (in the brackets) of logistic regressions. The standard errors are robust to heteroskedasticity. The dependent variable is *YSUP*, which has value 1 if the yield spread today is greater than yesterday, and value 0 otherwise. Country dummy variables are employed to control for individual country effects. The symbols *, **, and *** indicate significance level of 10 percent, 5 percent and 1 percent, respectively. *NEG_up* has value 1 if the negative sentiment score is above its five-day moving average, and value 0 otherwise. *tories_up* and *Words_up* are similar dichotomous variables created to indicate if the number of news stories or the number of words is above its five-day moving average.

and (6)). The increase of the number of news stories alone contributes to the likelihood of yield spreads' increase by 19 percent, significant at the 1 percent level (column (2)), but the effect vanishes if *NEG_up* is incorprated (column (3)). The simultaneous increase of media pessimism and the number of news stories significantly adds to the likelihood of yield spreads' increase by 19 percent, at the 1 percent significance level (column (4)). The upward movement of the number of words alone marginally contributes to the increase of yield spreads (column (5)), and the effect also vanishes if *NEG_up* is added (column (6)). The interaction between *NEG_up* and *Words_up* does not predict the direction of yield spread movement (column (7)). The liquidity measure and general risk aversion consistently and significantly predict the increase of yield spreads as expected.

Aggregating the results from table 4 to table 7, it can be concluded that textual sentiment or the concentration/volume of news does not directly determine sovereign bond yield spreads, but there is much more evidence that the interaction of media pessimism with the concentration or amount of news has significant impact on yield spreads. At the same time, although sentiment in the news alone does not have explanatory power on yield spreads, its direction of movement relative to its five-day moving average helps identify the direction of yield spread movement. If the number of news stories increases, or if both the number of news stories and media pessimism increase relative to their five-day average level, the likelihood that yield spreads widen relative to the previous day also increases.

C. What factors help predict the movement of textual sentiment?

As mentioned in Section III.A, few news stories in the sample recount previous market movements. Textual sentiment is exogenous to yield spreads, the *CDS* premium, market liquidity, and general risk aversion. The unreported results of *GMM* regressions show that none of these factors have direct explanatory power on textual sentiment. Nonetheless, the question remains: Do they help predict the trend of negative sentiment in the news stories? Do yield spreads affect the amount of information supply?

Logistic regressions are employed to examine these issues. *NEG_up*, *Stories_up* and their interaction term are employed separately as the dependent variable. The results are presented in table 8, in which the first row specifies the dependent variables. Greater yield spreads and

TABLE 8. Logit regressions: what factors help predict the movement of media pessimism?

Dependent variable	(1) NEG_up_t	(2) NEG_up_t	(3) $Stories\_up_t$	(4) $Stories\_up_t$	(5) $NEG\_up_t * Stories\_up_t$
$Stories\_up_t$		2.0580			
		(30.75***)			
NEG_up_t				2.0577	
				(30.74***)	
$D(YS)_t$	0.6817	0.4716	0.6777	0.4479	0.9903
	(3.12***)	(1.71*)	(2.83***)	(1.50)	(4.12***)
$D(CDS)_t$	0.0225	0.0277	–0.0010	–0.0136	–0.0036
	(1.03)	(0.93)	(–0.05)	(–0.45)	(–0.19)
$D(LIQ)_t$	0.0327	–0.0169	0.1060	0.1121	0.1610
	(0.36)	(–0.16)	(1.14)	(1.10)	(1.59)
$D(RISK)_t$	0.7715	0.6450	0.6006	0.3226	0.4782
	(2.69***)	(2.05**)	(2.06**)	(0.99)	(1.50)
Constant	–0.2501	–1.0180	–0.5542	–1.5865	–1.1561
	(–3.95***)	(–12.94***)	(–8.39***)	(–17.93***)	(–15.56***)
Observations	5,195	5,195	5,195	5,195	5,195
Pseudo R^2	0.02	0.17	0.01	0.17	0.01
Country dummies	Yes	Yes	Yes	Yes	Yes

Note: 1) This table reports the estimated coefficients and their respective z-statistics (in the brackets) of logistic regressions. The standard errors are robust to heteroskedasticity. The dependent variables are specified in the first row. Country dummy variables are employed to control for individual country effects. The symbols *, **, and *** indicate significance level of 10 percent, 5 percent and 1 percent, respectively. NEG_up has value 1 if the negative sentiment score is above its five-day moving average, and value 0 otherwise. $Stories\_up$ is a similar dichotomous variable created to indicate if the number of news stories is above its five-day moving average.

general risk aversion help predict that negative sentiment moves upward relative to its five-day moving average (column (1)). One unit of their separate increase leads to 97 percent and 116 percent increase of the likelihood respectively. Their effects diminish if *Stories_up* is added to the independent variables. If the number of news stories is above its five-day average, the likelihood that media pessimism moves upwards would increase by 6.83 times (column (2)). Greater yield spreads and general risk aversion indicate higher possibility that the number of news stories reported in a day is above its five-day average level (column (3)). These effects completely vanish if *NEG_up* is also employed as an independent variable. If media pessimism is above its five-day average level, the likelihood that the number of news stories moves upwards would increase by 6.83 times (column (4)). Greater yield spreads increase the possibility that both media pessimism and the number of news stories are above their latest five-day average level by 1.69 times (column (5)). Replacing *Stories_up* with *Words_up* and/or using *CAB* or IND to control for credit risk yields qualitatively similar results.

To summarize, both yield spreads and general risk aversion have obvious predictive ability over the movement of media pessimism, but their effects are less influential than the number of news stories. Although the correlation between the detrended sentiment score (*NEG_DTR*) and the detrended number of news stories is only 0.15, the increase of media pessimism is strongly associated with more news stories being reported. If media pessimism (the number of news stories) is above its latest five-day average level, the likelihood that the number of news stories (media pessimism) moves upwards would increase by almost 7 times. The results reveal that sentiment are strongly associated with the concentration/volume of news and information.

V. Summaries and Conclusions

Although there is an increasing body of literature studying the role of textual sentiment in equity markets, no previous research has been conducted for debt markets. The association between textual sentiment and yield spreads has yet to be linked. This paper examines to what extent the well-documented effect of textual sentiment on equities also applies to the sovereign debt markets. In comparison to the literature exploring the determinants of sovereign yield spreads, this research focuses on examining the impact of media pessimism and its interaction

with the concentration/volume of news stories on yield spreads, instead of testing different measures of credit risk, liquidity and general risk aversion. Greece, Ireland, Italy, Portugal and Spain are selected to form the sample. The European sovereign debt crisis, from January 2009 to December 2012, is selected as the data range, due to the surge in news stories regarding each country's credit situations during this period. Media pessimism is extracted and quantified using the dictionary-based approach from a total of 18,610 country-specific news stories, and two measures of sentiment are constructed.

The study finds that although media pessimism does not directly determine sovereign bond yield spreads, its interaction with the concentration/volume of news is significantly associated with yield spreads. Linear regressions by *GMM* show that if both media pessimism and the number of news stories/words increase and other factors remain unchanged, yield spreads would move upwards, causing prices to fall. Logistic regressions show that the likelihood that yield spreads widen relative to the previous day increases by 21 percent if media pessimism is above its five-day moving average. The increase of the concentration of news (the number of news stories) alone contributes to the likelihood of yield spread increase by 19 percent, but the effect vanishes if controlling for the upward movement of media pessimism. The simultaneous increase of media pessimism and the number of news stories relative to their latest five-day average level also adds to the likelihood of yield spreads widening.

Greater changes in yield spreads and/or general risk aversion have an obvious predictive ability on the rise of media pessimism relative to its five-day moving average, but their effects are less influential than the number of news stories. The heightening of media pessimism is strongly associated with more news stories being reported, suggesting that "no news is good news": more news stories during the Eurozone crisis are often associated with a higher level of media pessimism. This in turn drives up sovereign bond yield spreads, as media pessimism and the concentration/volume of news collectively determine yield spreads in addition to the traditionally recognized variables.

A limitation of this study is that the same methodology is nearly impossible to be applied to a non-crisis period, because there are usually very few relevant news stories during such times, often fewer than 10 in one year.[13] Therefore, the implications of this study are most relevant

---

13. I checked the number of news stories for each country over 2002–2007.

to the countries of GIIPS during the Eurozone crisis. Overall, this research confirms that textual sentiment could convey additional pricing-relevant information for sovereign bonds over the proxies of sovereign credit risk, liquidity risk and general risk aversion. Particularly, media pessimism, coupled with the concentration/volume of news stories, plays an important role in determining sovereign bond yield spreads. This research has implications for both the rational and the behavioral paradigms of asset pricing, and suggests a novel perspective of studying the determinants of bond yield spreads or interest rates in the future.

Accepted by: *Prof. P. Theodossiou, Editor-in-Chief, August* 2014

References

Afonso, A.; Furceri, D.; and Gomes, P. 2012. Sovereign credit ratings and financial markets linkages: Application to European data. *Journal of International Money and Finance* 31(3): 606–638.

Ahmad, K.; Hutson, E.; Kearney, C.; and Liu, S. 2014. Media pessimism and stock returns: A time-varying analysis at the firm level. Available at http://eprints.soton.ac.uk/365557/.

Attinasi, M-G.; Checherita, C.; and Nickel, C. 2009. What explains the surge in euro area sovereign spreads during the financial crisis of 2007–09?. ECB Working Paper, No.1131.

Barbosa, L., and Costa, S. 2010. Determinants of sovereign bond yield spreads in the euro area in the context of the economic and financial crisis. Banco de Portugal, Working Paper, No.22.

Barrios, S.; Iversen, P.; Lewandowska, M.; and Setzer, R. 2009. Determinants of intra-euro area government bond spreads during the financial crisis. European Commission Economic Papers No. 388.

Beber, A.; Brandt, M. W.; and Kavajecz, K. A. 2009. Flight-to-quality or flight-to-liquidity? Evidence from the Euro-area bond market. *Review of Financial Studies* 22: 925–957.

Bernoth, K.; von Hagen, J.; and Schuknecht, L. 2006. Sovereign risk premiums in the European government bond market. Governance and the Efficiency of economic systems (GESY), Discussion Paper, No. 151.

Breitung, J. 2000. The local power of some unit root tests for panel data. *Advances in Econometrics* 15: 161–177.

Codogno, L.; Favero, C.; and Missale, A. 2003, Yield spreads on EMU government bonds. *Economic Policy* 18(37): 503–532.

Ejsing, J. W. and Sihvonen, J. 2009. Liquidity premia in German government bonds. ECB Working Paper, No.1081.

Engelberg, J. 2008. Costly information processing: evidence from information announcements. AFA 2009 San Francisco Meetings Paper. Available at SSRN: http://ssrn.com/abstract=1107998.

Faini, R. 2006. Fiscal policy and interest rates in Europe. *Economic Policy* 21(47): 443–489.

Favero, C.; Pagano, M.; and vonThadden, E.-L. 2010. How does liquidity affect government bond yields? *Journal of Financial and Quantitative Analysis* 45(1): 107–134.

Ferguson, N. J.; Guo, J.; Lam H. Y. T.; and Philip, D. 2014. Media content and stock returns: The predictive power of press. Available at SSRN: http://ssrn.com/abstract=2111352.

Garcia, D. 2013. Sentiment during recessions. *Journal of Finance* 68(3): 1267–1300.

Gerlach, S.; Alexander, S.; and Guntram, G. B. 2010. Banking and sovereign risk in the Euro area. Deutsche Bank, Discussion Paper, No.09/2010.

Geyer, A.; Kossmeier, S.; and Pichler, S. 2004. Measuring systematic risk in EMU government yield spreads. *Review of Finance* 8: 171–197.

Gomez-Puig, M. 2006. The Impact of monetary union on EU-15 sovereign debt yield spreads. University of Barcelona, Working Paper in Economics 147.

Harris, R. D. F., and Tzavalis, E. 1999. Inference for unit roots in dynamic panels where the time dimension is fixed. *Journal of Econometrics* 91: 201–226.

Haugh, D.; Ollivaud, P.; and Turner, D. 2009. What drives sovereign risk premiums? An analysis of recent evidence from the euro area. OECD Economics Department, Working Paper No. 718.

Hayashi, F. 2000. *Econometrics*. New Jersey: Princeton University Press.

Heppke-Falk, K., and Hófner, F. 2004. Expected budget deficits and interest rate swap spreads - evidence for France, Germany and Italy. Deutsche Bundesbank Discussion Paper No. 40/2004.

Im, K. S.; Pesaran, M. H.; and Shin, Y. 2003. Testing for unit roots in heterogeneous panels. *Journal of Econometrics* 115: 53–74.

Jankowitsch, R.; Mosenbacher, H.; and Pichler, S. 2006. Measuring the liquidity impact on EMU government bond prices. *European Journal of Finance* 12(2): 153–169.

Jegadeesh, N., and Wu, A. D. 2013. Word power: A new approach for content analysis. *Journal of Financial Economics* 110(3): 712–729.

Kearney, C., and Liu, S. 2014. Textual sentiment in finance: A survey of methods and models. *International Review of Financial Analysis* 33: 171–185.

Levin, A.; Lin, C. F.; and Chu, C. 2002. Unit root test in panel data: Asymptotic and finite sample properties. *Journal of Econometrics* 108: 1–25.

Li, F., 2010. The information content of forward-looking statements in corporate filings - A Naive Bayesian machine learning algorithm approach. *Journal of Accounting Research* 48: 1049–1102.

Loughran, T., and McDonald, B. 2011a. When is a liability not a liability? *Journal of Finance* 66, 35–65.

Loughran, T., and McDonald, B. 2011b. Barron's red flags: Do they actually work? *Journal of Behavioral Finance* 12 (2): 90–97.

Loughran, T., and McDonald, B. 2013. IPO first-day returns, offer price revisions, volatility, and Form S-1 language. *Journal of Financial Economics* 109(2): 307–326.

Manganelli, S., and Wolswijk, G. 2009. What drives spreads in the Euro area government bond market? *Economic Policy* 24(58): 191–240.

Mody, A. 2009. From Bear Stearns to Anglo Irish: How Eurozone sovereign spreads related to financial sector vulnerability. IMF Working Paper, No.09/108.

Pagano, M., and VonThadden, E.-L. 2004. The European bond markets under EMU. *Oxford Review of Economic Policy* 20(4): 531–554.

Pesaran, M. H. 2004. General diagnostic tests for cross section dependence in panels. Cambridge Working Papers in Economics, 0435, University of Cambridge.

Schuknecht, L.; vonHagen, J.; and Wolswijk, G. 2011. Government bond risk premiums in the EU revisited the impact of the financial crisis. *European Journal of Political Economy* 27(1): 36–43.

Schwarz, K. 2014. Mind the gap: disentangling credit and liquidity in risk spreads. Working Paper, Department of Finance, The Wharton School, University of Pennsylvania.

Sgherri, S., and Zoli, E. 2009. Euro area sovereign risk during the crisis. IMF Working Paper, No.09/222.

Sinha, N. R. 2010. Underreaction to news in the US stock market. Available at SSRN: http://ssrn.com/abstract=1572614.

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

Tetlock, P. C.; Saar-Tsechansky, M.; and Macskassy, S. 2008. More than words: Quantifying language to measure firms' fundamentals. *Journal of Finance* 63: 1437–1467.