European Commission

JRC TECHNICAL REPORTS

Anatomy of a Sovereign Debt Crisis: CDS Spreads and Real-Time Macroeconomic Data

Lucia Alessi

Pierluigi Balduzzi

Roberto Savona

2019

JRC Working Papers in Economics and Finance, 2019/3

This publication is a Technical report by the Joint Research Centre (JRC), the European Commission's science and knowledge service. It aims to provide evidence-based scientific support to the European policymaking process. The scientific output expressed does not imply a policy position of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use that might be made of this publication.

EU Science Hub
https://ec.europa.eu/jrc

JRC115389

PDF ISBN 978-92-76-00120-1 ISSN 2467-2203 doi:10.2760/761074

Luxembourg: Publications Office of the European Union, 2019

How to cite: Alessi, L., Balduzzi, P. and Savona, R., *Anatomy of a Sovereign Debt Crisis: CDS Spreads and Real-Time Macroeconomic Data*, *JRC Working Papers in Economics and Finance*, Publications Office of the European Union, Luxembourg, 2019, ISBN 978-92-76-00120-1, doi:10.2760/761074, JRC115389

Anatomy of a Sovereign Debt Crisis: CDS Spreads and Real-Time Macroeconomic Data

LUCIA ALESSI[*] PIERLUIGI BALDUZZI[†] ROBERTO SAVONA[‡ §]

This draft: February 2019

[*]European Commission - Joint Research Centre, Italy.

[†]Boston College. Corresponding author: Pierluigi Balduzzi, Carroll School of Management, Boston College, 140 Commonwealth Avenue, Chestnut Hill, Massachusetts, 02467; Tel: (617) 552-3976; Fax: (617) 552-0431; email:balduzzp@bc.edu.

[‡]Department of Economics and Management, University of Brescia, Italy.

[§]The views in this paper are those of the authors, and do not necessarily reflect those of the European Commission. We thank seminar participants at the 2018 Annual Conference of the Community of Practice in Financial Research (CoPFiR) for useful comments.

Anatomy of a Sovereign Debt Crisis: CDS Spreads and Real-Time Macroeconomic Data

ABSTRACT

We construct a unique and comprehensive data set of 19 real-time daily macroeconomic indicators for 11 Eurozone countries, for the 5/11/2009–4/25/2013 period. We use this new data set to characterize the time-varying dependence of the cross-section of sovereign credit default swap (CDS) spreads on country-specific macro indicators. We employ daily Fama-MacBeth type cross-sectional regressions to produce time-series of macro-sensitivities, which are then used to identify risk regimes and forecast future equity market volatility. We document pronounced time-variation in the macro-sensitivities, consistent with the notion that market participants focused on very different macro indicators at the different times of the crisis. Second, we identify three distinct crisis risk regimes, based on the general level of CDS spreads, the macro-sensitivities, and the GIPSI connotation. Third, we document the predictive power of the macro-sensitivities for future option-implied equity market volatility, consistent with the notion that expected future risk aversion is an important driver of how CDS spreads impound macro information.

JEL G12

Introduction

The Eurozone sovereign debt crisis clearly exhibits *three* fundamental turning points. The first turning point is in October 2009, when Greek Finance Minister George Papaconstantinou discloses that the true budget deficit for 2009 was 12.5% of GDP, more than twice the previously announced figure. From this point in time on, after a decade of disconnect, different macroeconomic fundamentals begin to translate into very different assessments of sovereign default probabilities and recovery rates.

The second turning point is in April 2010, when Greece activates a 45 billion Euros EU-IMF bailout, and S&P downgrades Greek debt to junk status.[1] The yields on Greek long-term debt jump immediately in response to news about a future potential default, soon followed by spreads of sovereign bonds of Eurozone countries facing similar fiscal troubles (Ireland, Italy, Portugal, Spain).

The third turning point is at the time of Mario Draghi's "whatever it takes" pledge, made on July 26, 2012, and the subsequent announcement of the Outright Monetary Transactions (OMT) program (August 2012), through which the European Central Bank will make purchases (outright transactions) in the secondary bond markets of Eurozone member states. As a result, spreads on the sovereign bonds of more vulnerable Eurozone countries start trending down and the sovereign debt crisis begins to subside.

What we know from the extensive literature on the Eurozone sovereign debt crisis is that most of the increase in the price of sovereign risk was due to a deterioration in countries' fundamentals coupled with fundamentals' contagion (Beirne and Fratzscher, 2013) and feedback

[1] On April 25, 2010, namely 2 days before the S&P's downgrading, the Financial Times wrote "This is going to be the most important week in the 11-year history of Europe's monetary union. By the end of it we will know whether the Greek fiscal crisis can be contained or whether it will metastasize to other parts of the Eurozone."

loops between sovereign and domestic bank risks (Acharya et al., 2018; Bolton and Olivier, 2011). We also know that the OMT program was successful in lowering spreads of sovereign bonds issued by more exposed European countries (Krishnamurthy et al., 2018). However, it has proven so far very challenging to explain the bulk of the high-frequency variation in sovereign spreads or Credit Default Swap (CDS) premia by means of fundamentals.

We show that what matters is ultimately not the level of macro fundamentals, which indeed evolve very smoothly. Rather, it is the *importance* that markets attach to different fundamentals. We show that this is very low at the outset of the crisis: in this phase, markets "panic," with some countries paying for their mere belonging to a set of vulnerable countries. This explains why an abrupt, substantial repricing of risks may take place against unchanged, or only marginally deteriorated, economic conditions. On the contrary, at the height of the crisis, attention to economic fundamentals becomes extreme. Past the peak, we get back to a virtual disconnect between market developments and macro fundamentals, with spreads mostly driven by sentiment. Our results suggest that monetary policy intervention can reduce spreads across the board in a time of crisis by providing the proverbial "tide that lifts all boats" (e.g., De Grauwe and Ji, 2013). However, individual countries can ultimately improve their funding costs only by intervening on their own macro fundamentals.

In order to study the "anatomy" of the Eurozone sovereign debt crisis, we carry out a detailed analysis of the relation between the pricing of sovereign risk and a comprehensive set of macroeconomic fundamentals, as they are disclosed to the public in *real-time*. We are the first to perform this type of exercise. In fact, most of the empirical work on the topic has used *revised* macroeconomic data.[2] The very few papers employing real-time data

[2] A "revised vs. real-time bias" has been brought to attention, for example, by Ghysels et al. (2018) in the context of bond return predictability, showing that real-time macroeconomic data have a much lower predictive power than final, revised data.

have focused on macro-news announcements (e.g., Kim et al., 2015; Beetsma et al., 2013), relating changes in sovereign CDS spreads to the "distance" between released and expected quantities. However, this approach suffers from two main limitations. First, given that macro announcements are not synchronized across countries, it is not possible to implement a pure cross-sectional analysis of the responses of CDS spreads to news. Hence, if there is time variation in these responses, the time variation needs to be modeled explicitly. Second, data on consensus forecasts are available for only few macroeconomic variables and countries. Hence, any feasible empirical analysis can only be limited in scope.

In our analysis, on the other hand, we are able to use data for *all* trading days and for a *rich* cross-section of macro indicators. This approach allows us to accommodate time-variation in the relation between spreads and fundamentals, without having to model such relation explicitly. This is important, as we document that the variation in macro-sensitivities—the coefficients relating the pricing of default risk to macro fundamentals—is substantial, with the set of relevant variables itself varying drastically from one "regime" to another. Moreover, we are able to include in the analysis a broad set of indicators, for which consensus forecasts are not available. This is also important because, at times, we find that less-known macro indicators play an important role in explaining the cross-section of CDS spreads.

Specifically, we use the ECB e-archives to construct a unique real-time, daily-frequency data set on 19 macroeconomic fundamentals, for 11 Eurozone countries over the period from 5/11/2009 to 4/25/2013. We then relate the cross-section of sovereign CDS spreads of different maturity to the macro fundamentals, employing data for all of the trading days in the sample.

While the literature agrees that during the European crisis sovereign bond prices and

CDS spreads exhibited excessive sensitivity to macroeconomic indicators (Aizenman et al., 2018; Bernoth and Erdogan, 2013), a clear understanding on how market prices incorporated information on country-specific fundamentals over time is still missing. This is our *first* contribution. Based on our real-time macroeconomic data set, we characterize and interpret the cross-section of sovereign CDS spreads. Analytically, we implement a Fama-MacBeth procedure in which, for each day, the sovereign CDS spreads for the 11 Eurozone countries at 3, 5, 7, and 10 years are regressed on the 19 country-specific macro fundamentals, controlling for the level and volatility of an indicator of banking risk, and for being part of the GIPSI group of countries. These covariates are employed individually, to capture "level" effects, and then interacted with the maturities of the CDS contracts, to capture "slope" effects. Given the large initial set of covariates, we implement a LASSO-type approach to the regression analysis (Tibshirani, 1996). In doing this, first, we reduce the dimensionality of the space of covariates, making the estimation procedure feasible; second, we implement a day-by-day variable selection procedure, discarding those covariates that make no contribution in explaining the cross-section of CDSs. The daily LASSO-type cross-sectional coefficients are then stacked together, producing time-series of macro-sensitivities whose values tell us which variables are important at any given time.

The daily time series of the LASSO-type coefficients are then used to detect homogeneous clusters of observations. Such groups of observations, i.e., "regimes", are identified through the "medoid" clustering procedure introduced in Kaufman and Rousseeuw (1990). This is our *second* contribution. Our approach differs from existing papers which instead identify non-crisis vs. crisis regimes based on the behavior of CDS/bond spreads alone (e.g., Blommestein et al., 2016; Delatte et al., 2017).

Our *third* contribution has to do with the explanatory power of the cross-sectional regression coefficients vis-à-vis the future option-implied volatility of the European equity market. In principle, our LASSO-type coefficients combine information on the cross-section of CDS spreads with information on macro fundamentals, in a way that reflects the risk attitudes of market participants.[3] To explore this conjecture, we focus on the Euro Stoxx 50 volatility index (VSTOXX). This type of indicator has taken a central role in the debate on the predictability of future economic activity, monetary policy stance, and financial instability (e.g., Bekaert and Hoerova, 2014; Bekaert et al., 2013). High expected financial market volatility is a signal of increased risk of adverse future economic conditions and, hence, of a potential impending crisis (Danielsson et al., 2018). In doing this, we complement and extend the analysis of Beber et al. (2015), who document significant explanatory power of real-time macroeconomic indicators for future realizations of the VIX index. Our approach differs from theirs, however, as our macro-sensitivities combine real-time macro data with financial market indicators (the CDS spreads) to explain future realizations of an equity volatility index. Indeed, we show that our approach is superior to that of using either the macro indicators or the CDS spreads alone.

In terms of empirical results, our strategy delivers very good cross-sectional fit, with an average cross-sectional R-square of 0.9850. For comparison, the model in Aizenman et al. (2018) explains the cross-section of CDS spreads with R-squares between 0.7 and 0.8, during the pre-crisis period, and between 0.45 and 0.60, during the crisis.[4]

[3] The explanatory power of cross-country macro-sensitivities is motivated by the systematic nature of sovereign risk and the forward-looking properties of sovereign CDSs. Theoretically, this view is consistent with the literature exploring how investors react to uncertainty about the state of the economy (Veronesi, 1999; Veronesi, 2014) and to expectations about future monetary policy (Faust et al., 2007).

[4] The papers by Nosbusch (2010), Longstaff et al. (2011) and Dieckmann and Plank (2011) share a similarly disappointing fit. A notable exception is the recent paper by Augustin (2018), who models the time-varying dynamics and cross-country heterogeneity of the term structure of sovereign CDS, showing that global (local) shocks are the primary source of time variation for spreads when the term structure is upward (downward)-sloping.

Second, as discussed above, a clear link between sovereign spreads and macro fundamentals is still elusive (Beber et al., 2014), because of the "revised vs. real-time bias" as well as the abundance of variables that can be used as potential risk factors. Our Fama-MacBeth procedure selects macroeconomic indicators based on their usefulness in pricing the cross section of sovereign CDS spreads, selecting only the variables with non-zero coefficients, i.e., those covariates that matter most on any given day. The stacked LASSO-based coefficients that we estimate over the entire time interval exhibit significant time variation.

An in-depth analysis of the behavior exhibited by the coefficients over time identifies three distinct crisis regimes. While our results are generally consistent with those of Bernoth and Erdogan (2013), we document pronounced volatility in the macro-sensitivities that is strongly at odds with their conclusions that the relation between CDS spreads and fundamentals is "changing gradually over time, rather than having a discrete break-point between regimes." By using real-time data, we find exactly the opposite pattern of variation, with significant jumps in the macro-sensitivities. We set the pre-crisis regime to be the 5/11/2009 –3/31/2010 period, characterized by moderate macro-sensitivities, with the fiscal deficit and the GIPSI dummy being the major determinants of sovereign risk, followed by our banking risk indicator. Starting from April 2010, the time-series of the macro-sensitivities lead us to identify three main crisis regimes: (i) the *first* regime is mainly characterized by the explanatory power of the GIPSI dummy and the loan-to-government indicator, covering the first phase of the crisis (April 2010–July 2011) and the Cyprus bank bailout (March–April 2013); (ii) during the *second* regime, June–August 2011 and September 2012–mid March 2013, imports and changes in inventories over GDP were the most influential variables, given their relevance for future GDP growth; (iii) the *third* regime (July 2011–August 2012), corresponds to the phase of highest risk, when the cross sections of sovereign CDS were mostly driven by GDP growth and employment. Interestingly, it is during this last regime that CDS spreads reflect macro fundamentals the most. In other words, it is precisely when the overall perception of sovereign default risk is greatest that country-specific macro information impacts prices the most, consistent with the notion that these are the times when the returns from information production are also the highest.[5]

Third, we document substantial out-of-sample explanatory power of the LASSO-type, Fama-Macbeth coefficients for future option-implied equity market volatility, as it is captured by the VSTOXX volatility index. Given the large number of macro-sensitivities, we employ the LASSO algorithm a second time to identify the most robust specification. Moreover, we perform the first-step cross-sectional regressions on the principal components of the original macro indicators, following Beber et al. (2015).[6] While macro indicators play a only a minor role in explaining the slope of the term structure of CDS spreads, the coefficients capturing slope effects have a substantial role in this out-of-sample predictive exercise.

The remainder of this paper is organized as follows. Section I describes our real-time, macro-indicator data set. Section II derives the empirical specification used in the cross-sectional regression analysis. Section III presents the results of the Fama-Macbeth-style cross-sectional regressions. Section IV discusses the methodology and evidence on the different regimes. Section V presents the out-of-sample predictive exercise. Section VI concludes.

[5]This result stands in partial contrast with the evidence in Pasquariello (2014), who shows that market dislocations—i.e., conditions where the markets cease to price assets correctly on an absolute and relative basis—are more prevalent at the time of crisis.

[6]The principal components are constructed separately for sub-groups of macro indicators: (1) Labour market; (2) Prices and costs; (3) Money and credit; (4) Output; (5) Banking.

I Macroeconomic data in real-time

The dependent variable is the premium on sovereign CDSs at the 3, 5, 7 and 10 year maturity, over the 5/11/2009–4/25/2013 period, for a set of 11 Eurozone countries (see Figure 1a). CDSs are financial contracts aimed at protecting the buyer of a bond from the default risk of the issuer. For example, the seller of a CDS on a particular sovereign agrees to compensate the buyer in case of a loss emanating from a credit event in the reference sovereign. Notice that, unless an explicit ban is foreseen in the regulation of a specific jurisdiction, CDSs can also be "naked," i.e., bought by investors who do not own the underlying bond, and are hence simply 'betting' against the issuer. Being sovereign CDSs similar to insurance contracts protecting against a default of the counterparty, their price is a measure of the perceived default risk associated with a particular sovereign entity. Our analysis covers the following Eurozone countries: Austria, Belgium, Cyprus, Germany, Finland, France, Ireland, Italy, the Netherlands, Spain, and Portugal. Figure 1a shows developments in the sovereign CDS spreads for these countries during the sample. As in other studies using sovereign CDS data (e.g., de Santis, 2015), Greece is not covered as data on Greek CDSs is not available after September 2011.

We aim at explaining the dynamics in sovereign CDS prices based on a novel real-time data set comprising country-specific macroeconomic fundamentals and macro-financial indicators. Several studies have tried to explain sovereign debt crises by linking some measure of sovereign stress to macro fundamentals. While earlier papers in this literature have used standard, low-frequency data sets (e.g., Manasse and Roubini, 2009), only recently higher-frequency data have been used for this purpose (Beber et al., 2014). More broadly, there is a growing literature showing the importance of using real-time macroeconomic data for forecasting not only macroeconomic variables themselves (e.g., Giannone et al., 2005, and, more recently, Beber et al., 2014), but also financial variables (see Ghysels et al., 2018). Real-time data sets are also needed to construct credible early warning models for financial crises (e.g., Alessi and Detken, 2011 and Alessi and Detken, 2014), as these models are intended to be used by policymakers based on the information set that is available to them at each point in time.

For these reasons, real-time data sets are becoming increasingly popular also beyond the macro-econometric literature. However, to our knowledge, there are no existing real-time data sets covering a large number of countries and a large number of macroeconomic variables at high frequency. For the Euro area, in particular, the most well-known real-time database is the one maintained by the Euro Area Business Cycle Network (EABCN, see Giannone et al., 2012), which also started publishing real-time data for some individual European countries. This data set, however, reports macroeconomic indicators at a monthly frequency, and hence it is not suitable to pin down the high-frequency impact of data releases on the financial markets.

Against this background, we construct a novel real-time data set at a daily frequency, covering 11 Eurozone countries and including 19 macroeconomic and macro financial indicators reported in Table 1. Our real-time data set covers the following indicators:

- Labour market indicators: unemployment and employment rates.
- Prices and costs: inflation rate, industrial producer prices (% change), hourly labour cost (% change).
- Money, credit and debt: growth of M3, loans to private sector, loans to government, total credit to private sector, and total credit to government, as well as public sector deficit over GDP.

- Output: real GDP, consumption, government consumption, investment, exports, imports and industrial production (all rates of growth) and changes in inventories over nominal GDP.

Descriptive statistics are summarized in Tables 2 and 3. Figures 1b–1e show developments in selected macro fundamentals during the crisis. Clearly, these variables change much more smoothly than the CDSs spreads plotted in Figure 1a. The panel of time series is balanced, with all of the above variables being available for each of the considered countries over the whole time-span. In contrast, the data set used by Beber et al. (2014) includes a comparable number of indicators for Germany, but many less for the other countries.

Our real-time data set is mostly based on European Central Bank e-archives. These e-archives contain historical records of the information supplied to the public by the ECB. In constructing the data set, we have taken into account the various lags with which new data are released by the ECB, compared to the moment they are released by national statistical institutes and national central banks. The latter date is the release date that matters, as it corresponds to when new information reaches the markets for the first time. For this purpose, official release dates have been retrieved or double-checked using information from Bloomberg, Money Market Services (MMS), as well as information from national central banks and statistics offices.

The structure of the data set differs from standard, lower-frequency, real-time data sets, as it does not exhibit "vintages." This is due to the different frequencies at which the variables of interest are released (up to quarterly) and of the data set itself (daily). In fact, the data set is structured as a standard panel of mixed-frequency time series. The difference with respect to a standard data set is that, at each date, each of the macro variables listed above takes the *latest released value*, instead of the value for the reference period, which is not known in real-time. For example, current Eurozone GDP growth will only be available 30 days after the end of this quarter (T+30), in the form of a preliminary flash estimate, which will be revised 15 days later (T+45), while the second GDP release will be published 60 days after the end of this quarter (T+60).[7] Some countries nowadays publish preliminary GDP flash estimates, while some only publish GDP figures 60-70 days after the end of the reference quarter.[8] Moreover, data for GDP components may be released together with a flash estimate or only with the second GDP release, depending on the country. Even monetary and credit aggregates, which are released in a more timely manner as compared to macroeconomic statistics, are published in the month following the reference month.

Given publication lags, market participants never really know the current state of the economy. They base their decisions on a continuous flow of information, where data on various macroeconomic and macro financial indicators are released with a different timeliness, and revised afterwards. The real-time data set that we develop reflects the information set available to market participants at each point in time, based on which they form expectations. In this respect, our data set is similar to those used in "news" studies, such as Balduzzi et al. (2001), Ehrmann and Fratzscher (2005) and, more recently, Beber et al. (2014).

Finally, we complement our macroeconomic and macro financial real-time data set with a market-based indicator, namely a proxy for country-specific banking risk. As documented in the existing literature, the doom-loop between sovereign and bank credit risk was indeed the hallmark of the 2009-2012 sovereign debt crisis in the periphery of the euro area (Brunnermeier et al., 2016). Computationally, in order to obtain a market-based, daily basis proxy for banking risk, we used the country-specific banking equity index. More specifically, since

[7]Eurostat improved the timeliness of the Eurozone GDP flash estimate including a preliminary release at T+30 days in 2016.

[8]Focusing on the countries covered in this study, no flash estimate is provided for Ireland.

sovereign CDSs and banking equity indices are strongly correlated, we orthogonalized the daily banking equity returns by regressing on the contemporaneous daily change in sovereign CDS premia. Hence, the cumulative sums of residuals were used to construct our measure of banking risk, which, by construction, is uncorrelated with the variations in CDS premia, thereby reflecting the health status of the banking systems as measured by the market. While the banking risk indicator, computed as a daily cumulative sum of residuals, reflects the "level" of health of the banking systems, we also computed the 20-day rolling-window realized volatility of the residuals, which gives us a measure of the uncertainty associated with banks' health.

II Sovereign risk pricing

We start from a general framework for the pricing of CDSs in an arbitrage-free setting. This general framework motivates the specification used in the empirical analysis and provides an economic justification for why the dependence of CDS premia on macro-fundamentals may be time-varying and why the coefficients of cross-sectional regression of CDS spreads on macro-fundamentals may be informative as to the expected future volatility of the equity market.

II.A The general framework

For given country n, let s_{nt} denote the one-period CDS spread, and let r_{nt} and π_{nt} denote the recovery rate and default probability, respectively. Consider the payoff of a one-period CDS with a \$1 face value, and assume, for simplicity, that default can only take place at time 1:

$$c_{nt} = \max\{1 - r_{nt}, 0\}. \tag{1}$$

Under standard assumptions, absence of arbitrage implies that there exists a stochastic discount factor m_{t+1} such that:[9]

$$\frac{s_{nt}}{1+i_t} = E_t(m_{t+1}c_{nt+1}), \tag{2}$$

where i_t denotes the rate of interest, and where $E_t(m_{t+1}c_{nt+1})$ can be broken into the CDS risk-neutral valuation and the CDS risk premium, namely:

$$E_t(m_{t+1}c_{n,t+1}) = \underbrace{\frac{E_t(c_{nt+1})}{1+i_t}}_{\text{risk-neutral valuation}} + \underbrace{\text{cov}_t(m_{t+1}, c_{nt+1})}_{\text{risk premium}}. \tag{3}$$

We have:

$$E_t(c_{nt+1}) = \pi_{nt} \times E_t\{1 - r_{nt+1}|r_{nt+1} \leq 1\} \tag{4}$$

and:

$$\text{cov}_t(m_{t+1}, c_{nt+1}) = \underbrace{\frac{\text{cov}_t(m_{t+1}, c_{nt+1})}{\text{var}_t(m_{t+1})}}_{\text{risk}_{=\beta_{nt}}} \times \underbrace{\text{var}_t(m_{t+1})}_{\text{market price of risk}_{=\lambda_t}}. \tag{5}$$

II.B Towards an empirical specification

We assume that the conditional expectation $E_t(m_{t+1}c_{n,t+1})$ is a function of market information available at time t. Specifically, we assume:

$$\begin{aligned}
\pi_{nt} &= \pi(x_{nt};\theta) & (6)\\
E_t\{1 - r_{nt+1}|r_{nt+1} \leq 1\} &= \gamma(x_{nt};\theta) & (7)\\
\frac{\text{cov}_t(m_{t+1}, c_{nt+1})}{\text{var}_t(m_{t+1})} &= \beta(x_{nt};\theta) & (8)\\
\text{var}_t(m_{t+1}) &= \lambda(y_t;\theta), & (9)
\end{aligned}$$

[9] For a generalization of the CDS pricing equation to a multi-period contract, see, for example, Augustin (2018).

where x_{nt} denotes a $(K+1) \times 1$ vector of country-specific macro variables (including a constant, the last element of the x_{nt} vector), y_t is a vector of systematic variables, and θ is a vector of coefficients. Therefore, we can express the one-period CDS spread as:

$$s_{nt} = \pi(x_{nt};\theta)\gamma(x_{nt};\theta) + \beta(x_{nt};\theta)\lambda(y_t;\theta)(1+i_t). \tag{10}$$

We approximate the general specification above with the linear specification:

$$s_{nt} = \delta_t^\top x_{nt}. \tag{11}$$

Note that even though we are assuming time-invariant relations linking the fundamentals x_{nt} and y_t to the determinants of CDS spreads, the time-variation in the market price of risk $\lambda(y_t;\theta)$ leads to time-varying relations linking s_{nt} to x_{nt}.

The specification (11) can be generalized to a τ-maturity CDS with spread $s_{n\tau t}$ as:

$$s_{n\tau t} = \delta_t(\tau)^\top x_{nt}, \tag{12}$$

where we assume:

$$\delta_t(\tau) \quad = \quad \delta_{1t} + \delta_{2t}\tau. \tag{13}$$

In this way, we provide a simple semi-non-parametric CDS term structure model, where the coefficients δ_{1t} capture a "level" effect and the coefficient δ_{2t} a "slope" effect. In the next section, we introduce a variable-selection approach aimed at selecting the possibly different factors that matter at different times in explaining the cross-section of CDS spreads.

II.C A Fama-MacBeth specification with variable selection

The empirical specification used in our analysis results directly from the expression derived for the τ-maturity CDS spread. Based on equations (12)–(13), we have:

$$s_{nmt} = \delta_{1t}^\top x_{nt} + \delta_{2t}^\top (x_{nt} \times \tau_m) + \epsilon_{nmt}. \tag{14}$$

Note that the dimension of each cross-section at time t is given by NM, where N is the number of countries and M is the number of CDS maturities. Therefore, since in our sample we have 11 countries and 4 maturities, we then have a total of $11 \times 4 = 44$ cross-sectional observations at every time t. This relatively small cross-section complicates the estimation process as the $K+1$ country-specific macro-variables in x_{nt} enter in equation (14) both alone and then interacted with τ.

We address the issue above by implementing a LASSO-type penalty regression (Tibshirani, 1996). This approach reduces the dimensionality of the covariate space, allowing us to estimate the cross-sectional regressions even when the initial set of regressors exceeds the number of cross-sectional observations. Moreover, the LASSO translates into a time-varying variable selection algorithm, as we discard those covariates that make no contribution in explaining the cross-section of CDSs at each time t. Relative to standard approaches to specification search, such as step-wise regression, the LASSO has the advantage that it optimizes the *out-of-sample* performance of the regression model, in the spirit of "calibrating" a pricing model on one set of CDS contracts and then applying that model to price another set of contracts.

The econometric procedure contemplates as many cross-sectional regressions as the number of time-series observations. In so doing, the procedure is analogous to the first-step of the Fama-MachBeth procedure. Indeed, after running T LASSO-type cross-sectional regressions, we focus on the time-t's coefficients by stacking them together, and hence obtaining the time-series of the sensitivities towards the macro-variables (alone and interacted with τ).

II.D LASSO: motivation and implementation

The motivation for using the LASSO algorithm is twofold.[10] First, as highlighted above, the initial number of covariates exceeds the number of cross-sectional observations. A natural solution to the problem is to apply the so-called "bet on sparsity principle", namely, to assume that the underlying true model contains only relatively few nonzero parameters. This principle is implemented via the LASSO algorithm, which constrains or regularizes the estimation process, leading to nonzero coefficients for a subset of few variables, and forcing the remaining coefficients to zero.

The second and related reason behind the use of the LASSO is the "philosophy" underlying this approach to regression analysis. It is well known that regularization seeks a compromise between interpretability and flexibility, by excluding covariates whose coefficients are close to zero.[11] In doing this, redundant and noisy information (covariates) are discarded, as they are not useful in characterizing the response variable.

Sparsity is a pervasive concept in our "data-driven" era, where ever-increasing amounts of data lead to natural questions, such as: "why go through so much effort to acquire all the data when most of what we get will be thrown away?" ((Donoho, 2006)). Questions to which common sense suggests to measure only the effects present in the portion of the data that will not end up being thrown away. This is exactly the perspective that we embrace, by aiming to characterize the cross-sections of sovereign CDS spreads while selecting a small number of covariates that explain most of the variation in the response. Agnostically, we

[10] There is increased usage of plain and advanced LASSO methods in economics (for e.g., (De Mol et al., 2008); (Song and Bickel, 2011); (Fan et al., 2011); (Kock, 2016); (Li and Chen, 2014); (Gefang, 2014); (Kock and Callot, 2015)), and finance ((Brodie et al., 2009); (Fan et al., 2012); (DeMiguel et al., 2009); (Scherer, 2015); (Bruder et al., 2013); (Freyberger et al., 2017); (Feng et al., 2019); Chinco et al. (2019)).

[11] This is the so-called bias-variance trade-off: the choice of a more exhaustive set of covariates reduces the estimation bias but increases the variance of the estimates.

assume that although all of the available macro variables are in principle equally important, we expect that in any cross-section only a small number really matters and that over time the sub-set of relevant covariates can change based on changing investors' views.

Analytically, the LASSO algorithm estimates regression parameters by imposing a constraint on the sum of the absolute values of the slope coefficients, namely on the total ℓ_1 norm of the parameter vector (excluding the constant). In our context, cross-sectional regressions (equation (14)) are estimated by solving the following problem:

$$\min_{\delta_{1t},\, \delta_{2t}} \{\frac{1}{2NM}\{\|s_{nmt} - [\delta_{1t}^{\top} x_{nt} + \delta_{2t}^{\top}(x_{nt} \times \tau_m)\|_2^2]\} \tag{15}$$

subject to $\|\tilde{\delta}_t\|_1 \leq c$, where $\tilde{\delta}_t$ is the vector of slope parameters, $\|\cdot\|_2$ is the vector Euclidean norm, and c is the tuning parameter which shrinks and forces coefficients equal to zero. Smaller values of c restrict the dimension of the parameter space by forcing more coefficients to zero, while larger values tend to include more covariates up until convergence to the OLS solution. Since c controls the complexity of the model, a key issue is how to select the best value for this parameter. As pointed out in Chinco et al. (2019), there is no *a priori* theoretically optimal value for c. Therefore, we rely on the standard cross-validation procedure, through which the data set (the cross-section of sovereign CDS contracts) is split into two sub-sets, using one sub-set (the *training set*) to estimate the model and then judge the goodness of the prediction based on the remainder of the data (the *test set*).

More specifically, we first randomly split the full data-set into n sub-sets, each containing the same number of observations. Typically n varies between 5 to 10. In our implementation, the LASSO estimation is run following the cyclical coordinate descent algorithm outlined in Friedman et al. (2010) and developed in the the `R` package `glmnet`, where we set n equal to default value of 10. Hence, the algorithm estimates the model based on $n-1$ $(=9)$ data sets and the remaining data set is used to evaluate the out-of-sample model performance in terms of root mean squared error (RMSE) of the predictions. This process is executed for different values of c, thereby obtaining predictions from a variety of models, ranging from the all-inclusive model ($c \to \infty$) to the model with no covariates ($c \to 0$). This process is repeated for all the n data sets, each one playing the role of the test set, while the remaining $n-1$ groups act as the training set.

We then obtain n estimates of the prediction error for different values of c, which are then averaged out producing the cross-validation error for each value of c. The LASSO solution corresponds to the model showing the minimum cross-validation error. In our implementation, we use an increasing sequence of values for c, starting at the lowest value c_{min}, for which the entire vector of slope parameters is set to zero, and then adding increments, up to the value c_{max}, for which we have the OLS solution. In total, for each cross-section, we use a sequence of 400 values for c, which is substantially higher than the default setting in `gmlnet` of 100, with the aim of improving the accuracy of the modelling choice.

III Time-varying sensitivities to the macro factors

The daily time series of the intercept and the most important cross-sectional LASSO-type coefficients are displayed in Figures 2 and 3, respectively.[12] Since the covariates are standardized by their cross-sectional standard deviation before running the procedure, the coefficients are scale-independent, which helps to assess the variable importance for each macro-fundamental based on the absolute value of the coefficient itself.

At first glance, the dynamics of the stacked coefficients exhibit significant time variation, with a prominent role played by the average term-structure level effect (δ_{01t}). The macro-

[12] These figures are further discussed in Section IV, where we identify regimes characterized by different dynamics of the macro-sensitivities.

sensitivities exhibit an increased variability in 2010, while they are essentially dormant in 2009, to explode later in 2011-2012, and calm down in 2013. Our findings are generally consistent with Bernoth and Erdogan (2013), who show that the impact of fiscal policy variables and general investors' risk aversion on sovereign yield spreads in Europe was not constant over time. However, our results are in contrast with the authors' view that it is plausible to think of the time-varying sensitivities as "changing gradually over time, rather than having a discrete break-point between regimes." Our results indicate, instead, that the time-varying macro-sensitivities exhibit substantial jumpiness in their dynamics, with sudden and discrete changes in "regime."

Tables 4 and 5 report summary statistics computed over the entire period for all the cross-sectional coefficient estimates. The column "Zeros" is informative as to the number of times, expressed as ratio over the total number of cross-sections, in which the variable made no contribution in explaining the cross-section of CDS spreads and was discarded by the LASSO algorithm. This number should be read carefully, as it is informative only about the "importance persistence" of the variable, regardless of how much the specific explanatory power of that variable was—this analysis is performed in the next section. The bank risk indicator was the most selected level-effect variable, being discarded in less than one third of the estimations, while among the slope-effect variables, GIPSI exhibits a ratio of 0.688, the lowest value among all slope-effect coefficient estimates.

A few findings are worthy of note. *First*, the coefficient for GDP growth has the highest average value (in absolute terms), both for level and slope effects.[13] GDP growth impacts negatively (positively), on average, the level (slope) of the term structure of CDS spreads.

[13]Since all regressors are cross-sectionally standardized and de-meaned, the intercept represents the average CDS spread across all countries and maturities, whereas the level and slope coefficients represent the effect, in basis points, of a one (cross-sectional) standard deviation increase in the corresponding covariates.

Assuming, in first approximation, a constant expected loss given default, this evidence is consistent with the notion that an increase in GDP growth reduces the (risk-adjusted) marginal probability of default, but that this effect is stronger at short horizons. *Second*, the term structure slope effect is flattening-oriented for most of the variables, except for GIPSI, inflation, credit to private sector, credit to government, and exports. For these macro-variables, all showing positive level effect-type coefficients, the longer the CDS maturity, the higher the impact on the CDS spread. *Third*, Min and Max denote high values for all coefficients, thereby reflecting substantial spikes and, in turn, "jumpy" dynamics of macro-sensitivities, also confirmed by Figure 3.

While the analysis above provides us with a full description of the time-series dynamics of the macro-sensitivities, the number of covariates, and their corresponding time-varying coefficients, complicate the understanding of the underlying economic developments. To deal with this dimensional problem, in the next section we introduce a simple statistical procedure to detect homogeneous groups of observations for the cross-sectional regression coefficients which, in turn, identify regimes in the macro-sensitivity behavior. These regimes allow us to come up with a synthetic characterization on the changing nature of sovereign risk in Europe during the 2009–2013 period.

IV Macro-sensitivity regimes

In this section, we present results on the identification of macro-sensitivity regimes based on the time-series dynamics of the LASSO-type coefficients. As discussed in the introduction, existing studies on sovereign CDS regimes look at CDS/bond spreads to identify regime changes and incorporate structural changes in the econometric relationships between the spreads and the macroeconomic covariates (e.g., Blommestein et al., 2016; Delatte et al., 2017). Other authors (e.g., Afonso et al., 2018) arbitrarily define regimes as time dummies based on the ECB policy intervention decisions and then explore how sovereign risk sensitivity changed once these measures took place. Our approach differs, as we use the data on macro-sensitivities to identify regimes conceived as homogeneous groups of observations over the entire observed time period.

IV.A Methodology

Macro-sensitivity regime identification is based on Kaufman and Rousseeuw (1990)'s clustering algorithm: Partitioning Around Medoids (PAM). This algorithm maps a distance matrix into a specified number of clusters using the concept of "medoids" as the representation of the cluster centers. Let $\delta_t^\top = \{\delta_{10t}, \delta_{20}, \delta_{11t}^\top, \delta_{12t}^\top, \delta_{21t}^\top, \delta_{22t}^\top\}$ denote the generic row vector of the $P = T \times [(K+1)M + 1]$ matrix containing the time-varying parameters from equation (14), and denote by $d(\delta_{t_i}, \delta_{t_j})$ the dissimilarity between parameter estimates at time t_i and time t_j. Let, now, $\boldsymbol{D}$ be the $P \times P$ symmetric matrix of dissimilarities.[14] Using data from such matrix $\boldsymbol{D}$, the algorithm minimizes the distance between δ_t and a center, i.e. the "medoid" of that cluster, chosen among the T rows of the matrix δ. Therefore, medoids are robust representations of the corresponding clusters and act as "mass points" in the space of parameters δ. In our context, these clusters denote homogeneous time dynamics of the macro-sensitivities around medoids, and as such identify specific "regimes." These regimes have distinctive features that we explore by focusing on the changing structure of the sovereign risk sources due to both a shift in macroeconomic fundamentals and changes in risk pricing.

Computationally, the procedure needs to pre-specify the number of clusters before run-

[14] A common distance one can use is, as in this paper, the Euclidean distance.

ning the algorithm. In principle, we may view this number as given and related to some *a priori* theoretical reasoning or empirical evidence. Alternatively, the number could be data-driven, based on some of the existing criteria proposed in the literature (see Kaufman and Rousseeuw, 1990). We choose the number of regimes by running a specific F-test-based clustering method which looks at the percentage of the explained variance—more precisely, the ratio of the between-group variance to the total variance—as a function of the number of clusters.[15] The criterion is commonly used in the literature and is based on the between-group variance, consistent with the concept of "distance" used to identify the homogeneous cluster of observations. Having the objective to pre-specify the number of crisis regimes, we run the test over the 4/1/2010–4/25/2013 period, thereby arbitrarily establishing the 5/11/2009–3/31/2010 sub-period as the "pre-crisis" regime. This is consistent with the empirical evidence we discussed in the introduction, as the surge of CDS/bond spreads of GIPSI countries occurs in April 2010, when Greece activates the 45 billion Euros EU-IMF bailout and S&P downgrades Greek debt to junk status.

After running the preliminary test to pre-specify the number of clusters, we next execute the PAM algorithm, and we then scrutinize each regime by measuring the contribution of each variable in cross-sectional variance of sovereign CDSs spreads. We do this by decomposing the explained cross-sectional variance of the CDS spreads, $\mathrm{var}_t^{cs}(\hat{s}_{nmt})$, according to the following expression:

[15] More specifically, the number of clusters is chosen by comparing the percentage of variance explained by the clusters against the number of clusters: the appropriate number of clusters corresponds to the point in which the marginal gain, expressed in terms of additional explained variance, drops, thereby signaling no significant information added by the last cluster.

$$
\begin{aligned}
\mathrm{var}_t^{cs}(\hat{s}_{n\tau_m t}) &= \sum_{n=1}^{N}\sum_{m=1}^{M}\left[\sum_{k=1}^{K+1}\mathrm{cov}_t^{cs}(\hat{s}_{n\tau_m t},\hat{\delta}_{11kt}x_{nkt}) + \mathrm{cov}_t^{cs}(\hat{s}_{n\tau_m t},\hat{\delta}_{12t}\mathrm{GIPSI}_n)\right] \\
&+ \sum_{n=1}^{N}\sum_{m=1}^{M}\left[\sum_{k=1}^{K+1}\mathrm{cov}_t^{cs}(\hat{s}_{n\tau_m t},\hat{\delta}_{21kt}x_{nkt}\tau_m) + \mathrm{cov}_t^{cs}(\hat{s}_{n\tau_m t},\hat{\delta}_{22t}\mathrm{GIPSI}_n\tau_m)\right]. \quad (16)
\end{aligned}
$$

Hence, the cross-sectional variance of CDS spreads is split into components due to the different $K+1$ explanatory variables, where the individual contribution by each variable is computed as:

$$
\frac{\sum_{n=1}^{N}\sum_{m=1}^{M}\mathrm{cov}_t^{cs}(\hat{s}_{n\tau_m t},\hat{\delta}_{11kt}x_{nkt})}{\mathrm{var}_t^{cs}(\hat{s}_{n\tau_m t})};\ \frac{\sum_{n=1}^{N}\sum_{m=1}^{M}\mathrm{cov}_{cs}(\hat{s}_{n\tau_m t},\hat{\delta}_{12t}\mathrm{GIPSI}_n)}{\mathrm{var}_t^{cs}(\hat{s}_{n\tau_m t})} \quad (17)
$$

for level effects, and:

$$
\frac{\sum_{n=1}^{N}\sum_{m=1}^{M}\mathrm{cov}_t^{cs}[\hat{s}_{n\tau_m,t},\hat{\delta}_{21kt}(x_{nkt}\times\tau_m)]}{\mathrm{var}_t^{cs}(\hat{s}_{n\tau_m t})};\ \frac{\sum_{n=1}^{N}\sum_{m=1}^{M}\mathrm{cov}_t^{cs}[\hat{s}_{n\tau_m t},\hat{\delta}_{22t}(\mathrm{GIPSI}_n\times\tau_m)]}{\mathrm{var}_t^{cs}(\hat{s}_{n\tau_m t})} \quad (18)
$$

for slope effects. While such a decomposition changes over time, we compute averages, conditional on the regime, thereby identifying the variables that matter most, on average, during the non-crisis and crisis regimes.

IV.B Results

The regimes obtained as the output of the clustering procedure detailed in the previous section tell us that, within each cluster, macro-sensitivities exhibited distinctive and homogeneous patterns of behavior. We now scrutinize the regimes, focusing on the summary statistics of the LASSO-type coefficients and the cross-sectional variance decompositions, equations (17) and (18). Each regime can be then identified with the key variables driving the cross-section of CDS spreads.

We display the regime time-line in Figure 2, where we see the three crisis regimes (regimes 1, 2, and 3), as well as the pre-crisis regime arbitrarily set from 5/11/2009 to 3/31/2010 (regime 0), where the different shaded areas identify the different regimes. In the same figure, we also display the time-varying intercept, which corresponds, by construction, to the cross-sectional average of sovereign CDS spreads.[16]

The regimes identified by our PAM procedure are broadly consistent with the turning points discussed in the introduction. The pre-crisis regime is characterized by moderate, but increasing CDS spreads, until April 2010. At that point we have the Greece-driven transition to the first crisis regime, which continues until the third quarter of 2011. We then have fluctuations between the first and second regime until November 2011, when we enter the third and most risky regime, in terms of the cross-sectional average of CDS spreads. Starting with September 2012, the average CDS spread comes down as we first transition to the second regime, and finally we revert back to the first regime, starting in March 2013, when the average CDS spread is around 220 basis points.

To better interpret the pre-crisis and crisis regimes, Tables 6–7 report summary statistics of the LASSO-type coefficients conditional on regimes, and Table 8 shows the value of the cross-sectional CDS spread variance explained by each variable according to equations (17) and (18). The values of explained variance are also normalized by the highest value as $\frac{v_{kt}}{v_{kth}} \times 100$, where v_k is the value of the cross-sectional CDS variance explained by variable k and v_{kth} is highest value of the explained variance out of all $K+1$ covariates. With the purpose of identifying the most informative variables for each regime, we set 50% as the cut-off value for this normalized measure, thereby highlighting only those variables showing an explanatory power not less than 50%, compared to the most informative variable associated.

[16] Remember that the LASSO-type regression was run by cross-sectionally de-meaning and standardizing each covariate.

Such an approach seems reasonable in order to obtain a parsimonious characterization of each regime.

The following patterns emerge:

- The pre-crisis regime, from May 2009 to March 2010, displays moderate, while increasing, macro-sensitivities, with the deficit and GIPSI variables being the main determinants of CDS spreads, followed by the banking risk measure and the banking risk volatility indicator.

- The first crisis regime covers the April 2010–July 2011 period—which includes the 40 billion Euros Greek bailout and the ensuing widening of the spreads of peripheral Eurozone countries—and the mid-March–April 2013 period, when Cyprus secured a 10 billion Euros bank bailout from the European Union and the IMF. During this regime, GIPSI and Loans-to-Government were the main drivers of the cross-section of sovereign CDS spreads, which averaged around 167 bps.

- The second crisis regime comprises the June–August 2011 and September 2012–mid March 2013, periods, namely the periods before the peak of the Eurozone crisis and following the third turning point (Draghi's July 2012 speech and the announcement of the OMT program in September 2012), during which we also have the agreement on the part of European leaders for the European Stability Mechanism to directly recapitalize banks, rather than having to act through national governments (October 19, 2012). This is an intermediate regime, going from regime 1 to 3 and also from 3 to 1 (see Figure 1a), during which imports and changes in inventories over GDP were the most influential variables, because of their effect on GDP growth through balance of payment pressures and increased macroeconomic volatility. The average value of of sovereign CDS spreads was around 248 bps. Our coefficient estimates for changes in inventories over GDP, on average positive during the regime, are consistent with recent evidence (European Commission, 2015) documenting how firms viewing their inventory stocks as "too large" are expected to react by cutting production in the following months, thereby exacerbating economic downturn during crisis periods.

- The third crisis regime takes place between July 2011 (when we have a rebound between regime 2 and 3) and August 2012. This regime corresponds to the highest risk phase when average spreads reached 400 bps and the the cross-section of spreads was mostly explained by GDP growth and employment. The regime includes the Italian government crisis (November 2011) and the release of the results of the second round of pan European stress tests (eight European banks failed the stress tests, while 16 were in a "danger zone").

Our characterization of non-crisis and crisis regimes offers new insights on the economic mechanisms underlying the Eurozone sovereign debt crisis. As pointed out in De Grauwe and Ji (2013), one view of the crisis is that the surging spreads from 2010 to mid-2012 were the result of deteriorating fundamentals and the market was just a messenger of bad news. A second view is that, beginning in 2010, the spreads were panic-driven away from country fundamentals. The first view would explain why austerity-based measures should be the right measure of policy intervention. The implications of the second view is that in times of market panic, central banks should act as liquidity providers.

Our findings accommodate both views. We show that fear and panic disconnected spreads from fundamentals, but only in regime 1, when being in the GIPSI group of countries was the key driver for the surge of CDS. In regime 3, on the other hand, at the peak of the crisis, markets restored a fundamental-based connection with GDP growth, even when the

ECB intervened to provide essentially unlimited support to the government bond markets. Then, we move towards regime 2, finally returning to regime 1, where GIPSI is again the main risk factor, but at a lower average level of spreads and lower GIPSI sensitivity. In this context, regime 2 appears to be a transition regime, in which we do have a connection between spreads and risk signals from imports and inventory dynamics.

Interestingly, the interpretation suggested above is also consistent with the effects of the macro-fundamentals on the CDS spreads during the different stages of the crisis (see Table 7). During the pre-crisis period, macro-fundamentals tend to impact positively both the level and the slope of the CDS spread curve ("steepening" effect). This is consistent with the notion that default risk premia, which are more relevant at longer maturities, drive much of the variation in the term structure of CDS spreads. On the other hand, when the crisis is the most acute (regime 3), macro-fundamentals tend to impact the level and slope of the spread curves in opposite directions. This is consistent with the notion that the expected occurrence of default, which is more relevant at short maturities, drives spread variation. Also consistent with the notion that fundamentals are most relevant at the height of the crisis is the evidence in Table 9, showing that the average cross-sectional R-square is highest in regime 3.

V Macro-sensitivities and financial market volatility

In this section, we explore the link between our LASSO-type coefficients (the macro-sensitivities) and future equity market volatility. The economic reason for this link is that in a CAPM setting equity market volatility is directly related to the volatility of the underlying pricing kernel and the aggregate market price of risk. Hence, we would expect our LASSO-type coefficients to correlate with future implied equity volatility for the European equity markets, as it is proxied by the one-month volatility index of the Euro Stoxx 50 index (VSTOXX).

Computationally, we run an out-of-sample exercise using observations for the 08/01/2012–04/25/2013 period to *dynamically* estimate, i.e., by adding one observation each day, several predictive models. As alternative instruments to the cross-sectional LASSO-type regression coefficients obtained in the analysis of the previous section, we consider:

- the coefficients of LASSO-type cross-sectional regressions of the CDS spreads on the country-specific first principal components extracted from each of the five groups of macro and financial indicators described in Section I;[17]
- the five-year country-specific CDS spreads;
- the GIPSI and NON-GIPSI first principal components extracted from the sovereign CDS spreads;
- the GIPSI and NON-GIPSI first principal components extracted from the real-time macroeconomic variables.

To assess the predictive ability of different approaches we rely on the Root Mean Squared Errors (RMSE) and the Mean Absolute Squared Errors (MAPE). The results in Table 10 show that the preferred approach is the "parsimonious" version of our LASSO-on-LASSO-type coefficients, namely the LASSO-type penalty regressions run on the LASSO-type Fama-MacBeth procedure on country-specific principal components. This approach dominates all alternatives in terms of both RMSE and MAPE.

As reported in the same Table 10, the Diebold-Mariano test confirms the robustness of our results. Another interesting finding has to do with the time dynamics of the estimated

[17] Principal components are estimated dynamically as well, as in Beber et al. (2015), within the following clusters of macro-variables: (1) Labour market; (2) Prices and costs; (3) Money and credit; (4) Output; (5) Banking.

coefficients for the best performer. In Table 11 we report the absolute values of the average coefficient estimates as well as the ratios of the absolute values of the individual coefficient estimates over the sums of the absolute values of all the coefficient estimates. While slope effects play a minor role in explaining the cross-section of CDS spreads, the slope-effect coefficients have a substantial role in this out-of-sample exercise, with an average weight around 33%, versus 55% for level-effect coefficients, while the remaining 10% is accounted for by the time-varying intercept.

VI Conclusions

We construct a new real-time, daily-frequency data set to examine the relation between sovereign CDS spreads and macro-economic fundamentals during the Eurozone sovereign debt crisis. We provide several new and important results. First, we document pronounced time-variation in the sensitivity of CDS spreads to the country-specific macro indicators. Second, we identify three distinct risk regimes based on the general level of CDS premia, the sensitivity of CDS premia to different macro indicators, and the GIPSI connotation. It is during the regime corresponding to the most intense phase of the crisis that CDS spreads reflected macro fundamentals the most, whereas before the crisis it was only the GIPSI connotation to matter. Third, we show how the macro-sensitivities predict future equity market volatility better than competitive sets of instruments, consistent with the notion that expected future risk is an important driver of how CDS spreads react to macro information. We also show that slope effects are important in predicting implied equity volatility. In summary, we provide a new and complete characterization of the links between CDS spreads, macro fundamentals, and default risk. We trust that this characterization will prove useful to both market participants and policy makers.

References

Acharya, V., Eisert, T., Eufinger, C., and Hirsch, C. (2018). Real effects of the sovereign debt crisis in europe: Evidence from syndicated loans. *Review of Financial Studies*, 34:2855–2896.

Afonso, A., Arghyrou, M., Gadea, M. D., and Kontonikas, A. (2018). whatever it takes to resolve the european sovereign debt crisis? bond pricing regime switches and monetary policy effects. *Journal of International Money and Finance*, 86:1–30.

Aizenman, J., Hutchison, M., and Jinjarak, Y. (2018). What is the risk of european sovereign debt defaults? fiscal space, cds spreads and market pricing of risk. *Review of Financial Studies*, 34:2855–2896.

Alessi, L. and Detken, C. (2011). Quasi real time early warning indicators for costly asset price boom/bust cycles: A role for global liquidity. *European Journal of Political Economy*, 27:520–533.

Alessi, L. and Detken, C. (2014). On policymakers' loss functions and the evaluation of early warning systems: Comment. *Economics Letters*, 124:338–340.

Augustin, P. (2018). The term structure of cds spreads and sovereign credit risk. *Journal of Monetary Economics*, 96:53–76.

Balduzzi, P., Elton, E., and Green, C. (2001). Economic news and bond prices: Evidence from the u.s. treasury market. *Journal of Financial and Quantitative Analysis*, 36:523–543.

Beber, A., Brandt, M., and Luisi, M. (2014). Eurozone sovereign yield spreads and diverging economic fundamentals. *Working Paper*.

Beber, A., Brandt, M., and Luisi, M. (2015). Distilling the macroeconomic news flow. *Journal of Financial Economics*, 117:489–507.

Beetsma, R., Giuliodori, M., de Jong, F., and Widijanto, D. (2013). Spread the news: The impact of news on the european sovereign bond markets during the crisis. *Journal of International Money and Finance*, 34:83–101.

Beirne, J. and Fratzscher, M. (2013). The pricing of sovereign risk and contagion during the european sovereign debt crisis. *Journal of International Money and Finance*, 34:60–82.

Bekaert, G. and Hoerova, M. (2014). The vix, the variance premium and stock market volatility. *Journal of Econometrics*, 183:181–192.

Bekaert, G., Hoerova, M., and Lo Duca, M. (2013). Risk, uncertainty and monetary policy. *Journal of Monetary Economics*, 60:771–788.

Bernoth, K. and Erdogan, B. (2013). Sovereign bond yield spreads: A time-varying coefficient approachs. *Journal of International Money and Finance*, 34:639–656.

Blommestein, H., Eijffinger, S., and Qian, Z. (2016). Regime-dependent determinants of Euro area sovereign CDS spreads. *Journal of Financial Stability*, 22:10–21.

Bolton, P. and Olivier (2011). Sovereign default risk and bank fragility in financially integrated economies. *IMF Economic Review*, 59:162–194.

Brodie, J., Daubechies, I., De Mol, C., Giannone, D., and Loris, I. (2009). Sparse and stable markowitz portfolios. *Proceedings of the National Academy of Sciences*, 106(30):12267–12272.

Bruder, B., Gaussel, N., Richard, J.-C., and Roncalli, T. (2013). Regularization of portfolio allocation. *SSRN Electronic Journal*.

Brunnermeier, M., Garicano, L., Lane, P., Pagano, M., Reis, R., Santos, T., Thesmar, D., Nieuwerburgh, S. V., and Vayanos, D. (2016). The sovereign-bank diabolic loop and esbies. *American Economic Review*, 106:508–512.

Chinco, A., Clark-Joseph, A. D., and Ye, M. (2019). Sparse signals in the cross-section of returns. *The Journal of Finance*.

Danielsson, J., Valenzuela, M., and Zer, I. (2018). Learning from history: Volatility and financial crises. *Review of Financial Studies*, 31:2774–2805.

De Grauwe, P. and Ji, Y. (2013). Panic-driven austerity in the eurozone and its implications. *VOX*.

De Mol, C., Giannone, D., and Reichlin, L. (2008). Forecasting using a large number of predictors: Is bayesian shrinkage a valid alternative to principal components? *Journal of Econometrics*, 146(2):318 – 328.

de Santis, R. (2015). A measure of redenomination risk. *Working Paper*.

Delatte, A.-L., Fouquau, J., and Portes, R. (2017). Regime-dependent sovereign risk pricing during the Euro crisis. *Review of Finance*, 21:363–385.

DeMiguel, V., Garlappi, L., and Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *The Review of Financial Studies*, 22(5):1915–1953.

Dieckmann, S. and Plank, T. (2011). Default risk of advanced economies: an empirical analysis of credit default swaps during the financial crisis. *Review of Finance*, 16:903–934.

Donoho, D. L. (2006). Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306.

Ehrmann, M. and Fratzscher, M. (2005). Exchange rates and fundamentals: new evidence from real-time data. *Journal of International Money and Finance*, 24(2):317–341.

European Commission (2015). Highlight: Revisiting the relation between inventories and production using survey data. *European Business Cycle Indicators*, 1st quarter:7–11.

Fan, J., Lv, J., and Qi, L. (2011). Sparse high-dimensional models in economics. *Annual Review of Economics*, 3(1):291–317.

Fan, J., Zhang, J., and Yu, K. (2012). Vast portfolio selection with gross-exposure constraints. *Journal of the American Statistical Association*, 107(498):592–606.

Faust, J., Rogers, J., Wang, S.-Y., and Wright, J. (2007). The high-frequency response of exchange rates and interest rates to macroeconomic announcements. *Journal of Monetary Economics*, 54:1051–1068.

Feng, G., Giglio, S., and Xiu, D. (2019). Taming the factor zoo: A test of new factors. Working Paper 25481, National Bureau of Economic Research.

Freyberger, J., Neuhierl, A., and Weber, M. (2017). Dissecting characteristics nonparametrically. Working Paper 23227, National Bureau of Economic Research.

Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software, Articles*, 33(1):1–22.

Gefang, D. (2014). Bayesian doubly adaptive elastic-net lasso for var shrinkage. *International Journal of Forecasting*, 30(1):1 – 11.

Ghysels, E., Horan, C., and Moench, E. (2018). Forecasting through the rearview mirror: data revisions and bond return predictability. *Review of Financial Studies*, 31:678–714.

Giannone, D., Henry, J., Lalik, M., and Modugno, M. (2012). An area-wide real-time database for the Euro area. *Review of Economics and Statistics*, 94:1000–1013.

Giannone, D., Reichlin, L., and Small, D. (2005). Nowcasting gdp and inflation: The real time informational content of macroeconomic data release. *Working Paper*.

Kaufman, L. and Rousseeuw, P. (1990). Finding groups in data: An introduction to cluster analysis. *Wiley*.

Kim, S.-J., Salema, L., and Wu, E. (2015). The role of macroeconomic news in sovereign cds markets: domestic and spillover news effects from the u.s., the Eurozone and China. *Journal of Financial Stability*, 18:208–224.

Kock, A. B. (2016). Consistent and conservative model selection with the adaptive lasso in stationary and nonstationary autoregressions. *Econometric Theory*, 32(1):243–259.

Kock, A. B. and Callot, L. (2015). Oracle inequalities for high dimensional vector autoregressions. *Journal of Econometrics*, 186(2):325 – 344.

Krishnamurthy, A., Nagel, S., and Vissing-Jorgensen, A. (2018). Ecb policies involving government bond purchases: impact and channels. *Review of Finance*, 22:1–44.

Li, J. and Chen, W. (2014). Forecasting macroeconomic time series: Lasso-based approaches and their forecast combinations with dynamic factor models. *International Journal of Forecasting*, 30(4):996 – 1015.

Longstaff, F., Pan, J., Pedersen, L., and Singleton, K. (2011). How sovereign is sovereign credit risk? *American Economic Journal*, 3:75–103.

Manasse, P. and Roubini, N. (2009). “rules of thumb” for sovereign debt crises. *Journal of International Economics*, 78:192–205.

Nosbusch, J. H. Y. (2010). Determinants of sovereign risk: Macroeconomic fundamentals and the pricing of sovereign debt. *Review of Finance*, 14:235–262.

Pasquariello, P. (2014). Financial market dislocations. *Review of Financial Studies*, 27:1868–1914.

Scherer, B. (2015). *Portfolio Construction and Risk Budgeting.* Risk Books.

Song, S. and Bickel, P. (2011). Large vector auto regressions. Available at arxiv.org/abs/1106.3915.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society*, 58:267–288.

Veronesi, A. D. P. (2014). Investors' and central bank's uncertainty embedded in index options. *Review of Financial Studies*, 27:1661–1716.

Veronesi, P. (1999). Stock market overreactions to bad news in good times: A rational expectations equilibrium model. *Review of Financial Studies*, 12:975–1007.

Table 1: Macroeconomic Variables in Real-Time

Variable	**Description**	**Cluster**
unempl	Unemployment rate	Labour market
empl	Employment rate, total	Labour market
infl	Inflation rate	Price and costs
ind_price	Industrial Producer Prices (% change)	Price and costs
labour	Hourly labour cost (price index) (% change)	Price and costs
m3	M3 (variation)	money, credit and debt
loan_priv	Loans to private sector (variation)	money, credit and debt
loan_gov	Loans to government (variation)	money, credit and debt
cr_priv	Credit to private sector (variation)	money, credit and debt
cr_gov	Credit to government (variation)	money, credit and debt
deficit	Public sector deficit over GDP	money, credit and debt
gdp	Real GDP growth	output
cons	Consumption growth	output
gov_cons	Government consumption growth	output
inv	Investment growth	output
invent_gdp	Changes in inventories over nominal GDP	output
ex	Exports growth	output
im	Imports growth	output
ind_prod	Industrial production growth (price index)	output
bank	Banking risk proxy	banking
vol_bank	20 days rolling windows realized volatility of bank	banking

The table reports the list of macroeconomic variables we collected in real-time and used in our analysis. Variables are grouped according to their ownership category reported in the column Cluster: (i) labour market; (ii) price and costs; (iii) money, credit and debt; (iv) output; (v) banking.

Table 2: Descriptive Statistics, Non-GIPSI

Variable	AT		BE		CY		DE		FI		FR		NL		All Non-GIPSI	
	avg	**std**	**avg**	**std**	**avg**	**std**	**avg**	**std**	**avg**	**std**	**avg**	**std**	**avg**	**std**	**avg**	**std**
infl	2.14	1.22	2.10	1.48	2.39	1.45	1.58	0.91	2.46	0.90	1.62	0.95	1.89	1.06	2.02	1.14
gdp	0.66	2.49	0.24	2.14	-0.61	1.78	0.57	3.54	-0.22	4.57	0.32	1.65	-0.48	2.43	0.07	2.66
unempl	4.42	0.43	7.73	0.54	8.31	2.58	6.42	0.87	8.10	0.48	9.95	0.41	4.51	0.74	7.06	0.86
ind price	0.83	5.15	0.59	9.20	0.27	10.14	-1.62	8.62	3.18	5.99	0.23	7.19	-0.96	11.52	0.36	8.26
empl	0.86	1.50	0.75	0.91	-0.82	1.76	0.79	0.76	-0.09	2.21	1.02	2.45	-0.16	1.09	0.34	1.53
cons	2.69	4.66	2.99	6.47	2.81	6.82	2.76	3.03	2.70	4.61	2.26	3.60	2.10	5.95	2.62	5.02
gov cons	3.76	4.30	5.02	7.03	7.65	13.11	3.10	1.91	5.62	7.77	4.69	4.83	7.12	7.11	5.28	6.58
inv	1.67	9.60	-0.71	5.20	-5.35	9.82	-1.73	6.54	-0.04	11.44	0.25	6.16	-1.06	10.86	-1.00	8.52
invent gdp	0.26	2.94	-0.25	2.24	0.68	8.79	0.10	2.75	0.15	2.88	-0.14	1.91	-0.09	1.28	0.10	3.26
export	2.45	12.58	1.24	12.27	0.72	14.03	4.76	13.76	-0.25	16.47	2.79	9.75	4.19	11.86	2.27	12.96
import	1.18	9.87	0.69	9.17	-3.98	9.13	2.77	8.45	1.02	11.12	-0.13	7.29	2.08	7.51	0.52	8.94
ind prod	-1.35	13.20	0.13	17.89	-8.41	6.61	0.30	12.65	-3.70	13.43	-1.71	7.42	-1.05	7.84	-2.25	11.29
m3	2.48	3.05	2.44	2.31	7.15	8.29	3.53	2.78	2.88	2.67	2.08	3.63	4.72	2.66	3.61	3.63
loan priv	2.51	1.93	-3.05	3.29	8.59	5.04	0.95	1.52	5.63	2.49	2.76	2.79	1.68	3.45	2.73	2.93
loan gov	2.76	3.14	2.23	13.33	-3.77	4.18	1.50	8.87	13.49	3.60	2.92	7.39	5.13	11.33	3.47	7.40
cr priv	2.82	1.49	-0.83	1.96	8.24	7.06	-0.53	2.34	5.33	2.20	1.95	2.75	1.40	2.24	2.62	2.86
cr gov	9.12	7.39	-1.62	4.10	12.68	48.88	5.79	8.67	18.19	12.56	2.00	11.74	7.93	8.69	7.73	14.57
labour	-1.18	6.38	-1.58	9.98	-6.89	15.01	-1.27	5.50	-4.35	12.26	-3.47	10.92	-5.03	12.38	-3.40	10.35
deficit	-3.02	1.00	-4.04	1.06	-4.88	2.14	-1.96	1.47	-0.93	2.06	-6.08	1.23	-3.89	2.03	-3.54	1.57
bank	75.51	20.65	84.73	29.70	77.47	32.19	72.73	16.31	96.21	14.33	77.78	20.24	107.33	14.71	84.54	21.16
vol bank	0.02	0.01	0.03	0.01	0.03	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01

The table reports summary statistics of macroeconomic variables in real-time (see Table 1) for non-GIPSI countries computed over the entire period from 5/11/2009 to 4/25/2013. avg is the arithmetic average and std is the standard deviation. AT is Austria, BE is Belgium, CY is Cyprus, DE is Germany, FI is Finland, FR is France. Values are in percentage form.

Table 3: Descriptive Statistics, (G)IPSI

Variable	ES		IE		IT		PT		(G)IPSI	
	avg	**std**	**avg**	**std**	**avg**	**std**	**avg**	**std**	**avg**	**std**
infl	1.92	1.35	0.01	1.82	2.22	1.11	1.75	1.77	1.47	1.51
gdp	-1.08	1.65	-2.05	3.39	-1.38	2.45	-1.44	1.95	-1.49	2.36
unempl	21.60	2.59	13.85	0.99	8.88	1.25	12.38	2.52	14.18	1.84
ind price	-0.31	8.36	0.43	5.47	-1.12	8.28	0.45	8.00	-0.14	7.53
empl	-2.43	3.93	-3.93	3.42	-0.67	0.73	-2.35	1.21	-2.34	2.32
cons	2.92	8.81	2.32	9.68	2.71	6.92	1.80	6.10	2.44	7.88
gov cons	5.03	7.05	6.14	15.17	3.53	7.25	3.78	10.21	4.62	9.92
inv	-3.57	14.77	-5.01	25.97	-1.30	9.21	-3.16	5.64	-3.26	13.90
invent gdp	0.27	4.40	0.03	1.38	-0.20	1.46	-0.30	1.32	-0.05	2.14
ex	4.77	11.97	4.02	4.41	1.60	13.63	7.21	9.94	4.40	9.99
im	-3.07	9.88	-0.49	6.17	1.25	11.96	-0.61	8.10	-0.73	9.03
ind prod	-4.76	8.56	-5.72	13.06	-2.42	7.01	-2.89	5.08	-3.95	8.43
m3	-1.28	3.53	-5.06	9.51	2.01	3.71	-2.05	4.27	-1.60	5.26
loan priv	-1.88	3.01	-10.86	5.05	3.30	3.45	-0.70	4.08	-2.54	3.90
loan gov	22.06	8.39	81.94	193.86	3.64	2.05	22.98	42.50	32.65	61.70
cr priv	-0.13	5.68	-7.09	4.31	3.87	5.60	1.28	7.10	-0.52	5.67
cr gov	19.96	13.32	7.15	22.35	11.96	5.42	35.07	26.45	18.53	16.89
labour	-5.48	12.39	-8.15	12.84	-1.63	11.40	-4.83	9.89	-5.02	11.63
deficit	-9.11	1.92	-16.55	7.13	-4.26	0.67	-6.84	2.46	-9.19	3.05
bank	82.85	17.83	43.11	43.49	77.71	18.66	78.62	22.60	70.57	25.64
vol bank	0.02	0.01	0.04	0.02	0.02	0.01	0.02	0.01	0.03	0.01

The table reports summary statistics of macroeconomic variables in real-time (see Table 1) for GIPSI countries computed over the entire period from 5/11/2009 to 4/25/2013. avg is the arithmetic average and std is the standard deviation. ES is Spain, IE is Ireland, IT is Italy, PT is Portugal. Values are in percentage form.

Table 4: Summary Statistics, Level-effect LASSO Coefficients

	Zeros	Min	Max	Mean	StdDev
alpha, bank and GIPSI					
alpha	0.000	41.798	499.373	226.076	140.341
gipsi	0.688	0.000	158.968	15.380	32.197
bank	0.309	-222.888	27.936	-28.100	45.645
vol_bank	0.588	-13.894	357.487	23.674	54.555
employment					
unempl	0.634	-85.222	80.380	0.335	13.233
empl	0.603	-345.595	26.914	-30.741	60.036
prices and costs					
infl	0.622	-53.918	174.684	9.769	30.133
ind_price	0.492	-89.566	300.769	16.071	57.634
labour	0.597	-282.825	62.137	-5.672	25.868
money, debt, and credit					
m3	0.726	-167.288	71.262	-5.744	25.652
loan_priv	0.672	-135.109	56.167	-3.037	17.808
loan_gov	0.593	-110.797	253.166	23.120	59.678
cr_priv	0.740	-43.668	93.698	2.397	8.493
cr_gov	0.518	-209.265	98.281	2.517	37.375
deficit	0.602	-106.568	30.940	-8.031	16.444
output					
gdp	0.461	-380.469	30.770	-52.732	89.242
cons	0.591	-201.389	95.771	1.453	26.134
gov_cons	0.651	-234.760	248.033	3.978	47.828
inv	0.639	-205.101	68.453	-14.214	31.478
invent_gdp	0.492	-289.825	233.722	7.864	58.363
ex	0.606	-68.964	224.381	13.522	38.947
im	0.450	-274.770	55.389	-24.511	45.359
ind_prod	0.537	-189.701	41.456	-15.064	29.653
summary statistics - excluding alpha					
Mean	0.582	-168.708	126.853	-3.080	38.716
Min	0.309	-380.469	26.914	-52.732	8.493
Max	0.740	0.000	357.487	23.674	89.242
Mean (abs)				13.997	

This table reports summary statistics for the daily cross-sectional level-effect LASSO coefficient estimates computed over the entire period from 5/11/2009 to 4/25/2013. Cross-sectional regressions (equation (14)) are estimated by solving the LASSO problem (equation (15)). All regressors are cross-sectionally standardized and de-meaned. Therefore, the intercept (alpha) represents the average CDS spread across all countries and maturities, whereas the coefficients represent the effect, in basis points, of a one (cross-sectional) standard deviation increase in the corresponding covariates. Column Zeros is the number of times, expressed as ratio over the total number of cross-sections, in which the variable was discarded by the LASSO algorithm. Min, Max, Mean, StdDev are the minimum, the maximum, the arithmetic average and the standard deviation, respectively.

Table 5: Summary Statistics, Slope-effect LASSO Coefficients

	Zeros	Min	Max	Mean	StdDev	Slope
bank, GIPSI, and tau						
gipsi	0.597	-63.417	44.448	1.128	7.945	steep
bank	0.928	0.000	4.648	0.154	0.658	flat
vol_bank	0.763	-52.831	17.085	-2.638	9.558	flat
tau	0.787	-40.946	15.469	0.244	3.540	steep
employment						
unempl	0.949	-30.687	16.918	-0.208	1.863	flat
empl	0.624	-0.859	67.612	2.846	6.797	flat
prices and costs						
infl	0.780	-45.444	13.084	0.971	4.194	steep
ind_price	0.818	-54.111	7.685	-2.171	7.563	flat
labour	0.809	-18.248	24.839	1.017	3.355	flat
money, debt, and credit						
m3	0.711	-1.631	90.517	3.224	11.084	flat
loan_priv	0.897	-10.499	57.798	0.316	3.118	flat
loan_gov	0.666	-97.498	3.085	-7.290	16.823	flat
cr_priv	0.830	-22.589	14.696	0.157	2.647	steep
cr_gov	0.687	-10.459	47.307	3.673	9.655	steep
deficit	0.864	-0.736	70.762	2.552	9.678	flat
output						
gdp	0.791	-1.250	86.085	7.564	19.455	flat
cons	0.846	-66.662	24.350	-0.962	10.111	flat
gov_cons	0.811	-41.551	23.747	-0.289	4.952	flat
inv	0.859	-1.971	30.048	1.342	4.938	flat
invent_gdp	0.634	-41.738	36.978	-1.430	8.880	flat
ex	0.708	-27.494	29.648	1.654	7.458	steep
im	0.627	-6.610	61.308	2.265	8.840	flat
ind_prod	0.691	-6.829	78.093	2.039	7.313	flat
summary statistics						
Mean	0.768	-27.414	38.670	0.723	7.586	
Min	0.597	-97.498	3.085	-7.290	0.658	
Max	0.949	0.000	90.517	7.564	19.455	
Mean (abs)				2.006		

This table reports summary statistics for the daily cross-sectional slope-effect LASSO coefficient estimates computed over the entire period from 5/11/2009 to 4/25/2013. Coefficient estimates comes from the same cross-sectional regressions used for level-effect coefficients (equations (14))-(15)) and reported in Table 4, and relate to covariates interacted with CDS maturity as well as the maurity alone (tau). Column Zeros is the number of times, expressed as ratio over the total number of cross-sections, in which the variable was discarded by the LASSO algorithm. Min, Max, Mean, StdDev are the minimum, the maximum, the arithmetic average and the standard deviation, respectively. Slope denotes the flattening (flat) or steepeing (step) oriented effect for each variable: when the sign of the level- and slope-effect are the same, the term structure is steepening-oriented with higher impact on CDS spreads for longer maturities, otherwise (different sign of the level- and slope-effect coefficients) the term structure is flattening-oriented with higher impact on CDS spreads for shorter maturities.

Table 6: Level-effect LASSO Coefficients within Regimes

	Regime 0		Regime 1		Regime 2		Regime 3	
	Mean	**t-stat**	**Mean**	**t-stat**	**Mean**	**t-stat**	**Mean**	**t-stat**
alpha, bank and GIPSI								
alpha	64.229	10.447	166.738	8.966	247.012	4.319	423.974	21.672
gipsi	12.415	1.661	34.539	2.701	0.000	-	1.669	0.803
bank	5.239	3.984	-19.420	2.170	-13.097	1.765	-74.898	2.800
vol_bank	8.872	3.984	4.674	2.1698	22.712	1.7645	60.758	2.7997
employment								
unempl	-4.695	-2.405	-1.671	-0.846	9.854	2.451	1.837	0.680
empl	-0.102	-0.277	-18.063	-1.537	-5.446	-2.158	-86.083	-3.497
prices and costs								
infl	-0.294	-2.125	9.103	1.941	-7.339	-0.716	28.277	3.063
ind_price	1.473	2.592	-8.239	-2.121	4.904	1.248	65.388	2.393
labour	1.998	2.528	-0.129	-0.046	-5.637	-1.459	-19.117	-2.752
money, debt, and credit								
m3	2.498	0.828	-13.156	-1.839	-1.661	-1.582	-5.258	-0.453
loan_priv	4.453	2.957	-14.345	-2.860	6.418	1.034	0.123	0.135
loan_gov	-1.828	-2.898	27.421	0.923	30.913	1.274	33.874	1.639
cr_priv	6.828	5.894	0.345	0.750	1.767	0.701	1.727	1.132
cr_gov	3.841	3.500	22.484	4.456	-45.737	-1.146	2.244	0.933
deficit	-14.261	-8.946	-12.206	-1.506	-2.063	-1.089	-0.816	-1.717
output								
gdp	-0.390	-0.512	-23.965	-2.867	-14.566	-1.513	-153.630	-3.613
cons	-5.626	-1.804	14.129	1.465	-15.722	-1.324	0.428	0.314
gov_cons	-1.413	-1.886	-3.820	-0.859	-11.104	-0.708	26.681	0.824
inv	-3.409	-1.303	-6.698	-0.758	-38.646	-2.446	-19.369	-1.858
invent_gdp	3.953	1.419	-14.670	-0.421	72.742	2.313	4.454	0.269
ex	3.608	3.632	-1.263	-1.821	9.532	2.465	42.836	2.106
im	-3.469	-2.985	-9.166	-1.703	-107.656	-17.585	-15.981	-1.714
ind_prod	-0.750	-1.920	-18.461	-4.102	7.042	1.383	-34.600	-4.227

This table presents summary statistics for the daily cross-sectional level-effect LASSO coefficient estimates conditional on pre-crisis (Regime 0) and crisis regimes (Regime 1-2-3). Pre-crisis regime is arbitrarily set from 5/11/2009 to 3/31/2010, whereas crisis regimes are identified based on the Partitioning Around Medoids (PAM) clustering algorithm (Kaufman and Rousseeuw (1990)) executed on the time-varying parameters from equation (14). For each regime, the table reports the arithmetic average (Mean) and the corresponding t–stat computed with Newey-West robust standard errors (non-parametric kernel).

Table 7: Slope-effect LASSO Coefficients within Regimes

	Regime 0			Regime 1			Regime 2			Regime 3		
	Mean	t-stat	Slope	Mean	t-stat	Slope	Mean	t-stat	Slope	Mean	t-stat	Slope
bank, GIPSI, and tau												
gipsi	6.039	4.039	steep	0.393	1.372	steep	4.717	1.519	steep	-3.939	-1.842	flat
bank	0.503	1.727	steep	0.117	0.656	flat	0.000	-	-	0.000	-	
vol_bank	0.387	3.164	steep	0.274	0.614	steep	0.184	0.887	steep	-10.411	-2.956	flat
tau	0.334	3.172	steep	0.747	1.275	steep	1.650	2.111	steep	-1.245	-1.567	flat
employment												
unempl	0.000	-	-	-0.215	-2.364	steep	0.106	0.977	steep	-0.541	-1.019	flat
empl	1.994	2.663	flat	1.352	2.521	flat	2.127	2.589	flat	5.857	2.454	flat
prices and costs												
infl	0.003	1.007	flat	3.561	3.885	steep	0.000	-	-	-1.022	-0.601	flat
ind_price	-0.074	-1.094	flat	0.269	2.083	flat	-0.833	-1.205	flat	-7.760	-3.177	flat
labour	0.044	2.128	steep	0.812	0.869	flat	2.665	3.225	flat	1.179	1.528	flat
money, debt, and credit												
m3	0.390	4.250	steep	4.015	1.810	flat	-0.018	-1.448	steep	6.317	1.163	flat
loan_priv	0.227	1.402	steep	0.073	1.054	flat	-0.225	-1.405	flat	0.997	1.231	steep
loan_gov	-0.580	-1.441	steep	-7.956	-2.125	flat	-9.916	-1.401	flat	-10.522	-1.798	flat
cr_priv	0.153	0.874	steep	0.977	2.988	steep	-1.837	-1.770	flat	0.203	1.909	steep
cr_gov	0.011	1.345	steep	0.186	0.149	steep	15.690	6.899	flat	4.579	1.370	steep
deficit	-0.006	-2.039	steep	0.404	1.533	flat	3.208	0.806	flat	7.053	2.574	flat
output												
gdp	-0.050	-0.996	steep	0.168	1.140	flat	0.077	1.328	flat	27.419	2.304	flat
cons	-0.034	-1.641	steep	-0.242	-1.254	flat	4.059	2.241	flat	-5.398	-0.895	flat
gov_cons	0.227	2.292	flat	-0.641	-2.385	steep	2.858	1.076	flat	-1.983	-1.591	flat
inv	-0.087	-1.630	steep	2.459	0.851	flat	0.343	1.578	flat	1.634	2.055	flat
invent_gdp	0.617	1.814	steep	2.343	1.093	flat	-12.546	-3.356	flat	-1.868	-1.946	flat
ex	-0.358	-1.231	flat	1.129	0.756	flat	13.847	5.874	steep	-2.695	-2.639	flat
im	-2.296	-4.348	steep	-0.177	-1.339	steep	2.069	2.674	flat	9.259	3.252	flat
ind_prod	-0.822	-3.286	steep	0.480	1.142	flat	2.329	1.419	steep	6.236	2.803	flat

This table presents summary statistics for the daily cross-sectional slope-effect LASSO coefficient estimates conditional on pre-crisis (Regime 0) and crisis regimes (Regime 1-2-3). Pre-crisis regime is arbitrarily set from 5/11/2009 to 3/31/2010, whereas crisis regimes are identified based on the Partitioning Around Medoids (PAM) clustering algorithm (Kaufman and Rousseeuw (1990)) executed on the time-varying parameters from equation (14). For each regime, the table reports the arithmetic average (Mean) and the corresponding t–stat computed with Newey-West robust standard errors (non-parametric kernel). Column Slope denotes the flattening (flat) or steepeing (step) oriented term structure effect for each variable.

Table 8: Covariance Decomposition

Regime 0			**Regime 1**			**Regime 2**			**Regime 3**		
Var.	Exp.	Imp.	Var.	Exp.	Imp.	Var.	Exp.	Imp.	Var.	Exp.	Imp.
deficit	0.23	100.00	gipsi	0.21	100.00	im	0.34	100.00	gdp	0.30	100.00
gipsi	0.16	69.41	loan gov	0.10	50.18	invent gdp	0.24	70.84	empl	0.14	47.51
vol bank	0.14	58.55	gdp	0.09	43.19	cr gov	0.09	27.70	ind price	0.13	41.71
bank	0.13	57.71	cr gov	0.08	37.52	inv	0.09	25.81	bank	0.11	36.99
gipsi tau	0.09	40.68	empl	0.08	36.75	vol bank	0.06	17.48	vol bank	0.11	36.49
export	0.04	19.27	deficit	0.07	35.85	loan gov	0.05	15.99	gov cons	0.07	22.13
inv	0.04	18.57	invent gdp	0.07	34.18	cons	0.05	15.39	export	0.06	20.98
cons	0.04	17.74	ind prod	0.07	33.39	gdp	0.04	11.88	loan gov	0.05	17.47
cr gov	0.04	16.06	bank	0.07	32.07	infl	0.04	10.63	ind prod	0.04	13.50
invent gdp	0.03	11.83	loan priv	0.05	23.35	ind price	0.02	5.91	invent gdp	0.03	9.49

This table shows the average value of the cross-sectional CDS spread variance explained by each variable (column Expl.) according to equations (17), for level-effect coefficients, and (18), for slope-effect coefficients, in pre-crisis (Regime 0) and crisis regimes (Regime 1-2-3). The values are normalized by the highest value as $\frac{v_{kt}}{v_{kth}} \times 100$, where v_k is the value of the cross-sectional CDS variance explained by variable k at time t and v_{kth} is highest value of the explained variance out of all $K+1$ covariates (column Imp.). The values are reported in descending order according to the explanatory power by each variable. For each regime, only the first 10 variables are reported in the table.

Table 9: Average Cross-sectional R^2-s

Regime				Overall
0	**1**	**2**	**3**	
0.9771	0.9837	0.9897	0.9906	0.9850

This table presents the average cross-sectional R-squared (explained variance) of the model (equation (14)) in the different regimes.

Table 10: VSTOXX Out-of-sample Predictability

	RMSE	**MAPE**	**D-M test**
Lasso on Lasso	6.20	27.70	-2.994(0.003)
Lasso on Country PC Lasso FM	3.56	15.54	-
OLS on Country PC Lasso FM	3.82	16.51	-2.716(0.007)
PC on Lasso coeff	8.87	44.94	-8.974(0.000)
CDS 5yr	5.86	27.65	-5.982(0.000)
PC CDS GIPSI and NON-GIPSI	6.93	36.09	-8.917(0.000)
GIPSI and NON-GIPSI PC	6.29	30.35	-7.396(0.000)

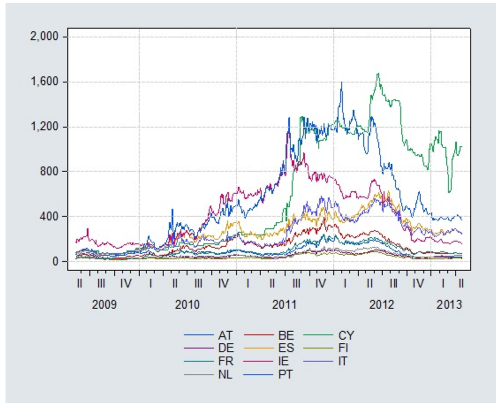
This table reports diagnostics of the out-of-sample forecasting ability of the LASSO coefficients for the one-month Euro Stoxx 50 index (VSTOXX) implied volatility index, three months ahead. Every day, we add one-t-ahead observation to the previous fit period t^{in} and we use the new estimation period to update the model estimates. Next, the new estimates are used to predict 3-month ahead. Mathematically we then have: $t^{in} = 1, \ldots, T_j^{in}$ and $t^{out} = T_j^{in} + 1, \ldots, T_j$. t^{in} is from 5/11/2009 to 7/31/2012, t^{out} is from 8/1/2012 to 4/25/2013, and predictions 3-month ahead $pred^{out}$ are from 10/24/2012 to 4/25/2013. The out-of-sample diagnostics computed using $pred^{out}$ are the Root Mean Squared Errors (RMSE), the Mean Absolute Squared Errors (MAPE) and the Diebold-Mariano (D-M) test. The table reports diagnostics for the following list of alternative models: (i.) LASSO-on-LASSO-type coefficients (Lasso on Lasso); (ii.) LASSO-on-LASSO-type Fama-MacBeth on country-specific principal components (Lasso on Country PC Lasso FM); (iii.) OLS-on-LASSO-type Fama-MacBeth on country-specific principal components (OLS on Country PC Lasso FM); (iv.) LASSO on principal components computed on LASSO-type coefficients (PC on Lasso coeff); (v.) OLS-on-Sovereign CDS 5yrs (CDS 5yr); (vi.) OLS on GIPSI and NON-GIPSI sovereign CDS principal components (PC CDS GIPSI and NON-GIPSI); (vii.) OLS on GIPSI and NON-GIPSI principal components extracted from the real-time macroeconomic data (GIPSI and NON-GIPSI PC). Principal components are estimated dynamically as in Beber et al. (2015) within the following clusters of macro-variables (see table 1): labour market; prices and costs; money and credit; output; banking.

Table 11: Lasso on Country PC Lasso FM, Level- and Slope-type Weigths

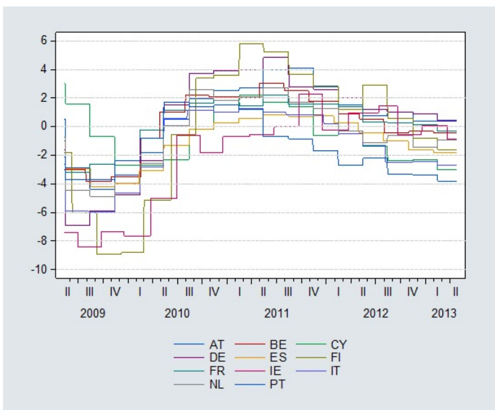
	Mean	**Sum**	**Weight**
Alpha	0.77	*0.77*	**0.10**
Level-type			
gipsi	0.02		
PC labour	0.14		
PC price	0.97		
PC money	1.69		
PC output	0.55		
PC banking	1.17		
		4.54	**0.57**
Slope-type			
gipsi	1.08		
PC labour	0.35		
PC price	0.19		
PC money	0.16		
PC output	0.28		
PC banking	0.24		
tau	0.36		
		2.64	**0.33**

This table presents the time dynamics of the estimated coefficients for the LASSO-on-LASSO-type Fama-MacBeth on country-specific principal components (Lasso on Country PC Lasso FM), selected as the best performer in the out-of-sample exercise (see Table 10). The values in table are the absolute values of the average coefficient estimates used to make the predictions out-of-sample as well as the ratios of the absolute values of the individual coefficient estimates over the sums of the absolute values of all the coefficient estimates. Alpha is the intercept; gipsi is the LASSO coefficient on the dummy variable denoting the peripheral connotation; PC labour, PC price, PC money, PC output, PC banking are the LASSO coefficients on the first country-specific principal component extracted from the clusters of macro-variables (table 1) following Beber et al. (2015); tau is the LASSO coefficient on the CDS maturities.

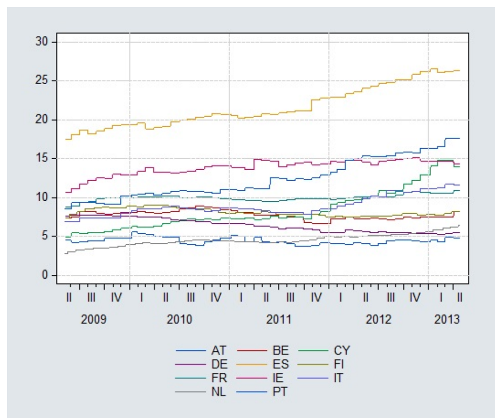
(a) Sovereign CDS spreads



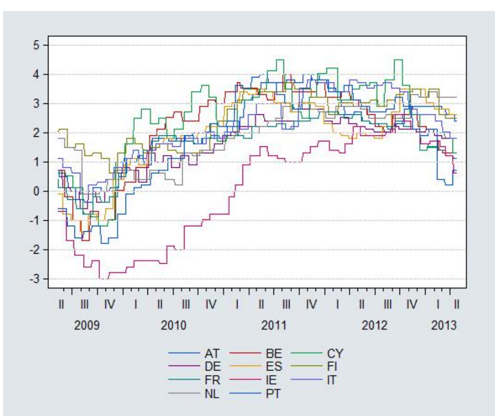
(b) GDP growth



(c) Unemployment rate



(d) Inflation rate



(e) Government deficit

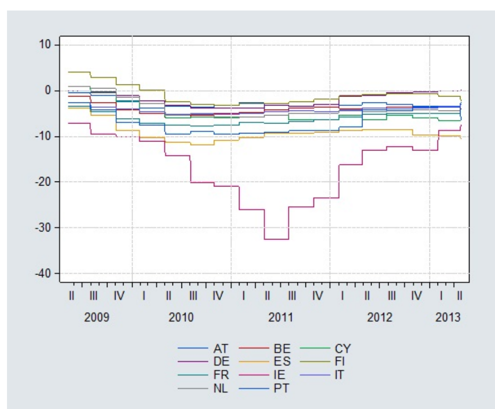


Figure 1: CDS spreads and macro fundamentals

The figure shows the 5-yr sovereign CDS (a), and the following key macroeconomic variables in real-time over the period from 5/11/2009 to 4/25/2013: GDP growth (b), unemployment rate (c), inflation rate (d), government deficit expressed as ratio over the GDP (e). AT is Austria, BE is Belgium, CY is Cyprus, DE is Germany, ES is Spain, FI is Finland, FR is France, IE is Ireland, IT is Italy, PT is Portugal.

Figure 2: Alphas and regimes

The figure displays the daily intercept estimation (Alpha) of the model (equation (14)) over the period from 5/11/2009 to 4/25/2013. Pre-crisis and crisis regimes are colored as grey (regime 0), pink (regime 1), yellow (regime 2) and light blue (Regime 3).

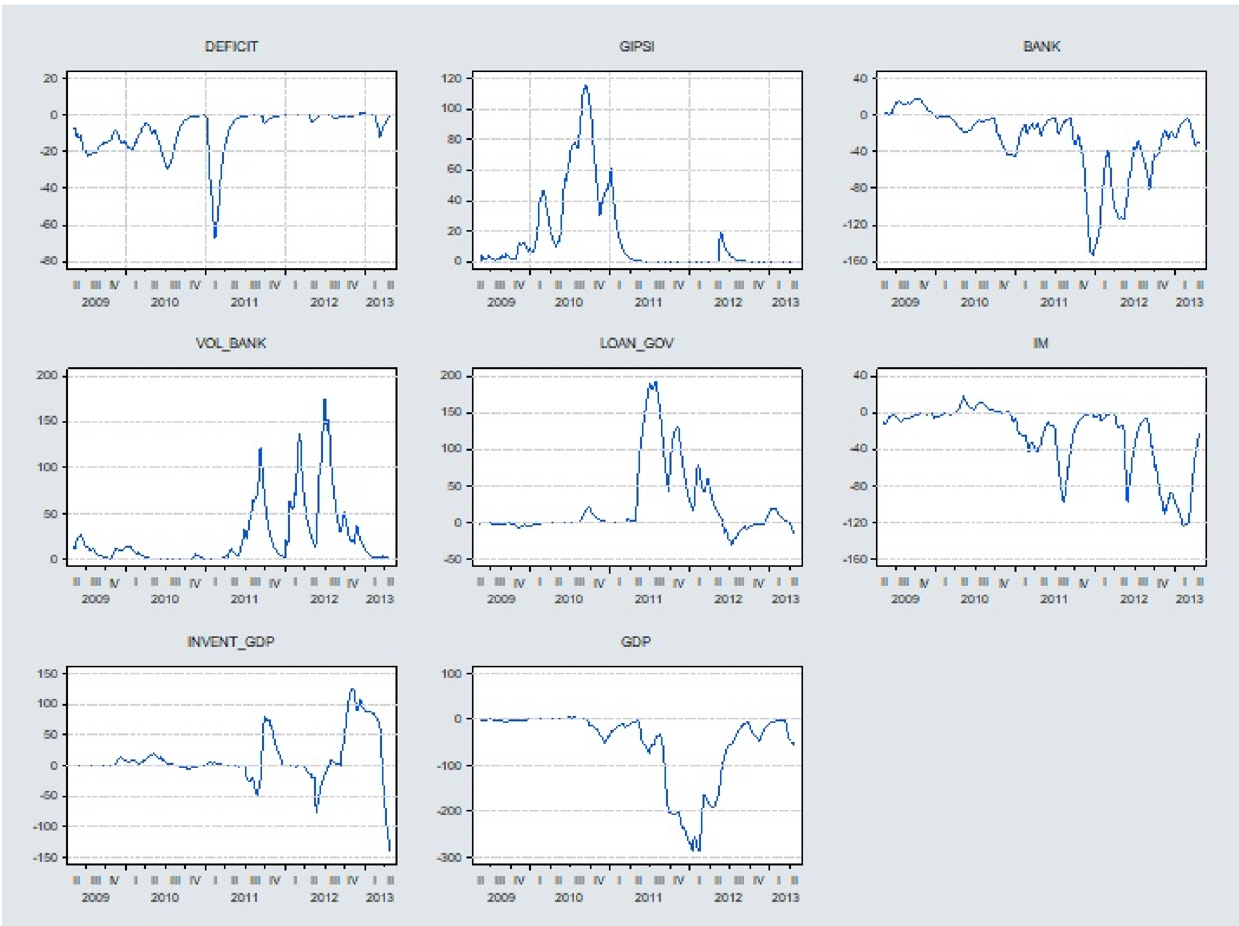


Figure 3: Most important coefficients

The figure shows the daily cross-sectional LASSO coefficients of the model (equation (14)) over the period from 5/11/2009 to 4/25/2013 for the most important covariates, as result of the covariance decomposition per regime (Table 8). The time patterns of the coefficients are recursively weighted with exponentially-decaying weights (the smoothing parameter is set at 0.95).

KJ-AE-19-003-EN-N

The European Commission's science and knowledge service

Joint Research Centre

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

EU Science Hub
ec.europa.eu/jrc

@EU_ScienceHub

EU Science Hub - Joint Research Centre

Joint Research Centre

EU Science Hub

Publications Office

doi:10.2760/761074
ISBN 978-92-76-00120-1