TrAffic LIght System for Systemic Stress: TALIS^3

Massimiliano Caporin
Department of Statistical Sciences
University of Padova

Laura Garcia-Jorcano
Instituto Complutense de Análisis Económico (ICAE)
Facultad de Ciencias Jurídicas y Sociales, Toledo
Universidad de Castilla-La Mancha

Juan-Angel Jiménez-Martin
Instituto Complutense de Análisis Económico (ICAE)
Facultad de Ciencias Económicas y Empresariales
Universidad Complutense de Madrid

Abstract

For the purposes of financial stability, it is important to identify financial institutions that, when in distress, could have a large adverse impact on financial markets. This paper proposes a TrAffic LIght System for Systemic Stress (TALIS-cube) that provides a comprehensive color-based classification for grouping companies according to both the stress reaction level of the system when the company is in distress and the company's level of stress. TALIS^3 has the ability to integrate multiple signals coming from the interaction between different risk metrics. In this paper, TALIS^3 builds on the standard VaR and the Conditional Value-at-Risk measure proposed by Girardi and Ergün (2013) extending this by introducing a Filtered Historical Simulation, nonetheless, TALIS^3 can easily be adapted to MES or SRISK. We classify companies by combining two loss functions, one for the system and one for each company, evaluated over time and as a cross-section. From the color-based classification of companies we also obtain an aggregated index. TALIS-cube can be used to enhance the performance and robustness of the current systemic risk measures. We provide an empirical analysis of the US market.

Keywords: (B) Risk Analysis, Systemic Risk, CoVaR, Company rankings, Aggregate index
JEL codes: G11, G21, G32, G38, C58

Acknowledgments. The authors wish to thank the Ministerio de Economía, Industria y Competitividad of Spain through the research project ECO2015-67305-P and the Banco de España through the research project PR71/15-20229. Juan-Angel Jimenez-Martin thanks the Department of Statistical Science for their hospitality during a research visit at the University of Padova and the Spanish Ministry of Education, Culture and Sport through the grant PRX17/00404. Massimiliano Caporin acknowledges financial support from the Project of Excellence 2018-2022 "Statistical methods and models for complex data" awarded by MIUR to the Department of Statistical Sciences of the University of Padova

TrAffic LIght System for Systemic Stress: TALIS^3

Abstract

For the purposes of financial stability, it is important to identify financial institutions that, when in distress, could have a large adverse impact on financial markets. This paper proposes a TrAffic LIght System for Systemic Stress (TALIS-cube) that provides a comprehensive color-based classification for grouping companies according to both the stress reaction level of the system when the company is in distress and the company's level of stress. TALIS^3 has the ability to integrate multiple signals coming from the interaction between different risk metrics. In this paper, TALIS^3 builds on the standard VaR and the Conditional Value-at-Risk measure proposed by Girardi and Ergün (2013) extending this by introducing a Filtered Historical Simulation, nonetheless, TALIS^3 can easily be adapted to MES or SRISK. We classify companies by combining two loss functions, one for the system and one for each company, evaluated over time and as a cross-section. From the color-based classification of companies we also obtain an aggregated index. TALIS-cube can be used to enhance the performance and robustness of the current systemic risk measures. We provide an empirical analysis of the US market.

Keywords: (B) Risk Analysis, Systemic Risk, CoVaR, Company rankings, Aggregate index
JEL codes: G11, G21, G32, G38, C58

1. Introduction

The recent financial crisis has fueled the search for precise measures of systemic risk. Several papers have proposed different systemic risk measures, such as Huang, Zhou and Zhu (2009), Segoviano and Goodhart (2009), Acharya, Pedersen, Philippon and Richardson (2010), Allen, Bali and Tang (2010), Zhou (2010), Brownless and Engle (2011), Benoit, Colliard, Hurlin, and (2017), Billio, Getmansky, Lo and Pelizzon (2012), Bisias, Flood, Lo and Valavanis (2012), Adrian and Brunnermeier (2016), and Silva, Kimura and Sobreiro (2017). Among the various approaches, two have attracted particular interest, these being the Marginal Expected Shortfall (MES) put forward by Acharya et al. (2010) and the Delta Conditional Value-at-Risk (ΔCoVaR) of Adrian and Brunnermeier (2016) (hereinafter AB). Recently, Girardi and Ergün (2013) (hereinafter GE) have produced a variation of the ΔCoVaR, based on a modification of the underlying risk measure, the CoVaR, i.e. the Value-at-Risk (VaR), of the financial system conditional on an institution being in financial distress (that is, at its Value-at-Risk level).

Adrian and Brunnermeier (2016) propose calculating $\Delta CoVaR$, as the difference between the CoVaR of the system when institution j is in financial distress and the CoVaR of the system when institution j is in a "normal" state, that is, when the company is in its median state. The financial distress of the conditioning institution corresponds to the institution being located exactly at its VaR. GE propose two changes. First, to measure the system's CoVaR allowing the financial institution to be at its VaR at most, thus taking into account the full density of the left tail of the conditional financial institution returns. Second, to define the normal state of the financial institution as one standard deviation about the mean. These modifications do not change the interpretation of $\Delta CoVaR$ as this systemic risk indicator still captures the marginal contribution of an institution to overall systemic risk. The higher the $\Delta CoVaR$, the larger the company's marginal contribution to systemic risk.

In this study we propose a new measure of systemic risk that enriches the ΔCoVaR proposed by GE. Even though Benoit et al. (2017) show that rankings of companies according to their contribution to systemic risk, as measured by ΔCoVaR estimated through quantile regression, might be equivalent to rankings using VaR in isolation, GE argue against that result and suggest a weak relationship between ΔCoVaR and VaR. Therefore, they suggest that monitoring only the company's tail risk might not be enough to forecast systemic risk. In the same line of reasoning we have Acharya et al. (2010). They state that an important part of the systemic risk of a company is its expected capital shortfall, an element that contains information about size, leverage and interconnectedness, and thereby provides a market-based measure of a firm's fragility. Through a panel regression analysis, GE also show that their ΔCoVaR contains information on the size, leverage and beta of the company. However, when they look at these variables, they find that they only account for 28% of the quarterly variation of ΔCoVaR.

These considerations lead us to design a new systemic risk measure that is an improvement on ΔCoVaR due to having additional information about a predictable dimension of risk included in companies' stock prices. Specifically, we consider the company's capital shortfall, (i.e. realized losses larger than the company VaR levels). Larger company financial losses imply larger expected amounts of capital that would need to be provided by the government in the event of a bail out. Furthermore, distress in a financial company would also have an effect on the services offered by that company. Consequently, we might have an impact on aggregate financial intermediation activity with knock-on consequences for the real economy. Therefore, the larger the capital shortfall, the larger the impact of the company's distress on the system and thus the more systemic the company. To examine this argument more explicitly, let us assume that two institutions, *j* and *i,* individually report the same $\Delta CoVaR$, they might even have the same *VaR*. Nonetheless, the capital shortfalls of the *j-th* company (losses larger than *VaR*) are systematically much larger than those of the *i-th* company. The *j-th* company is giving off a stronger signal of fragility than the *i-th* company. In the event of severe distress, the likelihood that the *j-th* company will change the level of its financial services offer is greater, and therefore it should be labelled as a larger contributor to systemic risk than the *i-th* institution.

The **TrAffic LIght System for Systemic Stress (TALIS^3)** proposed in this paper deals with the above issue. TALIS^3 makes use of loss functions frequently used in backtesting to provide a more accurate systemic risk ranking for financial companies. TALIS^3, which classifies financial institutions using a color code that will group companies into four categories according to the joint behavior of the company and the system exceedance magnitudes, will allow the monetary authorities to identify those financial institutions that pose a threat to financial stability.

TALIS^3 is an empirical methodology rank financial companies based on CoVaR calculations. Therefore, computing TALIS^3 requires the precise estimation of the relationship between the system and the company. GE report time varying in-sample CoVaR by estimating a bivariate GARCH model. In detail, they follow Engle's (2002) DCC specification fit on the returns of the institution and of the financial system to identify the volatility and correlation patterns. However, Mainik and Schaanning (2014) show that the dependence parameter between institutions and the system plays a key role in the accurate measurement of systemic risk under the ΔCoVaR. In particular, as the institution and the system become more highly correlated, the institution's systemic risk increases. This fact makes accurate measurements of the relationship extraordinarily important when analyzing financial distress.

Therefore, in this paper we extend GE's CoVaR computation using Filtered Historical Simulation and three different specifications for the evolution of conditional covariance. In addition to the DCC (Engle, 2002) with a GJR-GARCH (Glosten, Jagannathan and Runkle, 1993) specification for the marginal conditional variances, we use two other specifications: the BEKK model of Engle and Kroner (1995) and the Orthogonal GARCH (OGARCH) model introduced by Alexander and Chibumba (1997). In particular, the OGARCH model was designed to overcome the well-known curse of dimensionality linked to the Multivariate GARCH (MGARCH) models, while maintaining simplicity in estimating and interpreting their outcomes. See Silvennoinen and Teräsvirta (2009) for a discussion of the optimal features of MGARCH models. In contrast, the BEKK model at a bivariate level can be estimated in its full representation, thus without resorting to the common restrictions needed to overcome the curse of dimensionality. Consequently, the model becomes quite flexible and might also be preferable to the DCC, see Caporin and McAleer (2008, 2012). The comparison between competing models is thus the first contribution of our work.

The second innovation we provide focuses on the distributional assumption for the errors of the multivariate GARCH model. GE adopted both multivariate normal and multivariate skewed-t distributions. However, the risk measurement might suffer from the shortcomings involved in assuming a parametric error distribution. In the late 90s, Filtered Historical Simulation (FHS) emerged as one of the tools for calculating risk measures without relying on a distributional assumption. FHS, introduced by Barone-Adesi, Bourgoin and Giannopoulos (1998), Barone-Adesi and Giannopoulos (1996) and Barone-Adesi, Giannopoulos and Vosper (1999), uses a combination of volatility models and past (in-sample) innovations to build the probability distribution of future returns. Combining the MGARCH specifications and FHS, we produce 1-day CoVaR forecasts using a multi-step modelling approach: (1) a GJR-GARCH model is used to fit volatility; (2) DCC, BEKK and OGARCH are used to fit correlations; (3) FHS is used to simulate paths of 1-day forecast returns. With FHS, the distributional assumptions of the errors are relaxed, although the market conditions are taken into account by using the conditional volatility and correlations estimated in steps one and two; and (4) from the joint empirical distribution of the simulated returns for the system and company, we empirically evaluate the empirical CoVaR and $\Delta CoVaR$.

Our third contribution is the ranking system we introduce, TALIS[3]. Our system builds on loss functions and leads to a new color-based systemic indicator. TALIS[3] provides an evaluation of each company's risk, on a time-varying basis and conditional on the most recent financial company returns. TALIS[3] first evaluates two loss functions for each company, one at the system level, set to the system CoVaR, and one at the company level, set to the squared deviations between the returns of the financial company's equity and the corresponding Value-at-Risk. The two loss functions lead to the measurement of loss magnitudes, conditional on the stress states of the market and the company. Finally, the magnitudes are compared with time-varying threshold levels to obtain a color-based classification of the companies that includes four possible states. Beside the company classification, we provide a further contribution, i.e. the construction of an aggregate index based on the single company classifications. Finally, we stress that, by construction, TALIS^3 provides a time-varying one-day ahead company rankings forecast along with a one-day-ahead time varying aggregate index of systemic risk.

Our work includes a detailed empirical analysis that makes use of the same 73 companies used in GE and Acharya et. al. (2010), although we adopt an extended sample that spans from January 2000 to January 2018. This leads us to select a reduced number of companies, 58, for the last part of the sample, from September 2008 until January 2018. With these data we estimate TALIS^3 and the aggregated index. The losses, which are the base elements of TALIS^3, show that the scale of company distress is larger after the demise of Lehman Brothers, in a way that is consistent with the empirical evidence of the Aggregated Systemic Risk measure of the Volatility Laboratory (V-Lab). If we consider company rankings based on loss magnitudes and compare them with the rankings provided by ΔCoVaR, we find important differences, in particular for the Insurance sector, suggesting that TALIS^3 provides some improvement in the identification of systemically important companies. When moving to TALIS^3 outcomes, we observe how our approach provides an intuitive way to identify systemically important companies, and how the risk level changes in a sensible way over time, showing a diffuse risk increase around crises (Lehman) and a risk decrease after stabilizing events (such as the famous "whatever it takes" speech by Mario Draghi). We also analyze the predictive power of the aggregated version of TALIS^3. TALIS^3 not only shows an interesting pattern, in particular, an important correlation with the VIX, but also, as the "fear" index does, aggregate TALIS^3 provides a remarkable ability to predict futures changes in stock market indices, especially during period of financial distress.

We accompany the empirical analyses with several robustness checks. A first group focuses on the role played by the building blocks for TALIS^3, i.e. the multivariate GARCH model adopted, different choices for the rule for setting the thresholds, and different sizes of the rolling window used to determine the magnitude of losses. A second set of checks deals with the comparison between TALIS^3 rankings and those provided by other systemic risk indicators.

Finally, we highlight that TALIS^3 is a signal processing technique for clustering companies that manages information (signals) contained in the equity price. TALIS^3 allows the regulator to monitor different dimensions of risk and see the relationship between signals coming both from the system and the company. In this sense, TALIS^3 permits to manage the signals in order to provide more information than the original CoVaR and VaR individually do. Therefore, the usefulness of TALIS[3] lies, not only in the fact that it provides a measure of systemic risk, but also it increases the ability to integrate multiple signals coming from the interaction between different risk metrics, namely VaR, ES, CoVaR, MES and SRISK.

The remainder of the paper is organized as follows. In Section 2 we describe CoVaR and ΔCoVaR and the way to calculate it. In Section 3 we introduce the new TALIS[3]. In Section 4 we provide several empirical analyses based on the US market. Section 5 includes various robustness checks analyzing the sensitivity of TALIS[3] to the underlying volatility model and its tuning parameters. Section 6 contains the conclusions.

2. Measuring Systemic Risk with ΔCoVaR under Multivariate GARCH

Conditional Value-at-Risk (CoVaR) is a generalization of Value-at-Risk (VaR) that allows for the measurement of the risk of a financial system (as proxied by a market index) accounting for the impact of both control variables and a single, potentially important, financial institution. The classical formulation of Value-at-Risk implicitly defines it as the *q*-quantile of the return on an asset (or portfolio, or index…). If we focus on a financial company *j*, whose asset returns are R_t^j, we have

$$\Pr\left(R_t^j \leq VaR_{q,t}^j\right) = q \tag{1}$$

where $VaR_{q,t}^j$ is the *q%* Value-at-Risk for company *j*. CoVaR extends the classical VaR definition by adding a conditioning event, and is designed to calculate the conditional quantile of the financial system when a single company faces distress. As we have already highlighted, Adrian and Brunnermeier (2016) introduced CoVaR, and several papers have generalized and extended their original contribution. In this paper we use the definition of CoVaR proposed by Girardi and Ergün (2013) as it represents the most appropriate choice given our modeling framework. Following Girardi and Ergün (2013), we condition the evaluation of the system Value-at-Risk on the event that institution *j* is *at most at* its VaR. This differs from the best known definition of CoVaR by Adrian and Brunnermeier (2011, 2016). In fact, the latter use the condition of institution *j* being *exactly* at its VaR.

Girardi and Ergün (2013) define $CoVaR_{q,t}^{s|j}$ as the *q%*-quantile of the conditional probability distribution of the system returns conditional on a financial institution being in distress:

$$\Pr\left(R_t^s \leq CoVaR_{q,t}^{s|j} \mid R_t^j \leq VaR_{q,t}^j\right) = q \ . \tag{2}$$

Note that we use the same reference level, *q*, for both the system and the company; this is not a requirement and we can easily allow for different levels for the VaR of the company and the CoVaR of the system.

Adrian and Brunnermeier (2016), followed by Girardi and Ergün (2013), propose directly monitoring not CoVaR but a different quantity, namely the marginal contribution of a single institution to systemic risk. They suggest focusing on ΔCoVaR, which is the difference between the Conditional VaR of the system when the institution is in *distress* and the Conditional VaR of the system when the institution is in a *normal state*. Although we are using the GE framework to compute CoVaR, we stick to AB's ΔCoVaR definition:

$$\Delta CoVaR_{q,t+1}^{s|j} = CoVaR_{q,t+1}^{s|j} - CoVaR_{q,t+1}^{s|b^j} \tag{3}$$

where $CoVaR_{q,t}^{s|b^j}$ is the VaR of the financial system conditional on the *benchmark* state of institution *j*, b^j, defined as a one-standard deviation about the mean event, i.e. when $\mu_t^j - \sigma_t^j \leq R_t^j \leq \mu_t^j + \sigma_t^j$, and where μ_t^j and σ_t^j are the conditional mean and standard deviation of institution *j*, respectively, evaluated with a suitable model. Notably, this corresponds to a variation of equation (2) with the conditioning event being given by the company being in a normal or a benchmark state.

Although the absolute change metric used by AB to compute ΔCoVaR might be imperfect, because the change can depend, in part, on the value of CoVaR in the benchmark, the GE metric for ΔCoVaR proposed as a solution to this problem of dependency on the absolute change in the value of the benchmark, using the standardization of the absolute change to the benchmark state, has its own quirks. In fact, there is no identification of what happens if the CoVaR in the benchmark state happens to be 0. This event is quite improbable, but in a large cross-sectional and time series analysis we cannot easily exclude it. Furthermore, as the CoVaR in the benchmark state becomes increasingly small, the percentage change might become quite large. Therefore, we might find that during calm periods, when the benchmark CoVaR is becoming less positive (close to zero in absolute terms), the ΔCoVaR of the company is becoming unrealistically large, not because it is becoming more systemic, but just because the benchmark CoVaR is moving closer to zero. Finally, as with other ratios, even if there is no relationship between the absolute change (the numerator of the ratio) and the benchmark CoVaR, GE's definition of ΔCoVaR creates a relationship between them. By using the AB version of the ΔCoVaR we also maintain in the systemic risk measure information about the size of the possible losses suffered by the market when the conditioning institution enters a distressed state.

In Girardi and Ergün (2013), the only difference between CoVaR and VaR is given by the conditioning set. Therefore, the methodologies for forecasting CoVaR are mostly drawn from the VaR apparatus. Girardi and Ergün (2013) calculate $CoVaR_{q,t}^{s|b^j}$ and $CoVaR_{q,t}^{s|j}$, using a bivariate GARCH-DCC model assuming Gaussian and Skew-T distributions for the innovations. This choice goes towards capturing the heteroscedasticity present in financial returns, thus overcoming a potential limitation of Adrian and Brunnermeier's approach, in particular for analyses run at the daily level.

A first contribution of this paper comes from the more flexible model specification we adopt for the bivariate systems including the financial system returns and one company's returns. We deviate from Girardi and Ergün (2013) in several aspects. We adopt a semi-parametric approach and combine a Multivariate GARCH model augmented with volatility asymmetry, borrowing from Glosten, Jagannathan and Runkle (1993), with the Filtered Historical Simulation approach of Barone-Adesi et al. (1998), Barone-Adesi and Giannopoulos (1996) and Barone-Adesi et al. (1999). We evaluate the performance of three competing models in recovering the conditional covariance matrices, namely the DCC model used in Girardi and Ergün's strategy, the BEKK model of Engle and Kroner (1995), and the Orthogonal GARCH (OGARCH) model introduced by Alexander and Chibumba (1997). These three models represent the most common specifications found in the empirical finance literature. Clearly, other choices would also be possible, but the comparison of competing models is not the main focus of this paper. In fact, the MGARCH modeling step is just a necessary prerequisite for the subsequent analyses that represent the most important part of this paper. Alternative model specifications are only present as a robustness check analysis. Given the estimated conditional covariances, we proceed to the orthogonalization of the returns and obtain variance standardized and decorrelated residuals that are in turn used to generate future paths of the variances and future distributions of company and system returns. In this step, we follow the Filtered Historical Simulation (FHS) methodology. With FHS, the distributional assumptions for the errors are relaxed compared to Girardi and Ergün's choice. From the joint simulated distribution of system and company returns we can easily identify, by numerical integration, the empirical CoVaR and $\Delta CoVaR$. Note that by using FHS we estimate a predictive version of the $\Delta CoVaR$ systemic risk measure.

Appendix A includes details of our implementation of the approach used by Girardi and Ergün (2013), including the Filtered Historical Simulation, and also describes how we obtain the 1-day ahead $CoVaR_{q,t}^{s|j}$, $CoVaR_{q,t}^{s|b^j}$ and $\Delta CoVaR_{q,t}^{s|j}$ measures.

Building on these elements, in the following section we propose our TrAffic LIght System for Systemic Stress (TALIS[3]). We construct TALIS[3] by making use of loss functions frequently used in backtesting to provide a more accurate systemic risk ranking for financial companies.

TALIS^3 classifies financial institutions using a color code that will group companies into four categories according to the joint behavior of the company and system exceedance magnitudes. Further, the company-specific rankings contribute to the derivation of a composite index of systemic risk.

3. TrAffic LIght System for Systemic Stress (TALIS^3)

This paper's main contribution is an approach for reliably assessing financial companies' relative contributions to systemic market stress as implied by the distress states of the companies. Our methodological approach, summarized in a TrAffic LIght System for Systemic Stress (TALIS^3 – TALIS-cube), provides a comprehensive color-based classification that groups companies according to both the level of stress reaction of the system when the company is in distress and the level of stress of the company. In particular, TALIS^3 produces a quantitative assessment of the relative contribution to systemic stress based on the magnitude of exceedances while the company and/or the system are in distress. TALIS^3 uses appropriate loss functions to quantify the level of the company and system's stress and the severity of company and system losses.

TALIS^3 can be used to enhance the performance and robustness of current systemic risk measures. We can focus, for instance, on ΔCoVaR, which is the difference between the VaR of the system conditional on the institution being in *distress* and the VaR of the system conditional on the institution being in a *benchmark* state. Now take two companies, A and B, with the same, or very similar, ΔCoVaR. This seems to suggest that the two companies have a similar systemic impact. However, the two companies might have completely different risk profiles, going beyond the risk as summarized by the company VaR. In fact, the exceedances of the two companies (i.e. realized losses larger than the company VaR levels) might be significantly different, with the distress of the companies having a consequently different impact on system stability. Such a case is partly accounted for by the ΔCoVaR redefinition of Girardi and Ergün (2013). We therefore believe that the evaluation of company systemic risk should not just be associated with a quantitative evaluation of the co-occurrence of distress, but should also take into account the size of past realized exceedances at the company level. Furthermore, neither the approach of Adrian and Brunnermeier (2016) nor that of Girardi and Ergün (2013) consider the (historical) association between company exceedances and system exceedances. However, when the two events coincide, the systemic impact of a company might be larger. TALIS^3 tries to overcome these ΔCoVaR's shortcomings by accounting for the magnitude of both company and system distress when evaluating and ranking companies. TALIS^3 is thus a distressed state oriented systemic risk measure or systemic risk ranking scale.

Although, TALIS[3] is based on actual and past distress experiences, and may only give a partial picture of stress states, we believe that it provides a powerful tool for monetary authorities for the proper identification of systemic companies.

To introduce TALIS[3], let R_t^s and R_t^j be the returns of the financial system and of the *j-th* financial company at time *t,* respectively, where $j = 1, \cdots, N$ and $t = 1, \cdots, T$. Let $VaR_{q,t}^s$ and $VaR_{q,t}^j$ denote the *q%-quantile* for R_t^s and R_t^j, respectively. Note that both Value-at-Risk levels derive from the estimation of a bivariate MGARCH model. Moreover, as we adopt a forecasting perspective, the Value-at-Risk at time *t*, comes from a model estimated on a sample ending at time *t-1*. From an empirical perspective, the evaluation of the Value-at-Risk over time and over a cross-section of companies would require the adoption of rolling methods and the estimation of a large number of bivariate models (N estimates for each rolling sample).

We first consider two indicator functions that compare the system's and the *j-th* company's daily returns with system and company $VaR_{q,t}$, respectively. We define the indicator function as:

$$I_{t+1}^i = \begin{cases} 1 \ if \ \ R_{t+1}^i \leq VaR_{q,t+1}^i \\ 0 \ if \ \ R_{t+1}^i > VaR_{q,t+1}^i \end{cases} \tag{4}$$

for *i= s, j,* which takes the value of 1 if the system (company) has experienced a Value-at-Risk violation on that day, and 0 otherwise. To compute our new systemic measure, we will examine the relationship between stress states in the company and the system. As a result, every stress day in the sample, i.e. days with exceedances for both the company and the system and those days with a violation only in either the system or the company, will add particular value to identifying such co-movement. The new measure will combine system and company stress states to provide a more accurate measure of the systemic risk posed by a financial institution.

In addition to the presence of exceedances, the magnitude of system and company distress also matters and this can be proxied, for example, by the size of the exceedances. This element only partially enters in the definition of the $\Delta CoVaR$ of Adrian and Brunnermeier (2016) and Girardi and Ergün (2013) − through the joint distribution of the company and system. However, the effective loss made by the company in the presence of an exceedance and the contemporaneous impact on the system are not taken into account when evaluating the systemic impact of a specific financial institution or when ranking financial institutions on the basis of their systemic impact.

TALIS[3] tries to overcome this limitation by taking into account the magnitude of the distress states of the companies, and separately evaluating the loss suffered at the system level and the loss suffered by the company. In the system case, we suggest monitoring the systemic impact of a company falling into a distressed state by sticking to the $\Delta CoVaR$ evaluated when the system is in distress, i.e. when $I_{t+1}^{s} = 1$. The loss function thus becomes:

$$L_{t+1}^{s} = \Delta CoVaR_{t+1}^{s|j} \tag{5}$$

This amended, or stressed, version of ΔCoVaR can be seen as the contribution to a severe market decline because of a severe capital shortfall in a given company.

For the company loss function, we propose to introduce in the TALIS[3] methodology a quadratic numerical score that evaluates the squared excess returns of the company when the company is in distress, i.e. when the company's returns exceed their *VaR*. The general form of the loss function for the *j*-th company is thus:

$$L_{t+1}^{j} = \left(R_{t+1}^{j} - VaR_{t+1}^{j}\right)^{2} \tag{6}$$

A similar quadratic loss function was used by Lopez (1999) and Caporin (2008) for the evaluation of competing Value-at-Risk models. We stress that the evaluation of the loss at time *t+1* requires the observed return at time *t+1*, and the forecast of the Value-at-Risk made for time *t+1* using an information set up to time *t*.

Starting from the two loss functions, we define the magnitude of the distressed state at the system level as the expected loss conditional on the systemic event that the system is in distress:

$$MD_{t+1}^{s} = E\left[L_{t+1}^{s} \middle| I_{t+1}^{s} = 1\right] \tag{7}$$

Similarly, the magnitude of the distressed state in the *j-th* institution can be defined as the expected loss in the company conditional on the systemic event that the company is in distress:

$$MD_{t+1}^{j} = E\left[L_{t+1}^{j} \middle| I_{t+1}^{j} = 1\right]. \tag{8}$$

TALIS[3] provides a ranking of institutions based on the two above defined magnitudes. For ease of interpretation, the ranking uses different colors depending on whether the company and the system are in a relatively low or in a high state of stress.

The identification of the riskiest companies, from a systemic point of view, might not simply be based on rankings of the above mentioned expected losses as we might find a case where, at every point in time, some of the companies are always in a high risk state. Therefore, we decided to identify high or low financial stress by contrasting the value taken by the distress magnitudes in (7) and (8) with respect to properly defined threshold values, which we denote by $\overline{z}_{t+1}^{i},\ for\ i = j, s,$ where *j* is company and *s* is system. Clearly, we have several possible ways to estimate a threshold from the collection of magnitudes available at a given point in time from the cross-section of financial companies, as well as by calibration according to an economic evaluation of the magnitudes. Let us postpone until later the discussion on how we should define these thresholds and assume that they are available.

Given the thresholds, when $MD_{t+1}^{j} > \overline{z}_{t+1}^{j}$ the *j-th* institution is facing losses whose magnitude is larger than the threshold, and the company is said to be in a relatively severe stress state. In contrast, when $MD_{t+1}^{j} \leq \overline{z}_{t+1}^{j}$ the *j-th* institution is facing losses whose magnitude is smaller than the threshold and the company is now in a relatively minor stress state.

We define the magnitude of the losses of the system in a similar way. When $MD_{t+1}^{s} > \overline{z}_{t+1}^{s}$ the system is facing losses of a magnitude larger than the threshold. Thus a severe impact of the *j*-th institution on the system corresponds to the observation of $MD_{t+1}^{s} > \overline{z}_{t+1}^{s}$. In contrast, a minor impact of the *j*-th company on the system occurs when $MD_{t+1}^{s} \leq \overline{z}_{t+1}^{s}$. We stress that neither $CoVaR_{q,t}^{s|j}$ nor $\Delta CoVaR_{q,t}^{s|j}$ capture such a relationship between the stress of the company and the impact of the company on the system.

Given these two possible cases, identified on the basis of the comparison between the expected magnitudes and the reference thresholds, TALIS[3] ranks companies according to the four different possible states we could observe at a given point in time, namely

$$TALIS_{j,t+1,m}^{3} = \begin{cases} RED & when\ MD_{t+1,m}^{j} > \overline{z}_{t+1,m}^{j}\ and\ MD_{t+1,m}^{s} > \overline{z}_{t+1,m}^{s} \\ AMBER & when\ MD_{t+1,m}^{j} \leq \overline{z}_{t+1,m}^{j}\ and\ MD_{t+1,m}^{s} > \overline{z}_{t+1,m}^{s} \\ YELLOW & when\ MD_{t+1,m}^{j} > \overline{z}_{t+1,m}^{j}\ and\ MD_{t+1,m}^{s} \leq \overline{z}_{t+1,m}^{s} \\ GREEN & when\ MD_{t+1,m}^{j} \leq \overline{z}_{t+1,m}^{j}\ and\ MD_{t+1,m}^{s} \leq \overline{z}_{t+1,m}^{s} \end{cases} \quad (9)$$

According to TALIS[3], companies may be labelled using the traffic light signals for high (RED), medium-high (AMBER), medium-low (YELLOW) and low (GREEN) risk. Companies currently in a relatively high stress state and with a severe impact on the system take a *red* label, which indicates the highest risk. *Red* companies should be under intense scrutiny by the monetary authorities and supervisors as they might easily lead to large losses with potentially large and significant consequences at the system level. A company with a *green* label, which indicates the lowest risk, is associated with a low stress state of the company accompanied by the company's minor impact on the system. *Amber* and *yellow* represent companies in a middle position. For instance, an *amber* company has a severe impact on the system even when it is in a minor stress state. Companies in a severe stress state but which do not have a particularly severe impact on the system are labelled *yellow*. As our focus is on the evaluation of the risk at the systemic level, *amber* is riskier than *yellow*.

The ranking provided by TALIS[3] produces four groups of companies, and companies within each group are perceived as characterized by a similar degree of systemic impact. Obviously, a proper ranking within each group could be obtained using additional company-specific information, such as the company size (proxied by market value or book value) or leverage, to mention just two well-known quantities.

From a different viewpoint, we must also clarify that the ranking provided by TALIS[3] depends on the sample used for its evaluation. This aspect impacts on both the evaluation of the expected losses and, as we will discuss below, on the choice of the threshold. Both elements introduce time-dependence in the rankings, making them consistent with the changing nature of financial markets and of the risk of both single companies and the entire financial system. In the first case, for the expected magnitudes, a simpler approach is the use of a rolling window of a given size (reasonable values range from 252 days, i.e. one year, to 120, 60, 20, so a half-year, quarter or month, respectively). A second and more general approach is to evaluate expected losses for the companies as

$$MD_{t,m}^{j} = \sum_{i=1}^{m} w(i,m) L_{t-i}^{j} \tag{10}$$

and to adopt an equivalent formula for the system. In this equation, *m* identifies the length of the window used for the computation of TALIS[3], and $w(i,m)$ is a generic weighting function. The latter might include the simple average as a special case, but might also include an exponential decay function giving more weight to the most recent information.

In addition, a full characterization of TALIS[3] requires the computation of a threshold. There are several ways of defining this threshold, and here we suggest an approach that depends on the distress state history of the single company (system). We define the thresholds as follows:

$$\overline{z}_{t+1,m}^{i} = Max\left\{ Median\left\{ MD_{t+1,k}^{i} \right\}_{k=1}^{k=t}, \quad Median\left\{ MD_{t+1,k}^{i} \right\}_{k=t-m}^{k=t} \right\} \tag{11}$$

where, $i = j, s$ where *j* is the financial institution and *s* is the system. The threshold is defined as the maximum of two arguments: the median of the full history of expected losses and the median of the last *m-day* window. We stress that the threshold is company-dependent, time-dependent and computed as a long-run or short-run average of the losses. The company (system) will then be in a stress state when the magnitudes of losses are larger than the long-run averages if in the short-run the losses are relatively small, while the company (system) will be in a high stress state if the current expected magnitude of losses is larger than the short-run averages if in a recent period the company experienced an increase in losses, and therefore in its risk level. Note that this is similar to the use of two moving averages as commonly adopted in trading for the detection of signals for opening/closing trades.

Finally, we propose deriving a more aggregate view of the overall financial system, which is a time-varying aggregate version of TALIS[3]. We thus suggest creating an aggregate $\Delta CoVaR$ -related index. The aggregation will make use of the loss associated with the system level and of company-specific aggregation weights based on the TALIS[3] ranking. Following a strategy similar to Aikman et al. (2017), we define the aggregate index as:

$$A\_TALIS^3_{t+1,m} = \left[\prod_{j=1}^{N} \left| \Delta CoVaR^j_{t+1} \right|^{w_{j,t+1,m}} \right]^{\mathcal{E}}$$
$$\mathcal{E} = \left(\sum_{j=1}^{N} w_{j,t+1,m} \right)^{-1} \tag{12}$$

where $w_{j,t+1,m}$ is the weight of company *j* in period *t* and $\Delta CoVaR^j_t$ is the contribution of company *j* to the distress of the system in period *t*. We note that ΔCoVaR is always negative, the lower it is (the larger the absolute value), the larger the marginal contribution to systemic risk. Therefore, in order the build the aggregated index, the absolute value of $\Delta CoVaR^j_t$ is used, making the aggregate index positive, where an increase signals an accumulation of systemic risk.

From TALIS[3], we assign a quantitative weight, $w_{j,t+1,m}$ to each distress state as follows:

$$w_{j,t+1,m} = \begin{cases} 1 \quad when\ MD^i_{t,m} > \overline{z}^j_{t+1,m}\ and\ MD^s_{t+1,m} > \overline{z}^s_{t+1,m} \\ 0.75\ \ when\ MD^j_{t+1,m} \le \overline{z}^j_{t+1,m}\ and\ MD^s_{t+1,m} > \overline{z}^s_{t+1,m} \\ 0.50\ \ when\ MD^j_{t+1,m} > \overline{z}^j_{t+1,m}\ and\ MD^s_{t+1,m} \le \overline{z}^s_{t+1,m} \\ 0.25\ \ when\ MD^j_{t+1,m} \le \overline{z}^j_{t+1,m}\ and\ MD^s_{t+1,m} \le \overline{z}^s_{t+1,m} \end{cases} \tag{13}$$

Note that we calibrate weights to induce a larger role for companies in a riskier state. Furthermore, we compute the aggregate index using a weighted geometric mean instead of the weighted arithmetic mean because the former is robust to outliers and large differences between company $\Delta CoVaR$.

4. TALIS[3] for the US market

4.1. Data description

We work with daily log-returns for 73 US financial institutions in four industry groups, mimicking Girardi and Ergün (2013). Table 1 lists the financial institutions and their type based on two-digit SIC classification codes (Depositories Institutions, Securities Dealers and Commodity Brokers, Insurance, and Others). The Dow Jones US Financial Index (DJUSFN) is used as a proxy for the financial system, as in GE. We recovered all the data, company prices and the index level from Thomson Reuters Datastream. We extend GE's sample from January 3, 2000 through January 15, 2018 (4,706 trading days per company). Only 58 (in bold in Table 1) out of 73 companies remain for the full sample, as 15 companies (in red in Table 1) were delisted after September 15, 2008 (Lehman Brothers bankruptcy). This leaves us 73 companies with 2,271 observations until September 15, 2008 and 58 companies with 4,706 observations in the full sample. We will adopt a rolling window approach for model estimation, with a window size of 1,000 observations. This produces 3,706 (for 58 financial institutions) or 1,271 (for 15 financial institutions) one-day-ahead forecasts from November 3, 2003 up to the end of the sample. Table 2 reports the descriptive statistics for the daily returns of the financial system and the 73 and 58 financial institutions by industry group. All industry groups have mean and median returns close to zero. Others and Brokers-dealers are more volatile than the Depositories and Insurance groups. DJUSFN (financial system) and all industry groups have significant negative skewness, except Brokers-dealers which has positive skewness. The kurtosis is high for all industry groups and for the financial system, implying that the return distributions have much thicker tails than the Gaussian distribution. Kurtosis is especially large in Insurance.

4.2. Forecasting Results

Similar $\Delta CoVaR_{5\%}$ forecasts are produced using three MGARCH specifications, namely DCC, BEKK and OGARCH. Slight differences are found in the TALIS[3] company rankings and this will be analyzed in Section 4.4. To make the analysis clearer, sections 4.2 and 4.3 will only show the results when the DCC model specification is used to calculate the $\Delta CoVaR_{5\%}$ forecasts. In the evaluation of the systemic risk measures, we should recall that after the model estimation, conditional variances and covariances are used for computing orthogonal residuals which, in turn, represent the input for the generation of simulated innovations in the FHS approach. We generate 10,000 returns for each company and for the market index for each forecasting step. Using the FHS generated future returns, we determine the predicted quantile of the company along with the conditional quantile of the system.

Table 1.- Names and classifications of 73 US financial institutions.
58 US financial institutions: January 3rd, 2000 - January 15th, 2018 (4706 obs.)
15 US financial institutions: January 3rd 2000 – September 15th, 2008 (2271obs.)

28	Depositories	15	Others	23	Insurance	8	Broker-dealers
BAC	Bank of America	**ACAS**	American capital Strategies	**AFL**	AFLA Inc	**BSC**	Bear Stearns
BBT	BB & T	**AMTD**	Ameritrade Holding	**AIG**	American International Group	**ETFC**	E-Trade Financial
BK	Bank of New York Mellon	**AXP**	American Express	**ALL**	Allstate Corp	**GS**	Goldman Sachs
C	Citigroup	**BEN**	Franklin Resources Inc	**AON**	AON Corp	**LEH**	Lehman Brothers
CBH	Commerce Bancorp Inc NJ	**BLK**	Blackrock Inc	**BRKA**	Berkshire Hathaway Inc Del	**MER**	Merrill Lynch
CMA	Comerica Inc	**COF**	Capital One Financial	**BRKB**	Berkshire Hathaway Inc Del	**MS**	Morgan Stanley
HBAN	Huntington Bancshares Inc	**EV**	Eaton Vance Corp	**CB**	Chubb Corp	**SCHW**	Schwab Charles Corp
HCBK	Hudson City Bancorp Inc	**FNM**	Federal National Mortgage Assn	**CFC**	Countrywide Financial Corp	**TROW**	T Rowe Price
JPM	JP Morgan Chase	**FRE**	Federal Home Loan Mortgage	**CINF**	Cincinnati Financial Corp		
KEY	Keycorp New	**JNS**	Janus Cap Group Inc	**CNA**	Can Financial Corp		
MI	Marshall Isley	**LM**	Legg Mason Inc	**HIG**	Hartford Financial Svcs Group		
MTB	M & T Bank Corp	**LUK**	Leucadia national	**HUM**	Humana Inc		
NCC	National City Corp	**SEIC**	Sei Investments Company	**L**	Loews Corp		
NTRS	Northern Trust Corp	**SLM**	S L M Corp	**LNC**	Lincoln national Corp		
NYB	New York Community Bankcorp	**UNP**	Union Pacific	**MBI**	MBIA Inc		
PBCT	People United Financial			**MET**	Metlife Inc		
PNC	PNC Financial Services			**MMC**	Marsh and Mclennan Cos Inc		
RF	Regions Financials			**PGR**	Progressive Corp OH		
SNV	Synovus Financial Corp			**SAF**	Safeco Corp		
SOV	Sovereign Bancorp			**TMK**	Torchmark Corp		
STI	Suntrust Banks Inc			**TRV**	Travelers Companies Inc		
STT	State Street Corp			**UNH**	United Health Group		
USB	US Bancorp Del			**UNM**	Unum Group		
WB	Wachovia						
WFC	Wells Fargo						
WM	Washington Mutual						
ZION	Zions Bancorp						

Table 2.- Summary statistics for the percentage log-returns of the financial system before Lehman Brothers bankruptcy (73 US financial institutions) and after Lehman Brothers bankruptcy (58 US financial institutions) by industry group.

Sample 01/03/2000- 09/15/2008	DJUSFN	Depositories	Others	Insurances	Brokers-Dealers
Mean	-0.002	0.004	-0.001	0.028	-0.031
Median	0.000	0.000	0.000	0.000	0.000
St. Dev.	1.630	2.042	2.269	1.909	2.824
Skewness	0.213	0.126	-0.047	0.115	-0.016
Kurtosis	9.015	12.183	8.195	11.238	7.939
Sample 09/16/2008- 01/15/2018	**DJUSFN**	**Depositories**	**Others**	**Insurances**	**Brokers-Dealers**
Mean	0.041	0.021	0.043	0.038	0.033
Median	0.033	0.000	0.000	0.016	0.000
St. Dev.	1.971	2.749	2.521	2.023	2.352
Skewness	-0.445	-0.310	-0.307	-0.457	0.162
Kurtosis	20.436	22.203	17.284	30.005	22.041

Table 3 reports the summary statistics for $\Delta CoVaR_{5\%}$ forecasts of all the 73 institutions until the Lehman Brothers bankruptcy and for the other 58 institutions that remain in the sample after the demise of LEH. We report the summary by industrial group. In keeping with GE we compute two variants of the standard deviation. In the first, Std TS, we compute the standard deviation of the $\Delta CoVaR_{q,t}^{s|j}$ for an institution *j* and then average these standard deviations across institutions. In the second, Std CS, we compute the standard deviation of the mean of each individual institution' $\Delta CoVaR_{q,t}^{s|j}$ measures, and then compute the average.

Table 3.- Summary statistics for $\Delta CoVaR_{5\%,t}^{s|j}$ for all 73 US financial institutions (before Lehman Brothers bankruptcy) for all 58 US financial institutions (after Lehman Brothers bankruptcy) by industry group.

Sample 01/03/2000-09/15/2008	Mean	Std TS	Std CS	Max	Min
Overall	-1.97	1.37	0.04	-0.25	-8.31
Depositories	-2.12	1.49	0.04	-0.25	-8.93
Others	-1.85	1.27	0.04	-0.19	-7.82
Insurance	-1.70	1.25	0.03	-0.16	-7.43
Brokers-Dealers	-2.48	1.53	0.04	-0.62	-9.68
Sample 09/16/2008 -01/15/2018	**Mean**	**Std TS**	**Std CS**	**Max**	**Min**
Overall	-3.21	2.76	0.06	-0.22	-21.13
Depositories	-3.32	2.84	0.06	-0.25	-21.82
Others	-3.13	2.63	0.05	-0.14	-19.78
Insurance	-3.04	2.66	0.05	-0.15	-20.37
Brokers-Dealers	-3.64	3.15	0.06	-0.65	-24.49

Table 3 shows the results for the first part of the sample, before the Lehman Brothers bankruptcy, and for the second part of the sample starting on September 16, 2008. As reported by GE, Brokers-Dealers and Depositories show the largest $\Delta CoVaR_{5\%}$ and therefore they were the largest contributors to systemic risk during both periods. The insurance sector shows the lowest $\Delta CoVaR_{5\%}$. Clearly, results for the different sub-periods of this analysis might deviate from the ones we report. As defined in Section 3, TALIS[3] ranks companies based on two magnitudes. First, the degree of stress of the system $MD^{s}_{t+1,60}$ defined in (7) and computed as the 60-day average of the absolute value of the $\Delta CoVaR_{5\%}$ whose average values are shown above. Second, the degree of stress of the company $MD^{j}_{t+1,60}$ defined in (8) and computed as the 60-day average of the loss function defined in (6), i.e. the squared excess return of the company when the company is in stress. So far there is no theoretical guidance to help select the optimal window length, *m*, in the rolling computation of *MD*. TALIS[3] can be understood as a novel signal processing technique that uses equity prices **for extracting suitable features** for both visual and automatic classification of financial companies, according to their marginal contribution to the systemic risk. TALIS[3] monitors the evolution of company and system's MD. When one of these two indicators deviates from the "normal" behavior, a signal is released. With the aim of producing readily interpretable signals about the marginal contribution of a company to the market risk, we first choose *m =60* (three months) and we will later assess the sensitivity of TALIS[3] rankings to the choice of *m*. A longer rolling window size, for example *m=120* (half year), tends to yield a weaker signal, and a shorter size, *m=20* (one month) will produce a signal contaminated by noise. The 60-day average is not written or described in any regulation, but it seems a general practice in Basel recommendations. Our suggested approach to compute *MD* involving a window comprised of the last 60 days of operation is the same that the one used for calculating the capital charge for general market risk suggested by the Basel Committee on Banking Supervision (2019) (BCBS). BCBS sets the capital charge for general market risk to the maximum of the most recent observation of Expected Shortfall and the average estimate over the previous sixty trading days (approximately one quarter of the trading year) multiplied by a "scaling factor," which is generally equal to three. The 60-day cut-off period is also used to define what can be considered as proprietary trading in bank activity. Recently, the US regulators relaxed the Volker's 60-day rule announcing that positions that banks hold will automatically be counted as compliant with the proprietary trading ban if they are held for longer than 60 days.

Table 4 shows the summary statistics for $MD^{j}_{t+1,60}$ of the 73 companies before the Lehman Brothers bankruptcy and the statistics for the 58 companies remaining in the sample after Lehman's failure. The first column provides information about the average magnitude of the distress in the different sectors. As in the case of $\Delta CoVaR_{5\%}$, the average magnitude of company distress is larger after Lehman's demise. This result is in line with other systemic risk measures such as the Aggregate Systemic Risk (SRISK) time series provided by the Volatility Laboratory (V-Lab) webpage created by Engle, Capellini, Reis and Ruan that shows similar behavior, that on average SRISK increased after the second half of 2007 to the end of our sample, as can be seen in Figure 1.

Figure 1.- Systemic risk measure evaluated from 2004 to 2018. (Source: V-lab (2018), Risk Analysis overview, SRISK, https://vlab.stern.nyu.edu/welcome/risk/).

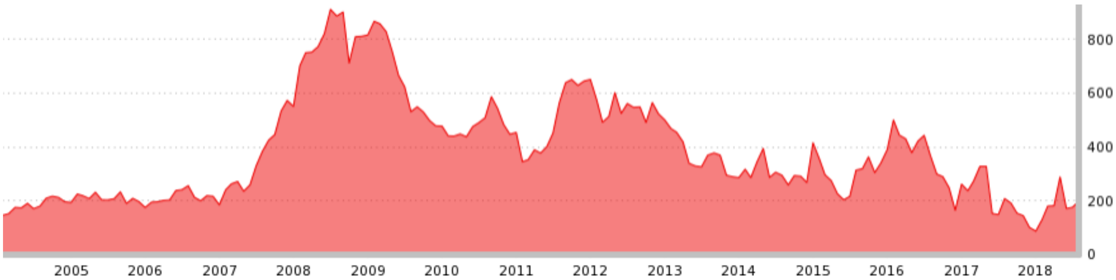


Table 4.- Summary statistics for the company expected loss for all 73 US financial institutions (before Lehman Brothers bankruptcy) and for all 58 US financial institutions (after Lehman Brothers bankruptcy) by industry group.

Sample 01/03/2000- 09/15/2008	Mean	Std TS	Std CS	Max	Min
Overall	4.0	10.4	0.3	65.9	0
Depositories	2.6	5.6	0.2	39.6	0
Others	5.8	16.3	0.5	101.4	0
Insurance	4.9	13.8	0.4	86.2	0
Brokers-Dealers	2.8	5.1	0.1	29.8	0
Sample 09/16/2008 -01/15/2018	**Mean**	**Std TS**	**Std CS**	**Max**	**Min**
Overall	5.6	12.0	0.2	85.6	0.0
Depositories	5.0	10.9	0.2	74.6	0.0
Others	8.2	15.1	0.3	93.3	0.0
Insurance	4.6	10.8	0.2	87.7	0.0
Brokers-Dealers	5.9	13.5	0.3	105.2	0.0

This might signal that the overall financial system has become more aware of the possible consequences of the risk implicit in financial companies and of the possible system impact that they could have if they were to enter into a distressed state. If we take a look at sector rankings, we first note that "Other" financial institutions show the largest expected losses during both periods. However, if we consider rankings according to the ΔCoVaR metric, this is the third systemic sector. We thus observe that sectors that rank in the top positions according to their contribution to systemic risk based on ΔCoVaR are not necessarily those with the largest magnitude of distress. For example, before Lehman's bankruptcy, the insurance sector ranked as the least systemic, but in second place because of the size of the capital shortfall. It seems that ΔCoVaR did not completely capture the risk that the system was facing because of the insurance sector's troubles (AIG played a significant role in the great financial crisis) reflected by the capital shortfall indicator. Therefore, there is room to introduce new information to improve the systemic risk measures. The use by TALIS^3 of two magnitudes when evaluating risk could lead to a better evaluation of systemic risk rankings since the size of the company's distress might contain important information for measuring the impact of companies on the market. In the following sections we carry out further analyses, focusing on single companies and on an aggregated systemic index.

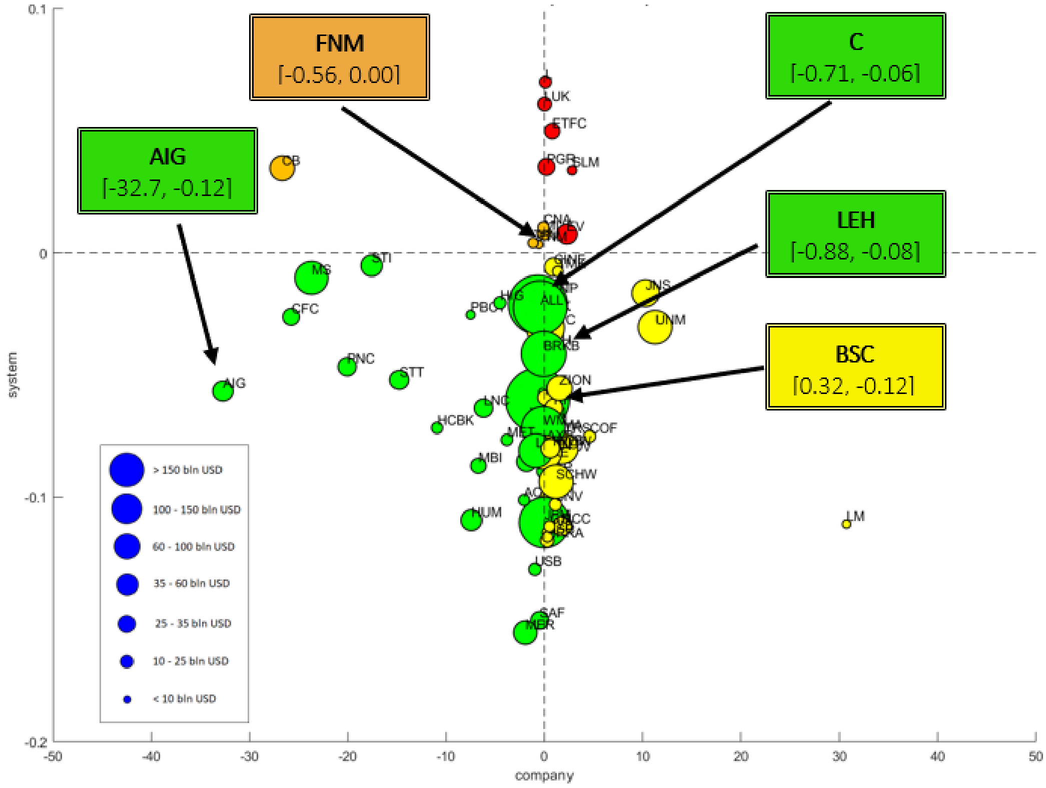
4.3. - TALIS^3 at work

We now use TALIS^3 to classify the 73 companies in our dataset. Figures 2-5 show a set of scatterplots summarizing company information at specific points in time. The horizontal axis represents $\left(MD_{t+1,60}^{j} - \overline{z}_{t+1,60}^{j}\right)$, i.e. the company's capital shortfall in stress states in deviation from the company's threshold. The vertical axis represents $\left(MD_{t+1,60}^{s} - \overline{z}_{t+1,60}^{s}\right)$, i.e. the magnitude of the capital shortage in the entire financial sector during stress states that exceed the system's threshold. Thresholds are computed using the historical values of the company or system's capital shortfalls. Consequently, positive (negative) values represents capital shortfall above (below) the historical thresholds. The figures include circles whose radii, based on the company's capitalization, allow the reader to easily compare and comprehend the difference in size for each company. We compute TALIS^3 for four different periods: 2005-2006 (pre-crisis), January 2007 to September 15, 2008 (crisis), September 16, 2008 to July 25, 2012 (European sovereign debt crisis), and finally from July 25, 2012 to the end of the sample. For each sub-period, *MD* is computed as the average of the daily loss functions during that sub-period. For example, for the January 2007-September 15, 2008 period, *MD* is computed as the average of the loss functions for that sub-period. The threshold is the maximum of the median of the loss functions for the previous 60 days (October, November and December in 2006) and the median of the loss functions for the full previous year (2006). Figures report end-of-period rankings.

Because of scale issues, companies with extremely large capital shortfalls (i.e. the largest values of $\left(MD_{t+1,60}^{j} - \overline{z}_{t+1,60}^{j}\right)$), do not appear in the plot. During the pre-crisis period, most of the large companies were in a green state and turned red, yellow or amber during the crisis. As an illustration, we follow five companies: Citigroup (C), Fannie Mae (FNM), American International

Group (AIG), Bear Stearns (BSC) and Lehman Brothers (LEH). All of these, according to the NYU Stern Systemic Risk Rankings webpage, were in the top 10 systemic firm group in July 2007. By March 2008, AIG had entered the top 10 rankings. TALIS^3 labelled C, AIG and LEH as green during the pre-crisis. During this period, FNM is amber and BSC is yellow. In the first part of the crisis, 2007/08, C turned yellow, AIG amber and FNM, BSC and LEH red. After Lehman's bankruptcy, FNM, BSC and LEH are not in our sample anymore and C is labelled as yellow and AIG as red. After Draghi's statement, "whatever it takes", the last part of the sample, C and AIG returned to green.

Figure 2.- Scatter plot: January 01, 2005 to December 31, 2006

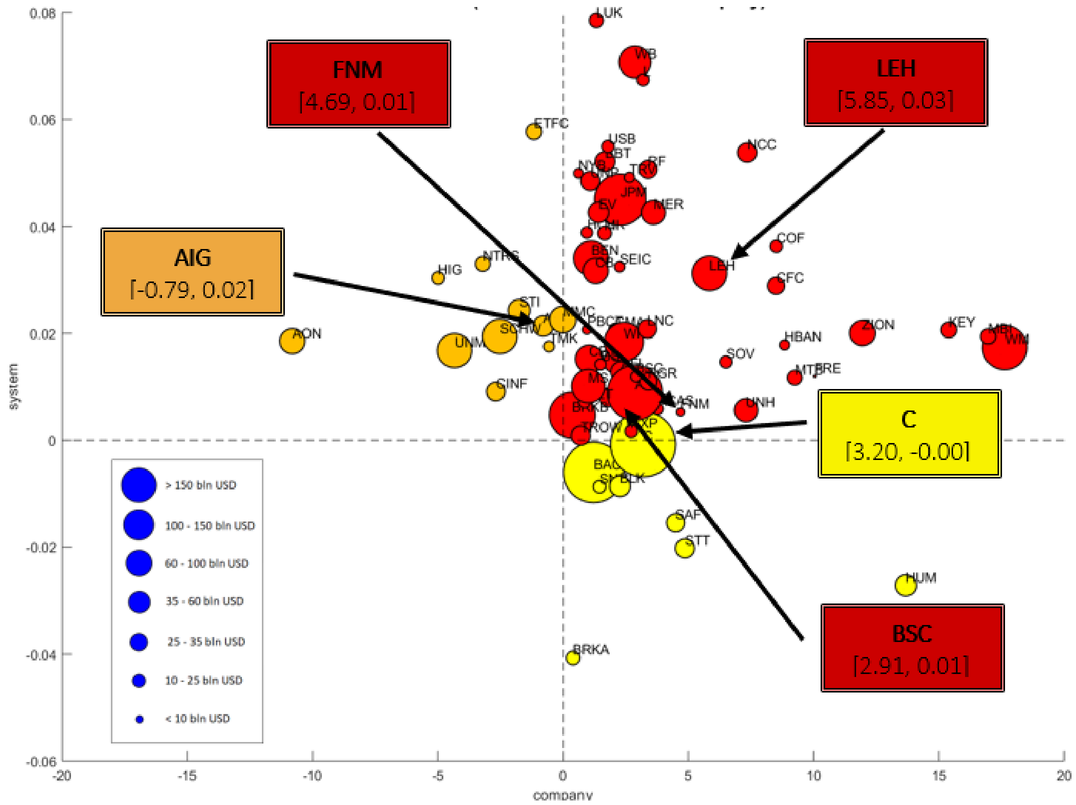


Note: Vertical (horizontal) axis represents $[MD^{s|j}_{t+1,60} - \bar{z}^{s|j}_{t+1,60}]$ ($[MD^{j}_{t+1,60} - \bar{z}^{j}_{t+1,60}]$), i.e. the system's (company's) capital shortfall in stress states that exceeds the system's (company's) threshold based on the own system's (company's) history of capital shortfalls (high stress state (+), low stress state (-)). Northeast (+ system, + company), red; Northwest (+, -), amber; Southwest (-, -), green; Southeast (-, +), yellow. The size of the circles is based on the company capitalization in the analyzed period. AON [-230.74, -0.02], AMTD [68.0, -0.04], MMC [-317.04, -0.01] and UNH [-55.66, -0.09] are not shown in the Figure because of scale issues. The Figure including all the companies is in the Appendix B.

As we can see in Figures 4 and 5, the number of red companies decreased during the European sovereign debt and post-crisis periods. For the full period, according to TALIS^3, green is the predominant color in the pre-crisis period, red in the first part of the crisis, yellow in the second part and green and yellow after the European sovereign debt crisis was curbed. Although scatter plots are snapshots of long periods of time, the general picture seems to follow the time series evolution of the SRISK measure by Brownless and Engle (2016). In July 2007, SRISK increased because of the subprime crisis and it reached its peak in September 2008. After March 2009, financial system capitalization improved and SRISK slightly decreased until the European sovereign debt crisis when a new increase is observed between 2011 and the summer of 2012. After Draghi's "whatever it takes" statement in July 2012, SRISK also decreased in line with the green and yellow color that dominates TALIS^3.

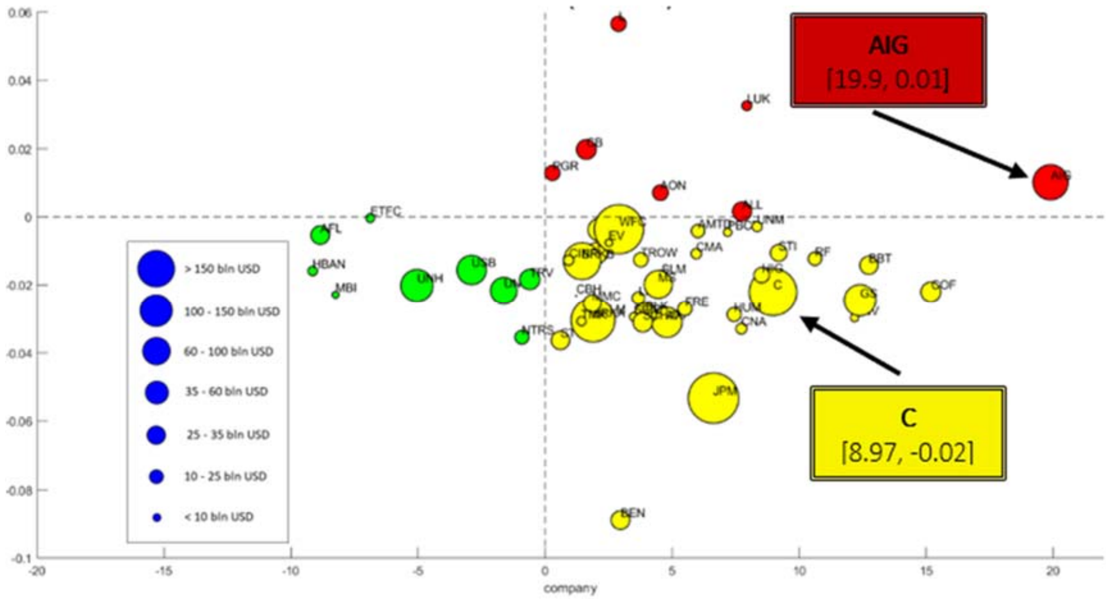
Since a monetary authority's goal is about the "early warning" detection of changes in the level of systemic risk, producing TALIS^3 at a daily level will be especially valuable for the rapid identification of institutions that are potentially most systematically important.

Figure 3.- Scatter plot: January 01, 2007 to September 15, 2008



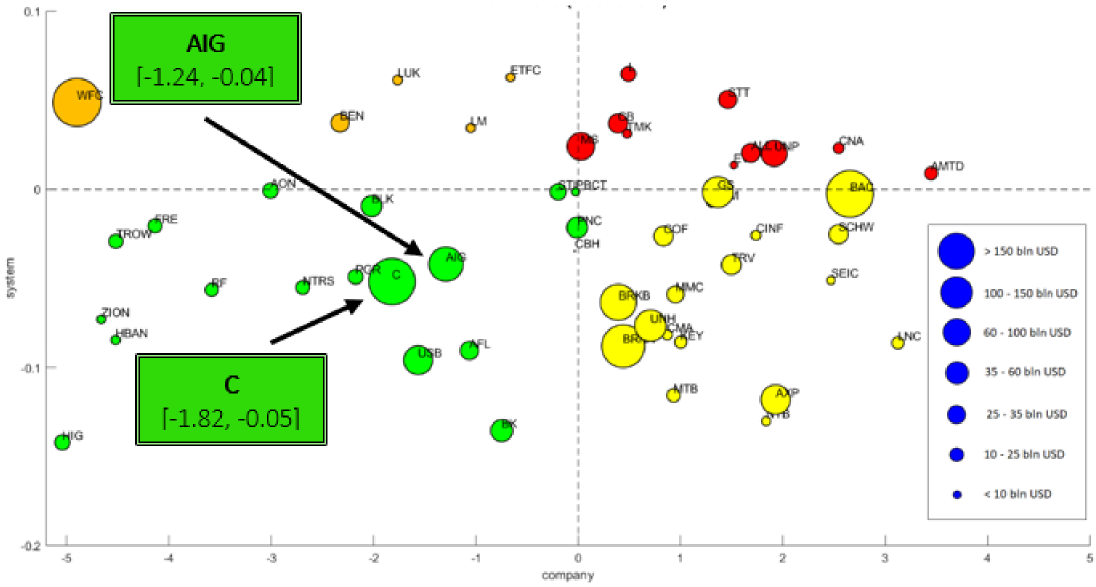
Note: Vertical (horizontal) axis represents $[MD^{s|j}_{t+1,60} - \bar{z}^{s|j}_{t+1,60}]$ ($[MD^{j}_{t+1,60} - \bar{z}^{j}_{t+1,60}]$), i.e. the system's (company's) capital shortfall in stress states that exceeds the system's (company's) threshold based on the own system's (company's) history of capital shortfalls (high stress state (+), low stress state (-)). Northeast (+ system, + company), red; Northwest (+, -), amber; Southwest (-, -), green; Southeast (-, +), yellow. The size of the circles is based on the company capitalization in the analyzed period. AMTD [-36.09, 0.04], CNA [20.25, 0.01], LM [-125.55, 0.02] and SLM [21.99, 0.01] are not shown in the Figure because of scale issues. The Figure including all the companies is in the Appendix B.

Figure 4.- Scatter plot: September 15, 2008 to July 25, 2012



Note: Vertical (horizontal) axis represents $[MD^{s|j}_{t+1,60} - \bar{z}^{s|j}_{t+1,60}]$ ($[MD^{j}_{t+1,60} - \bar{z}^{j}_{t+1,60}]$), i.e. the system's (company's) capital shortfall in stress states that exceeds the system's (company's) threshold based on the own system's (company's) history of capital shortfalls (high stress state (+), low stress state (-)). Northeast (+ system, + company), red; Northwest (+, -), amber; Southwest (-, -), green; Southeast (-, +), yellow. The size of the circles is based on the company capitalization in the analyzed period. BAC [23.20, -0.02], KEY [-71.99, -0.01], MTB [-26.59, -0.02] and ZION [-58.82, -0.01] are not shown in the Figure because of scale issues. The Figure including all the companies is in the Appendix B.

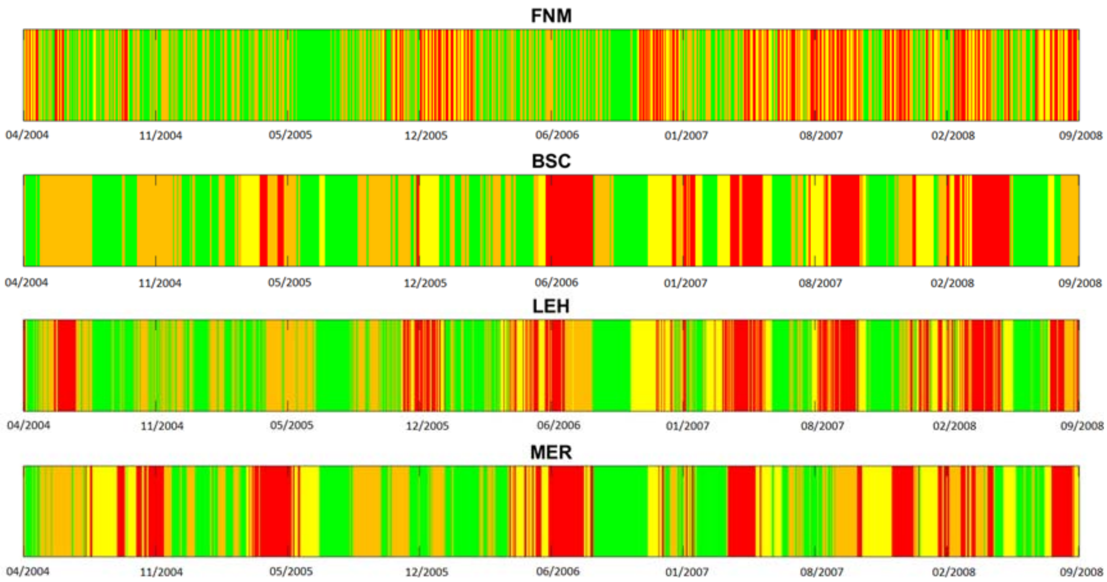
Figure 5.- Scatter plot: July 25, 2012 to the end of the sample



Note: Vertical (horizontal) axis represents $[MD^{s|j}_{t+1,60} - \bar{z}^{s|j}_{t+1,60}]$ ($[MD^{j}_{t+1,60} - \bar{z}^{j}_{t+1,60}]$), i.e. the system's (company's) capital shortfall in stress states that exceeds the system's (company's) threshold based on the own system's (company's) history of capital shortfalls (high stress state (+), low stress state (-)). Northeast (+ system, + company), red; Northwest (+, -), amber; Southwest (-, -), green; Southeast (-, +), yellow. The size of the circles is based on the company capitalization in the analyzed period. BBT [-7.76, -0.16], HUM [-30.49, -0.02], JPM [-17.97, 0.07], MBI [14.07, -0.01], SNV [-11.47, -0.07] and SLM [-33.96, 0.02] are not shown in the Figure because of scale issues. The Figure including all the companies is in the Appendix B.

Figures 6-10 show daily reports of TALIS^3 in equation (9). As described above, $MD^{s}_{t+1,m}\left(MD^{j}_{t+1,m}\right)$, the daily magnitude of the capital shortfall for the system (company) is computed as the average capital shortfall during the last 60 days (*m* = *60*), and $\overline{z}^{s}_{t+1,m}\left(\overline{z}^{j}_{t+1,m}\right)$, the system's (company) threshold, defined in (11), is computed as the Maximum of two arguments: the median of the full history of expected losses and the median of the last *60-day* window. Figure 6 shows daily heat maps of TALIS^3 for four companies during the first part of the sample: Fannie Mae (FNM), Bear Stearns (BSC), Lehman Brothers (LEH) and Merrill Lynch (MER).

Figure 6.- April 2004-September 15, 2018 Daily TALIS^3: Fannie Mae (FNM), Bear Stearns (BSC), Lehman Brothers (LEH) and Merrill Lynch (MER)



Note: The figure presents daily **TALIS^3** for the company with the name on the top. Let + (-) stand for high (low) stress state. Red (+ system, + company), amber (+ system, - company), yellow (- system, + company) and green (- system, - company) means that during that day this company is labelled as an EXTREMELY, VERY, MODERATELY or SLIGHTLY systemic company, respectively.

The top panel of Figure 6 shows the vulnerabilities of the system due to FNM's large losses on its retained portfolios in late 2007, especially on its subprime investments. In 2008, except for some days in May and June, FNM was mainly labelled as red, yellow or amber, as a sign of the troubles this company was facing and how potentially systemic it was. Looking more closely at what happened to FNM in 2008, we see that the huge size of its retained portfolios and mortgage guarantees led the American government's Federal Housing Finance Agency to conclude that FNM would soon be insolvent. On March 19, federal regulators allowed FNM to increase its debt in the hope of stabilizing the economy. By September 6, 2008, it was clear that the market believed that the firms were in financial trouble. Yellow and red labels for FNM in early September reflect these financial company vulnerabilities.

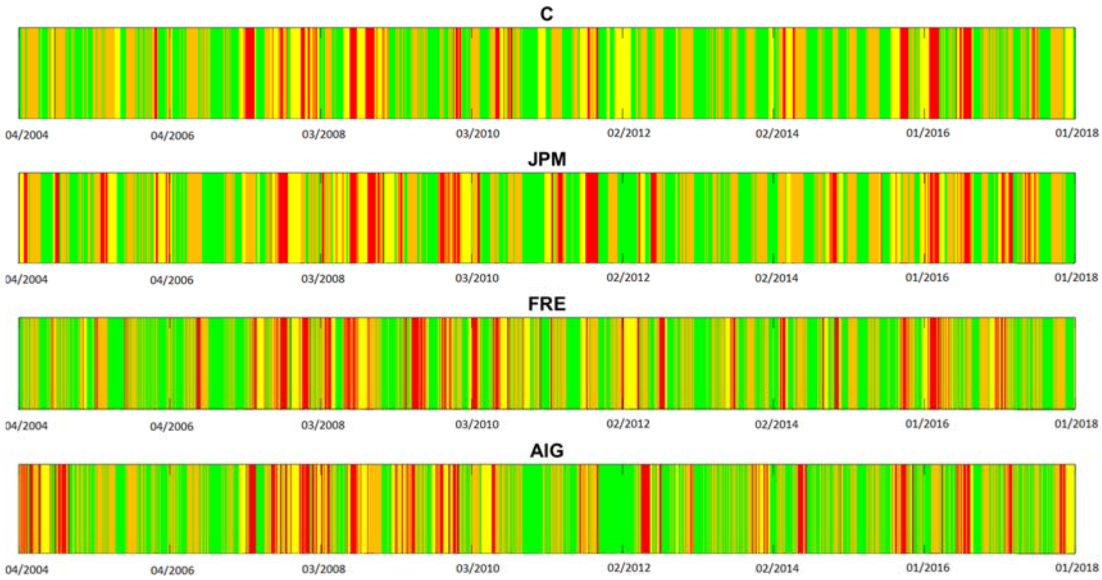
The middle of the top panels in Figure 6 shows the TALIS^3 for BSC. Bear Stearns is labelled as red from late June to late October in 2007. This pointed to the elevated vulnerability in the system because of the financial distress of BSC which on June 22, 2007 pledged a collateralized loan of up to $3.2 billion to "bail out" one of its funds. TALIS^3 documents a difficult 2008 for BSC starting on the morning of March 10, 2008 (red in TALIS^3) that served to trigger a run on Bear’s funding sources. First, the FED launched a $50-billion lending facility intended to support financial institutions in trouble. Second, a major rating agency downgraded a swathe of mortgage-backed securities issued by Bear. On March 16, 2008 Bear Stearns signed a merger agreement with JP Morgan Chase, but it took time for TALIS^3 to label BSC as green, remaining red and yellow until early June. June and July seem to be a calm period for BSC. Then in August, it is labelled as amber until Lehman’s bankruptcy.

The middle of the bottom panel in Figure 6 shows the TALIS^3 for LEH. Our measure shows the vulnerability in the system due to LEH starting in February 2007, just days after the company’s stock reached a record $86.18, giving Lehman a market capitalization of close to $60 billion. Cracks in the U.S. housing market were already becoming apparent as defaults on subprime mortgages rose to a seven-year high. On March 14, 2007 (yellow in TALIS^3), a day after the stock had its biggest one-day fall in five years, on concerns that rising defaults would affect Lehman's profitability, the firm reported record revenues and profits for its fiscal first quarter. In March 2017, LEH’s problems were becoming apparent and as seen in Figure 6, TALIS^3 started to label Lehman first yellow and then red. After some days of calm in June, between July and September 2007 TALIS^3 labelled Lehman first as amber and then red. Between mid-October and November TALIS^3 is green for LEH and from December 2007 to June 2008, LEH alternated between amber, red and yellow. July turned out to be a calm month for LEH, but in August and early September, just before Lehman’s collapse, our measure identified LEH as a systemically important company.

The bottom panel in Figure 6 shows the TALIS^3 for MER. It is labelled as red, yellow or amber on most days between mid-2007 and mid-2008. After a green June and mid-July, it turned yellow and then red in late-July until Lehman’s bankruptcy. Our measure shows evidence of the pronounced vulnerability in the system driven by the difficult situation of MER between 2007 and 2008 when it lost $19.2 billion, or $52 million daily. The company's stock price had also declined significantly during that time.

Figure 7 shows a daily heat map of the TALIS^3 for another four companies during the full period: Citi Group (C), JP Morgan (JPM), Freddie Mac (FRE) and American International Group (AIG). AIG, frequently ranked among the top ten systemic companies in 2008, is an interesting case. After starting the year 2008 in red, amber and yellow, it turned green in May until mid-July when amber, yellow and red colors were signs that more serious financial cracks were now becoming apparent.

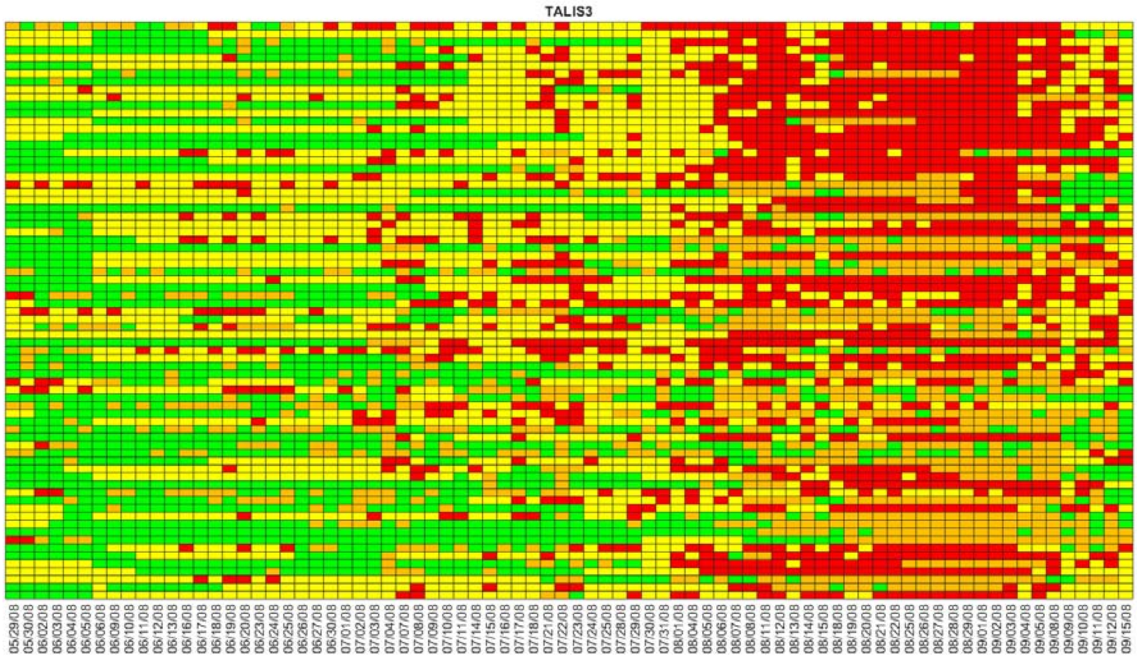
Figure 7.- Full period daily TALIS3: Citi Group (C), JP Morgan (JPM), Freddie Mac (FRE) and American International Group (AIG)



Note: The figure presents daily **TALIS^3** for the company with the name on the top. Let + (-) stand for high (low) stress state. Red (+ system, + company), amber (+ system, - company), yellow (- system, + company) and green (- system, - company) means that during that day this company is labelled as an EXTREMELY, VERY, MODERATELY or SLIGHTLY systemic company, respectively.

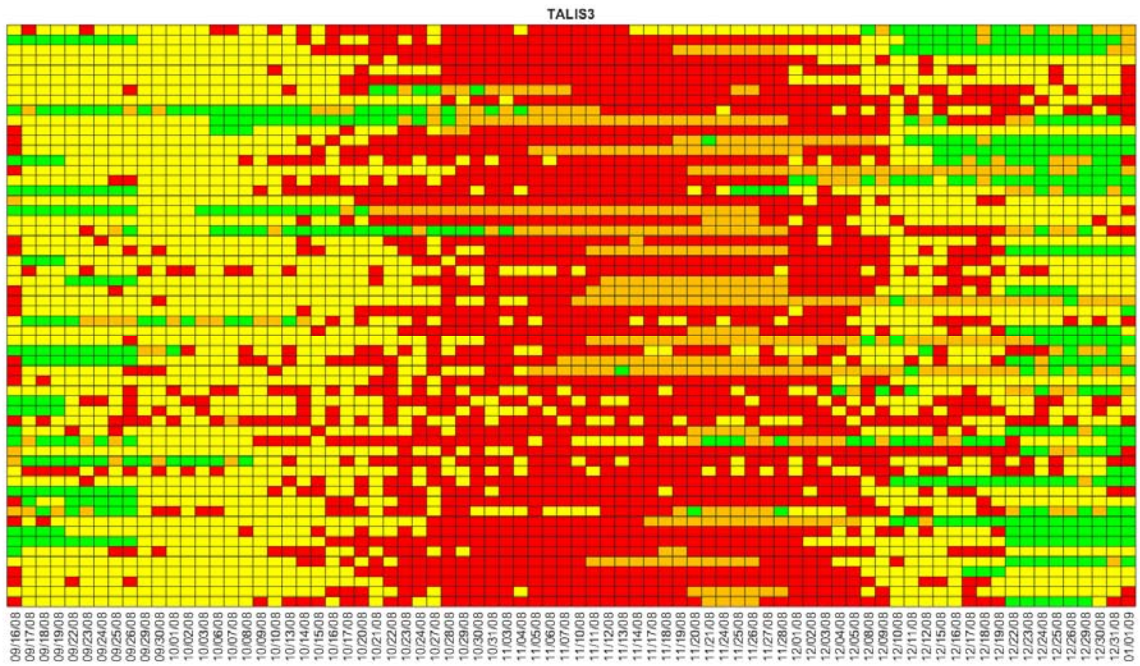
Figures 8 and 9 show the heat maps including all the companies for 70 days before and after LEH's bankruptcy on September 15, 2008, respectively. These figures show daily TALIS^3 rankings. Each row represents a company (name and order of the company is the same as reported in Table 1) while dates are on the x-axis. Red is the dominant color in August 2008. Although we find some green days, yellow, red and amber are the predominant colors during the days prior to the Lehman bankruptcy. The second half of September 2008 is yellow, while October and November are mainly red.

Figure 8.– TALIS³ for all companies. Period May 29 to September 15, 2008



Note: Each row in the figure presents 60-day **TALIS³** before Lehman bankruptcy for one of the 73 companies in Table 1. For example, first row from the top corresponds to Bank of America (BAC), and the fifth from the bottom to Lehman Brothers (LEH). Let + (-) sign stand for high (low) stress state. Red (+ system, + company), amber (+ system, - company), yellow (- system, + company) and green (- system, - company) means that during that day this company is labelled as an EXTREMELY, VERY, MODERATELY or SLIGHTLY systemic company, respectively.

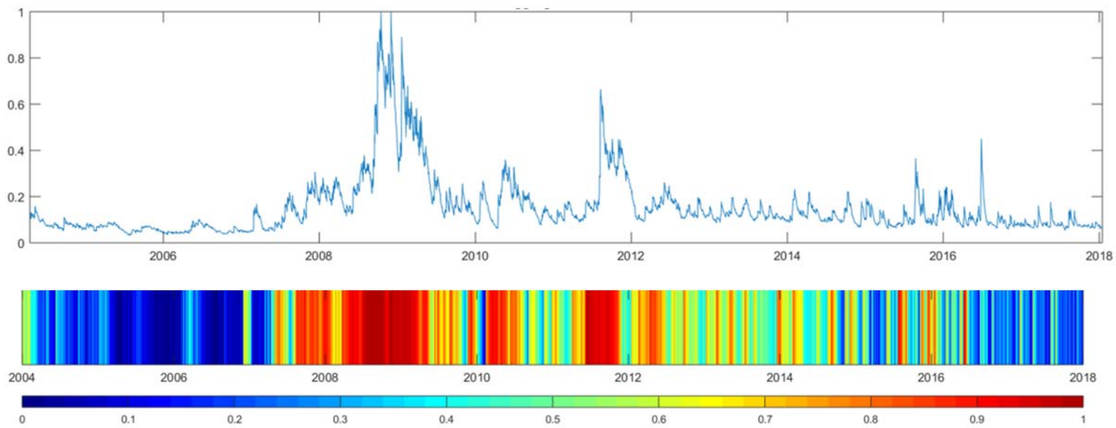
Figure 9.- TALIS³ for 58 companies. September 16, 2008/January 1, 2009



Note: Each row in the figure presents 60-day **TALIS³** for one of the 58 out of 73 companies remaining in the sample (in bold in Table 1) after Lehman bankruptcy. For example, first row from the top corresponds to Bank of America (BAC). Let + (-) stand for high (low) stress state. Red (+ system, + company), amber (+ system, - company), yellow (- system, + company) and green (- system, - company) means that during that day this company is labelled as an EXTREMELY, VERY, MODERATELY or SLIGHTLY systemic company, respectively.

Finally, we provide an aggregate view of the overall financial system through a time-varying aggregate version of TALIS^3 (A_TALIS^3) using (12), a weighted geometric mean where the weights are computed according to (13). The upper panel in Figure 10 shows A_TALIS^3 rescaled into probabilistic terms by using the kernel density estimate of its cumulative distribution. We follow the Aikman et al. (2017) strategy for presenting their aggregate index of vulnerabilities in the U.S. financial system. After computing A_TALIS^3, we estimate its empirical probability density function using a kernel distribution. Then each A_TALIS^3 observation is transformed into the [0, 1] interval based on its percentile in the empirical distribution. Therefore, each data point in the upper panel of Figure 10 is the estimated cumulative distribution function at the corresponding value of A_TALIS (i.e. the probability of observing values of A_TALIS^3 less than or equal to the raw data computed for that day).

Figure 10.– Heat map of aggregate TALIS^3



Note: Top panel of the figure shows aggregate TALIS^3 rescaled into probabilistic terms by using the kernel density estimate of its cumulative distribution. Each observation of A_TALIS^3's is transformed onto the [0, 1] interval based on its percentile in the empirical distribution. Therefore, each data point in the upper panel is the estimated cumulative distribution function at the corresponding value of A_TALIS (i.e. the probability of observing values of A_TALIS^3 less than or equal to the raw data computed for that day). Bottom panel presents the heat map of the index, a continuous color ramp assigning a color for each daily observation of the aggregate index. Colors are associated with the percentile of the distribution of aggregate TALIS^3. Low values of the index appear "cool" (deep blue) indicating low level of distress for the whole system, while large values appear "hot" (dark red), indicating high level of distress in the system.

The bottom panel of Figure 10 presents the heat map of the index**,** a continuous color ramp assigning a color to each daily observation in our index. Colors are associated with the percentile of the distribution of aggregate TALIS^3. Low values of the index appear "cool" (deep blue) indicating a low level of distress for the whole system, while large values appear "hot" (dark red), indicating a high level of distress in the system. Aggregate TALIS^3 and TALIS^3 can be updated on a daily basis once new information arrives in the market and illustrate the changes in risk and the depth of vulnerabilities of the full system. According to Figure 10, the more severe adverse scenarios seem to happen in the crisis period of the sample, in particular from mid-2007 to mid-2009 and from early 2011 to June 2012.In the next section we analyze whether our aggregate index can function as an early warning indicator for systemic risk.

4.4. Aggregate TALIS[3] as a predictor of financial distress

We now consider the predictive power of the Aggregated TALIS3 index, in particular with respect to the future changes in the stock market index. In doing that, we are particularly interested in the prediction of stressful market states compared to periods when the market is at a normal state. A comprehensive picture of the predictive ability of aggregated systemic risk measures on the stock markets when conditioning on the location of the stock market index returns over their density support can be obtained by using Quantile Regression (QR). We will consider the h-step-ahead predictive quantile regression model:

$$Q_{y_{t+h}}(\tau|Risk_t) = \alpha(\tau) + \sum_{i=1}^{p} \beta_i(\tau) Risk_{t-i+1} + u_t \tag{14}$$

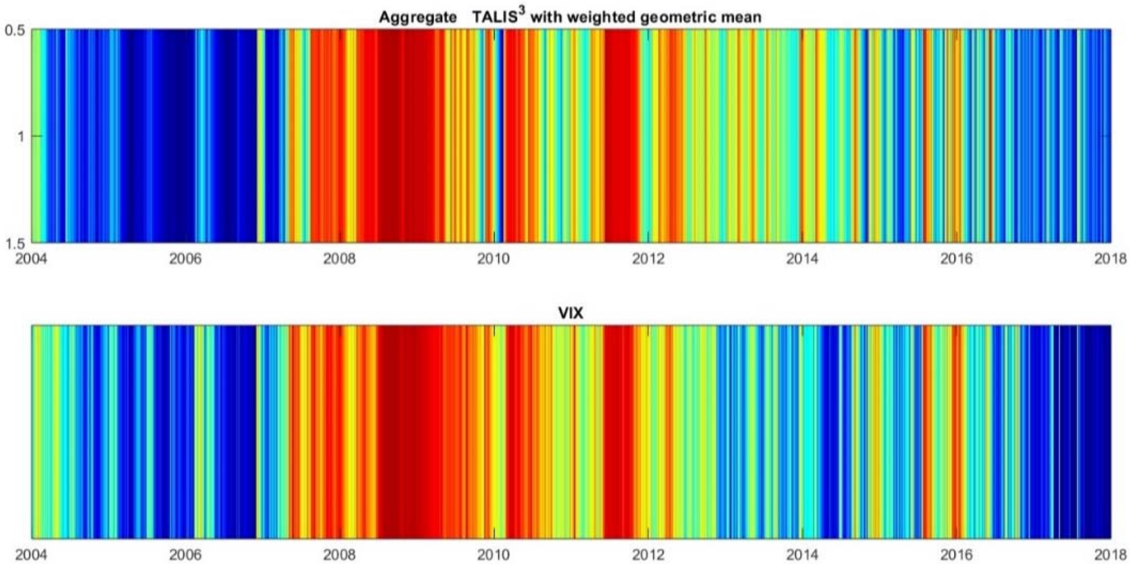
where $Q_{y_{t+h}}(\tau|.)$ denotes the $\tau - th$ quantile of the h-step ahead daily log change in the S&P100. Risk is one of the indicators we are adopting for the prediction of the equity market index lower tail. Finally, $\beta_i(\tau)$ measures the percentage change in the $\tau - th$ quantile of the S&P100 changes produced by the change in the predictor $(h + i - 1)$ days before. Note that with this approach we evaluate the informative content of the predictors for the target variable when the two are separated by at least h steps.

Among the predictors we include three proxies for Risk: i) the aggregate TALIS[3], ii) the VIX index, and iii) the Cleveland Financial Stress index (CFSI)[1] of Oet, Eiben, Bianco, Gramlich, and Ong (2011). The VIX captures variation in conditional moments of assets returns (see Adrian and Brunnermeier, 2016) and, as Brownless and Engle (2011) state, it also measures systematic risk that is highly correlated with systemic risk. The Chicago Board Options Exchange VIX (Chao, Härdle, and Wang, 2012; Exchange C.B.O., 2009), also known as the "fear" index contains information about future expectations of market volatility embodied in prices of the S&P100 at the money put and call options. Aggregate TALIS[3] and VIX share certain similarities as shown in Figure 11 that shows aggregate TALIS[3] and VIX index heat maps computed using the kernel

[1] The CFSI is based on publicly available data describing a six-market partition of the US financial system comprising credit, funding, real estate, securitization, foreign exchange, and equity markets. Units of CFSI are expressed as standardized differences from the mean (z-scores). Source: Federal Reserve Bank of St. Louis (FRED): https://fred.stlouisfed.org/series/CFSI. Sample: 04/21/2004 – 05/04/2016.

density estimate of their respective cumulative distribution functions, as explained above. Aggregate TALIS^3 and VIX heat maps seem to report similar colors during the crisis, although VIX reports colder color states at the end of the sample. The correlation between aggregate TALIS^3 and VIX is very high, reaching a value of 0.90.

Figure 11.- Heat map of aggregate TALIS^3 versus VIX



Note: Top panel of the figure shows the heat map of the aggregate TALIS^3 and the bottom panel presents the heat map of the VIX index, both are rescaled into probabilistic terms by using the kernel density estimate of their cumulative distribution. Colors are associated with the percentile of their distributions. Low values of aggregate TALIS^3 appear "cool" (deep blue) indicating low level of distress for the whole system, while large values appear "hot" (dark red), indicating high level of distress in the system. Regarding the VIX index (*investor fear gauge*), low values (deep blue) indicate lower expectation of rising volatility, while large values (dark red) would indicate higher expectation of rising volatility and times of financial stress. Correlation between both indices is **0.90.**

Regarding the Cleveland Financial Stress index (CFSI)[2] of Oet et al. (2011) is one of the few financial stress indices that uses daily observable financial-market data to capture the level of financial stress. Following indicators of stress of different markets (equity, foreign exchange, funding, real state and securitization), the CFSI allows to identify the level of stress in specific markets and provides insight into the nature of the stress in the market (see Oet, Dooley and Ong, 2015, Mananperi, 2013 and Kliesen, Owyang and Vermann, 2012).

Table 5 shows the results for a subset of the predictive QR that we run using a 1000-observation rolling window for five different quantiles, 1%, 5%, 10%, 20% and 50%. We report summary results for equation (14) for three forecasting horizons, 1 day, 1 week (5 days), and 1

[2] The CFSI is based on publicly available data describing a six-market partition of the US financial system comprising credit, funding, real estate, securitization, foreign exchange, and equity markets. Units of CFSI are expressed as standardized differences from the mean (z-scores). Source: Federal Reserve Bank of St. Louis (FRED): https://fred.stlouisfed.org/series/CFSI. Sample: 04/21/2004 – 05/04/2016.

month (22 days). The table reports the median of the estimated coefficients (MED), the percentage of times, out of N (the number of estimated QR), in which the coefficient $\beta_i(\tau)$ is statistically significant at the 5% confidence level (PS) – the parameter's standard deviation is computed using bootstrapping for each QR. We also report the average Pseudo R2 (Av-Ps-R2) of Koenker and Machado (1999).[3]

At the median, the stress indices do not seem to contain information that can be used to predict the financial index: $\beta_i(0.5)$ estimates are close to zero and the PS is below 16% for all the three indices. However, the predictor coefficient becomes negative and larger as we move downward into the left tail of the S&P100 returns distribution. Therefore, we found that the prediction of the average future decline in the S&P100 following an increase in the systemic risk index is much larger when the market is in distress than during normal times.

For Aggregate TALIS3 we note that the lower the percentile of the S&P100, the larger both the PS and the Av-Ps-R2, which clearly indicate that the predictive power is larger at low percentiles of the S&P100 changes. This further confirms that an increasing aggregate TALIS^3 predicts future drops in the market. In the QR estimate of Equation (14), TALIS^3 is strongly significant for the 1% and 5% quantiles of the 5-day horizon, while increasing the horizon up to 22 days (1 month) we note a decrease in both the PS and Av-Ps-R2. At the 1% and 5% quantiles, both VIX and CFSI indices show the largest PS at the 1-day horizon. The Av-Ps-R2 decreases when the horizon increases for the three indices. The CFSI index shows the lowest Av-Ps-R2 for all the quantiles and horizons. Overall, the predictive ability of Aggregate TALIS^3 for equity index quantiles is better exploited at extreme quantiles and for horizons beyond the single day. In particular, we note that at the 22-days horizon the TALIS^3 performances are always better than those of the VIX. This is somewhat counterintuitive as, in a predictive quantile regression setting, we might expect that the VIX performs better in long-term prediction due to the methodology used for its construction. The differences we detect highlight that the VIX performances are inferior, and we link these results to the different information set adopted in the construction of TALIS^3.

5. Robustness analysis

5.1. Changing the GARCH model

The results shown in the previous sections were obtained using a DCC specification for the bivariate GJR-GARCH model. In addition to DCC, we examine two other specifications, the

[3] The table reports the results for the QR estimates using one lag of the dependent variables. Although not shown in the table, we observed that the Av-Ps-R2 slightly increased by introducing a larger number of lags.

more flexible full BEKK of Engle and Kroner (1995) and the OGARCH model, introduced by Alexander and Chibumba (1997), designed to overcome the well-known curse of dimensionality. In this section we examine how TALIS[3] behaves when the MGARCH model is modified.

Table 5.- Forecasting power of TALIS[3]

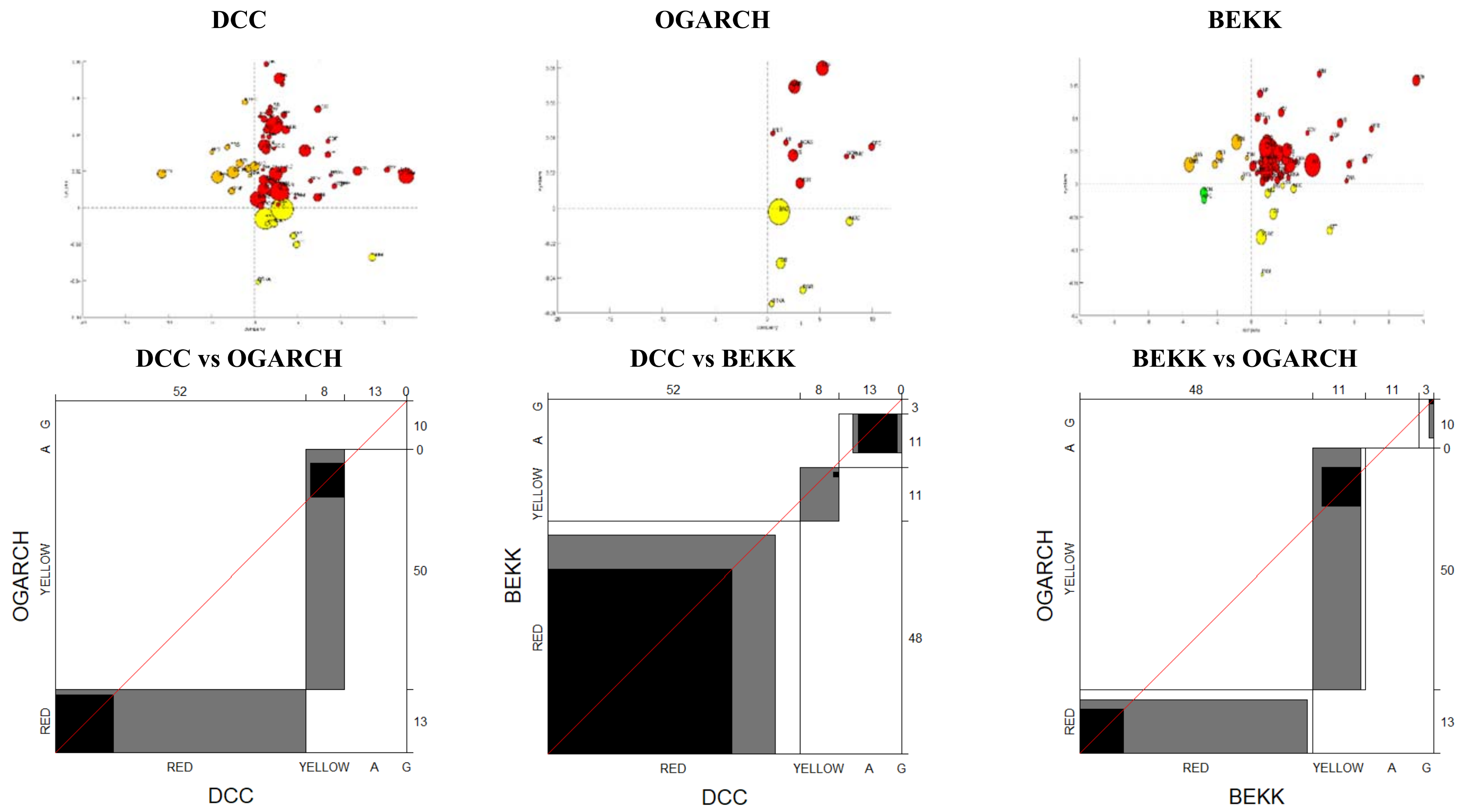
	$h = 1$			$h = 5$			$h = 22$		
				$\tau = 1\%$					
	MED	**PS**	**Av-Ps-R^2**	**MED**	**PS**	**Av-Ps-R^2**	**MED**	**PS**	**Av-Ps-R^2**
TALIS	-0.0129	83.00%	0.2273	-0.013	89.93%	0.1871	-0.0086	69.14%	0.1037
VIX	-0.0149	98.27%	0.2694	-0.014	70.93%	0.2128	-0.0048	55.03%	0.0740
CFSI	-0.0060	73.24%	0.1050	-0.005	70.83%	0.0886	-0.0039	55.83%	0.0638
				$\tau = 5\%$					
	MED	**PS**	**Av-Ps-R^2**	**MED**	**PS**	**Av-Ps-R^2**	**MED**	**PS**	**Av-Ps-R^2**
TALIS	-0.0089	91.97%	0.1444	-0.008	92.23%	0.1244	-0.0068	98.40%	0.0921
VIX	-0.0107	99.95%	0.1564	-0.009	99.67%	0.1422	-0.0070	94.57%	0.0762
CFSI	-0.0048	82.91%	0.0636	-0.004	78.14%	0.0549	-0.0035	74.04%	0.0449
				$\tau = 10\%$					
	MED	**PS**	**Av-Ps-R^2**	**MED**	**PS**	**Av-Ps-R^2**	**MED**	**PS**	**Av-Ps-R^2**
TALIS	-0.0079	92.15%	0.0972	-0.007	91.67%	0.0838	-0.0055	92.21%	0.0608
VIX	-0.0078	100.00%	0.1019	-0.007	97.52%	0.0917	-0.0048	55.03%	0.0480
CFSI	-0.0041	86.74%	0.0486	-0.003	77.67%	0.0402	-0.0027	62.15%	0.0306
				$\tau = 20\%$					
	MED	**PS**	**Av-Ps-R^2**	**MED**	**PS**	**Av-Ps-R^2**	**MED**	**PS**	**Av-Ps-R^2**
TALIS	-0.0050	90.10%	0.0516	-0.004	88.11%	0.0448	-0.0036	81.26%	0.0333
VIX	-0.0048	99.53%	0.0478	-0.004	90.68%	0.0441	-0.0029	71.21%	0.0234
CFSI	-0.0026	80.29%	0.0262	-0.002	61.00%	0.0235	-0.0021	55.78%	0.0189
				$\tau = 50\%$					
	MED	**PS**	**Av-Ps-R^2**	**MED**	**PS**	**Av-Ps-R^2**	**MED**	**PS**	**Av-Ps-R^2**
TALIS	0.0000	1.35%	0.0008	0.000	3.98%	0.0008	-0.0006	7.60%	0.0016
VIX	0.0003	16.07%	0.0014	-0.000	0.00%	0.0003	-0.0006	2.64%	0.0011
CFSI	0.0001	0.79%	0.0005	0.000	1.83%	0.0004	0.0001	1.88%	0.0004

Note: Table reports estimation results of the predictive QR defined in Equation (14) using a 1000-observation rolling window. We estimated the conditional quantiles for five quantile levels 1%, 5%, 10%, 20% and 50%. The regression measures the change in the S&P100 based on the change in the predictors *h* days before. Each panel, from left to right, reports the median of the estimated coefficients (MED), percentage of times, out of *N*, in which they are statistically significant at the 5% level of confidence (PS). The Av-Ps-R^2 is the average Pseudo R^2 computed as one minus the ratio of the residual absolute sum of weighted differences using the selected model to the residual absolute sum of weighted differences using a model with only the intercept.

Table C.1a (C.1b) in Appendix C shows summary statistics of $\Delta CoVaR_{q,t}^{s|j}$ for all 73 financial institutions during the first part of the sample (before Lehman's Bankruptcy) for the BEKK (OGARCH) model. Table C.2a (C.2b) shows summary statistics of $\Delta CoVaR_{q,t}^{s|j}$ for the 58 companies remaining in the sample after Lehman's Bankruptcy. The results are similar to those obtained in Tables 4 and 5 when the DCC specification was used. Brokers and dealers are the most systemic companies followed by Depositories. Others and Insurance companies are in last position as the least systemic. As for the DCC case, the average $\Delta CoVaR_{q,t}^{s|j}$ is larger after Lehman's bankruptcy, -3.21 percent compared to -2.00 percent before the Lehman crash in September 15, 2008.

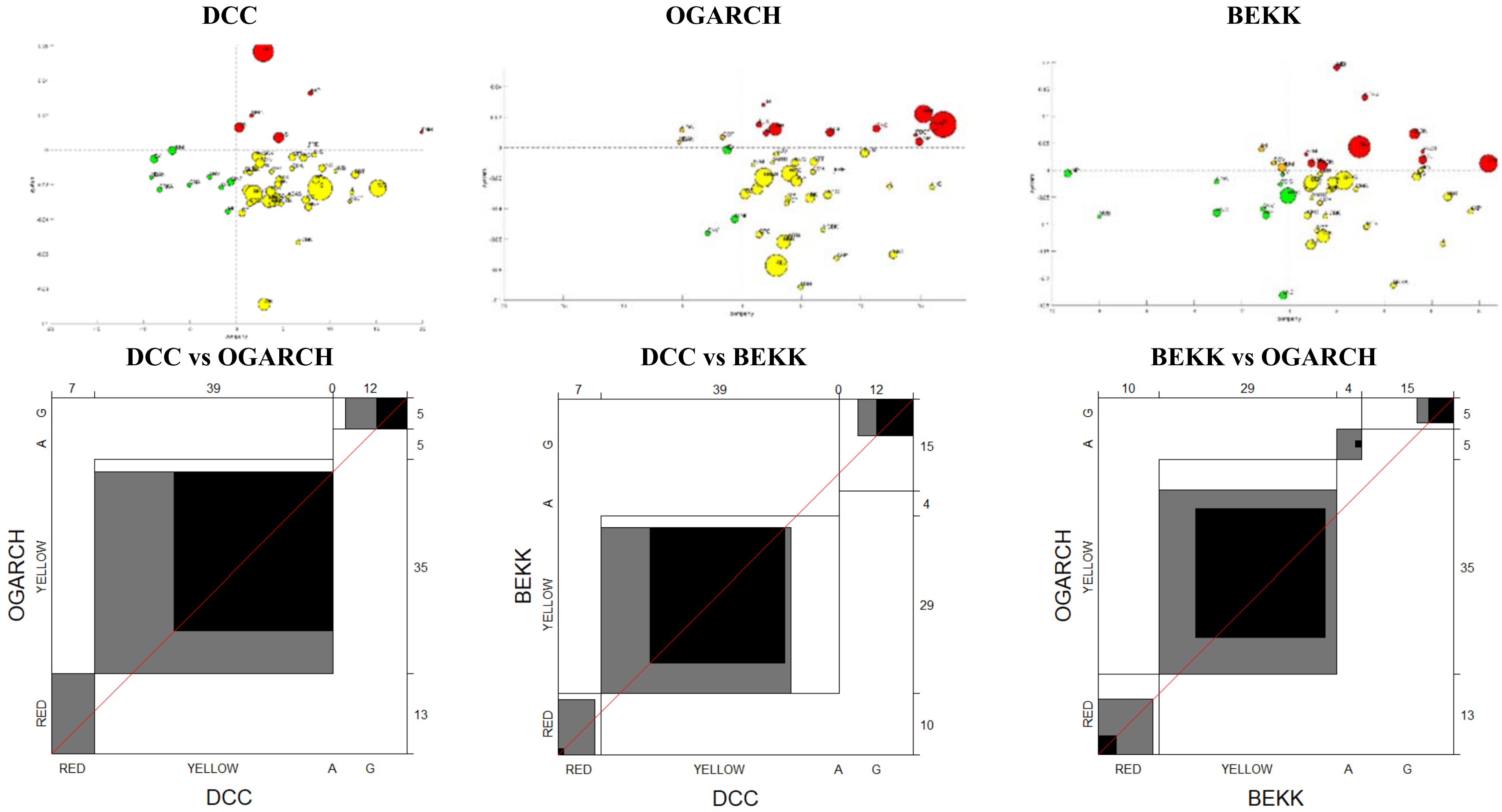
We repeat the exercise and compare the behavior of TALIS[3] computed under the BEKK and OGARCH specifications. TALIS[3] company rankings allow us to perform model agreement exercises based on the popular Cohen kappa (1973) and Bangdiwala (1988) tests, along with the well-known Bangdiwala (1988) agreement charts.

Figure 12. - TALIS[3] as reported by DCC, OGARCH and BEKK model specifications and Bangdiwala's (1988) Observer Agreement Charts for the crisis period until Lehman Brothers bankruptcy.



Note. The first row of the table shows TALIS[3] scatter plots during the period analyzed when using DCC, OGARCH or BEKK specification. In these sub-plots, vertical (horizontal) axis represents $\left[MD_{t+1,60}^{s} - \bar{z}_{t+1,60}^{s}\right]$ ($\left[MD_{t+1,60}^{j} - \bar{z}_{t+1,60}^{j}\right]$), i.e. the system's (company's) capital shortfall in stress states that exceeds the system's (company's) threshold based on the own system's (company's) history of capital shortfalls (high stress state (+), low stress state (-)). The second row of the table shows Bangdiwala's (1988) agreement charts to compare TALIS[3] rankings provided by the three alternative MGARCH specifications. The agreement between models is represented by the size and level of coincidence of the superposed white, grey and black rectangles inside the charts for each category, in our case, colors: red, yellow, amber (A) and green (G). To illustrate the information provided by these charts, consider position (2, 2) that represents the agreement chart between TALIS[3] rankings based on DCC and BEKK during 2007-2008 period. The size of the large square of this chart is N by N, where N is the total number of companies available in the sample during this first part of the sample. The vertical (horizontal) axis represents the number of companies ranked as red, yellow, A or G by TALIS[3] based on the BEKK (DCC) specification. Inside this, we can find a set of four rectangles, one for each color. Looking at the first set of rectangles on the lower left vertex starting from (0,0), which refers to the red color category, we can see a white big rectangle of size 52 by 48 (marginal totals). It reports that DCC (BEKK) ranked 52 (48) companies as red. The black square placed inside of side 38 by 38, reports that both DCC and BEKK ranked the same 38 companies as red. Finally, the grey rectangle size takes into account partial agreement counting it differently with decreasing weights for cells further away from the diagonal cells. For example, in addition to the strict agreement of 38 companies ranked as red, we found that 9 out of 52 companies ranked as red by DCC, were ranked as yellow, 3 as amber and 2 as green by BEKK. 7 out of 48 companies ranked as red by BEKK were ranked as yellow and 3 as amber by DCC. None of these 48 companies was ranked as green by DCC. An additional matrix of weights must be considered. Two commonly-used patters are those based on equal spacing weights (Cicchetti and Alison, 1971) (CA) that are defined by $1-|i-j|\,/(r-1)$, where r is the number of columns/rows, and Fleiss and Cohen (1973)'s weights (FC), defined by $1-|i-j|^2\,/(r-1)^2$. Fleiss-Cohen weights attach greater importance to near disagreements. In the case of perfect agreement, black, grey and white rectangles are exactly equal. Lesser agreement is visualized by comparing the area of the blackened squares to the area of the rectangles, while model bias is visualized by examining the 'path of rectangles' and how it deviates from the 45° diagonal line within the larger $N \times N$ square. A detailed description of these agreement plots can be found in Bangdiwala and Shankar (2013).

Figure 13.- TALIS[3] as reported by DCC, OGARCH and BEKK model specifications and Bangdiwala's (1988) Observer Agreement Charts for the crisis period after Lehman Brothers bankruptcy, September 16, 2008 to July 24, 2011.



Note. See notes to Figure 12.

Figures 12 and 13 show scatterplots of TALIS^3 and Bangdiwala's (1988) agreement charts during the first and second parts of the crisis for the three MGARCH specifications, DCC, BEKK and OGARCH. The first row in the tables shows TALIS^3 scatter plots and the agreement charts are shown in the second row. In the first row, Figure 12 shows TALIS^3's scatter plots from January 2007 to September 15, 2008 produced using DCC, OGARCH and BEKK. For example, the scatter plot in position (1,1) is the same as the one in Figure 3. Regardless of the model, TALIS^3 ranks most of the companies as red and yellow during this period. In the second row, in Figure 12, we find the agreement charts. We first produce a two-way contingency table from the company rankings provided by TALIS^3 computed using two different MGARCH specifications. Each specification classifies the N companies as red, yellow, amber or green. From these tables we produce the agreement charts, whose square layout captures all the information contained in the two-way contingency table. Agreement between models is represented by the size and coincidence level of the superposed white, grey and black rectangles inside the charts for each category which, in our case, are colors: red, yellow, amber (A) and green (G). In the case of perfect agreement, the black, grey and white rectangles are exactly equal. This would happen if the TALIS^3 produced using two different MGARCH specifications exactly agreed, that is, all the off-diagonal cells of the two-way table were zeroes. If there is no agreement whatsoever, that is, if all the diagonal cells of the two-way table were zeroes, there would be no blackened area. Lesser agreement is visualized by comparing the area of the blackened squares to the area of the rectangles, while model bias is visualized by examining the 'path of rectangles' and how it deviates from the 45° diagonal line within the larger $N \times N$ square. To illustrate this, just look at position (2,2) in Figure 12. Inside the large square, of a size 73 by 73, where 73 is the number of companies in the sample, in the first set of rectangles on the lower left vertex starting from (0,0), which refers to the red color category, we can see a large white rectangle of the size 52 by 48 (marginal totals). This reflects the fact that DCC ranks 52 companies in the red state while BEKK ranks 48 companies in the red state. The black square inside, of the size 38 by 38, shows that both DCC and BEKK ranked the same 38 companies as red. Finally, the size of the grey rectangle takes into account partial agreement between DCC and BEKK when ranking companies. Partial agreement is counted differently, with decreasing weights for cells further away from the diagonal cells. For example, in addition to the strict agreement of 38 companies ranked as red, we found that 9 out of 52 companies ranked as red by DCC were ranked as yellow by BEKK, with an additional 3 as amber and 2 as green. Seven out of the 48 companies ranked as red by BEKK were ranked as yellow by DCC and 3 as amber. None of these 48 companies were ranked as green by DCC.

We might also be interested in the agreement across major categories in which there is a meaningful difference. For example, as we have four categories, "red", "yellow", "amber" and "green" we may care less about one model ranking a company as "red" while another ranks it as "yellow" than we do about one model ranking a company as "red" while the other ranks it as "green". Therefore, we can also account for partial agreement by considering the information contained in the off-diagonal cell entries in the two-way contingency table. We can construct an additional matrix of weights that assigns less weight to agreement for categories that are further apart. There are two commonly-used patterns, one based on equal spacing weights (linear) (Cicchetti and Alison, 1971) that are defined by $1 - |i - j| / (r - 1)$, where r is the number of columns/rows and i and j are the index of the two-way contingency table. The other weighting strategy is to square the difference between groups, as suggested by Fleiss and Cohen (1973), defined by $1 - |i - j|^2 / (r - 1)^2$. Fleiss-Cohen's weights attach greater importance to near disagreements. A detailed description of these agreement plots can be found in Bangdiwala and Shankar (2013)

The main conclusions from the agreement charts is that during calm periods, before the great crisis (2004-2006) and then after it (2011-2018), green, yellow and amber are the predominant colors in the TALIS[3] rankings. Nonetheless, as seen in Figures 12 and 13, regardless of the model, during the crisis period, before (2007-September 15, 2008) and after (September 15, 2008-July 25, 2011) Lehman's bankruptcy, TALIS[3] ranks most of the companies as red and yellow. Rectangles for red and yellow categories are larger than for the other two colors, green and amber.

Statistics of agreement are reported in Tables 6-8, namely unweighted (Bangd-U) and weighted (Bangd-W) Bangdiwala statistics along with the unweighted (K-U) and weighted (K-W) Kappa statistics. Weighted statistics are computed using Fleiss and Cohen`s (1973) weights. They are all common measures of agreement between two models when classifying the N companies into the four mutually exclusive categories: red, yellow, amber and green. Table 6 shows the agreement statistics for the correspondence between DCC and OGARCH, Table 7 shows these statistics when comparing DCC and BEKK, and Table 8 for BEKK and OGARCH. In Table 6, the Bangd-W statistic shows values above 0.88, which can be interpreted as a good degree of agreement between DCC and OGARCH specifications. In general, for the three models being compared, Bangd-W shows larger values, above 0.65, during the two crisis periods. This can be interpreted as the degree of agreement being larger during crises than during calm periods.

To obtain a global picture of these statistics we compute geometric averages by column, seeking a measure through which to rank the degree of agreement between MGARCH specifications. Geometric averages of the four statistics (Bangd-U, Bangd-W, K-U and K-W) during the four periods (Pre-C, Crisis 1, Crisis 2 and Post-C) show reasonable similarity between DCC and OGARH (0.42), followed by BEKK and OGARCH (0.27), and finally BEKK and DCC

(0.23). Searching for the period showing the largest agreement, we compute geometric averages by row (four tests for the three available combinations). Globally, no large differences were found between the first three periods (crisis 2 (0.37), pre-crisis (0.36) and crisis 1 (0.35)).

Table 6.- Test of correspondence DCC/OGARCH

	Bangd-U	Bangd-W	K-U	K-U_ASD	K-W	K-W_ASD
Pre-C	0.31	0.89	0.22	0.06	0.74	0.03
Crisis 1	0.18	0.91	0.07	0.04	0.57	0.06
Crisis 2	0.46	0.89	0.15	0.10	0.64	0.08
Post-C	0.45	0.88	0.31	0.08	0.73	0.07

Note. Bang-U and Bang-W are the unweighted and weighted Bangdiwala statistics, respectively, K-U and K-U_ASD are the unweighted Kappa statistics and its asymptotic standard deviation, respectively and K-W and K-W_ASD are the weighted Kappa statistics and its asymptotic standard deviation, respectively. Agreement statistic measures the degree of agreement between the methods above that expected by chance alone. It has a maximum of 1 when agreement is perfect, 0 when agreement is no better than chance, and negative values when agreement is worse than chance. Other values can be roughly interpreted as: < 0.20- Poor, > 0.40 - Fair, > 0.60 – Moderate, > 0.80 - Good, 1 - Very good.

Table 7.- Test of correspondence DCC/BEKK

	Bangd-U	Bangd-W	K-U	K-U_ASD	K-W	K-W_ASD
Pre-C	0.15	0.47	0.15	0.08	0.30	0.12
Crisis 1	0.55	0.81	0.27	0.10	0.42	0.13
Crisis 2	0.38	0.65	0.15	0.09	0.35	0.12
Post-C	0.08	0.49	-0.01	0.06	0.01	0.13

Note. Bang-U and Bang-W are the unweighted and weighted Bangdiwala statistics, respectively, K-U and K-U_ASD are the unweighted Kappa statistics and its asymptotic standard deviation, respectively and K-W and K-W_ASD are the weighted Kappa statistics and its asymptotic standard deviation, respectively. Agreement statistic measures the degree of agreement between the methods above that expected by chance alone. It has a maximum of 1 when agreement is perfect, 0 when agreement is no better than chance, and negative values when agreement is worse than chance. Other values can be roughly interpreted as: < 0.20- Poor, > 0.40 - Fair, > 0.60 – Moderate, > 0.80 - Good, 1 - Very good.

Table 8.- Test of correspondence BEKK /OGARCH

	Bangd-U	Bangd-W	K-U	K-U_ASD	K-W	K-W_ASD
Pre-C	0.26	0.66	0.30	0.08	0.54	0.09
Crisis 1	0.12	0.77	0.27	0.10	0.42	0.13
Crisis 2	0.38	0.76	0.21	0.09	0.24	0.14
Post-C	0.11	0.44	0.02	0.07	0.21	0.12

Note. Bang-U and Bang-W are the unweighted and weighted Bangdiwala statistics, respectively, K-U and K-U_ASD are the unweighted Kappa statistics and its asymptotic standard deviation, respectively and K-W and K-W_ASD are the weighted Kappa statistics and its asymptotic standard deviation, respectively. Agreement statistic measures the degree of agreement between the methods above that expected by chance alone. It has a maximum of 1 when agreement is perfect, 0 when agreement is no better than chance, and negative values when agreement is worse than chance. Other values can be roughly interpreted as: < 0.20- Poor, > 0.40 - Fair, > 0.60 – Moderate, > 0.80 - Good, 1 - Very good.

5.2. The role of thresholds and the size of the rolling window

In this section we analyze the sensitivity of the TALIS[3] rankings to the size of the window, *m,* chosen to compute the rolling statistics on which TALIS[3] is based, $MD^{s}_{t+1,m}\left(MD^{j}_{t+1,m}\right)$, the daily magnitude of the capital shortfall for the system (company) and $\overline{z}^{s}_{t+1,m}\left(\overline{z}^{j}_{t+1,m}\right)$, the system's (company) thresholds, computed as the maximum of two arguments: the median of the full history of expected losses and the median of the last *m-day* window.

For four different periods: pre-crisis (from January 2005 to December 2007), pre-Lehman crisis (January 2007 to September 15, 2008), post-Lehman crisis (from September 16, 2008 to July 24, 2011) and post-crisis (from July 25, 2011 to February 2018), Figures D.1-D.3 in Appendix D show Bangdiwala's (1988) agreement charts described above. Figure D.1 compares rankings using $m_1 = 20$ and $m_2 = 60$. Figure D.2 presents the agreement between the rankings for $m_1 = 60$ and $m_2 = 120$ and Figure D.3 shows the analysis for $m_1 = 20$ and $m_2 = 120$. For example, for the pre-Lehman crisis period (January 2007 to September 15, 2008), $MD^{s}_{t+1,m}\left(MD^{j}_{t+1,m}\right)$ is computed as the average of the loss functions during this period, and $\overline{z}^{s}_{t+1,m}\left(\overline{z}^{j}_{t+1,m}\right)$ will be the maximum median losses during the previous year, 2006, and the median of the previous *m-day* window (if m = 20 (60) [120], the last 20 (3) [6] days (months) [months] of December, 2006).

Agreement between TALIS[3] rankings is represented by the size and level of coincidence of the superposed white, grey and black rectangles inside the charts for each category, in our case, colors: red, yellow, amber (A) and green (G). In the case of perfect agreement, the black, grey and white rectangles are exactly equal. From Bangdiwala's chart it is difficult to reach a conclusion. Tables D.1-D.3 in Appendix D show the unweighted (Bangd-U) and weighted (Bangd-W) Bangdiwala along with the unweighted (K-U) and weighted (K-W) Kappa statistics. Table D.1 shows these statistics for the comparison between $m_1 = 20$ and $m_2 = 60$, Table D.2 for $m_1 = 60$ and $m_2 = 120$, and Table D.3 for $m_1 = 20$ and $m_2 = 120$. When we look at these statistics, the main conclusion is that the agreement between 20 and 60 and 60 and 120 seems to be closer than when we compare 20 and 120. On average, Bangd-U and Bangd-W provide larger values for the four periods. K-U and K-W report extraordinarily low values for the comparison of $m_1 = 20$ and $m_2 = 120$ in Table D.3. Agreement is greater during the two crisis periods (just before and after Lehman's demise). Regardless of the size of the window, yellow is the predominant color after the Lehman Brother bankruptcy. The system seems to improve with respect to the financial disaster of the previous period and ΔCoVaR moved to below average, which explains why several companies moved from red to yellow. Lesser agreement is found for the other two periods, 2005-2006, and 2011-2018, as shown in Figures D.1-D.3 in the Appendix. The statistics analyzing the agreement between rankings using windows of 20 and 120 days are the smallest. Furthermore, by varying the window size we do not observe any impact on the aggregated index, thus still leaving the identification of the best window size unresolved.

More interesting is the sensitivity analysis of daily TALIS[3] rankings on a company basis, i.e. we will analyze the agreement between TALIS[3] computed with windows of different sizes, m_1 and m_2. Now the two-way contingency tables will record the number of days that a company is ranked as red, yellow, amber or green using TALIS[3] computed with rolling statistic based on windows of size m_1 or m_2 during 1,032 available days before Lehman's bankruptcy. Figures D.4 – D.7 in Appendix D show Bangdiwala's (1988) agreement charts for four companies: BSC, FNM, LEH and MER. For each company, the left panel shows the comparison between the daily rankings of TALIS[3] based on windows of size $m_1 = 20$ and $m_2 = 60$, middle panel for size $m_1 = 60$ and $m_2 = 120$ and the right panel for windows of size $m_1 = 20$ and $m_2 = 120$. A larger level of agreement is found for windows of size 60 and 120, as can be seen by looking at the middle panel of the figures.

Agreement statistics are reported in Tables D.4-D.7 in Appendix D, namely unweighted (Bangd-U) and weighted (Bangd-W) Bangdiwala statistics along with the unweighted (K-U) and weighted (K-W) Kappa statistics. Weighted statistics are computed using Fleiss and Cohen`s (1973) weights. They all are common measures of agreement between the rankings based on two different sizes of the rolling window when classifying the 1,032 days into the four mutually exclusive categories: red, yellow, amber and green. Table D.4 shows the agreement statistics for BSC, Table D.5 for FNM, Table D.6 for LEH and finally, Table D.7 shows these statistics for MER. In the four tables the Bangd-W statistic shows a value above 0.48, which can be interpreted as a close to fair degree of agreement between daily TALIS[3] rankings based on windows of size 60 and 120 days. Lesser agreement is found while comparing the 20-day window against either 60 or 120-day windows. In general, for the four companies we observe the least agreement when comparing the 20-day versus the 120-day windows. We also note that the shorter the window, the more frequently a company is ranked as green. This can be seen by examining the 'path of rectangles' and how it deviates from the 45° diagonal line within the larger $N \times N$ square. We find some kind of bias through less systemic risk rankings when reducing the size of window, i.e, the shorter the window, the larger the number of days that a company is labelled as green. Closer analysis of the behavior of TALIS[3] when changing the size of the window used to calculate it leads us to conclude that a small window size has three main drawbacks: i) $MD^{s}_{t+1,m}\left(MD^{j}_{t+1,m}\right)$ will be highly dependent on unusually large movements in returns, ii) the thresholds, $\overline{Z}^{s}_{t+1,m}\left(\overline{Z}^{j}_{t+1,m}\right)$, will also change more rapidly, and iii) many windows might exhibit rolling statistics with zero value and no signal (information) being provided. TALIS[3] releases a signal when company losses are greater than the company's forecast VaR and this signal remains *on* for at least *m* days after it was emitted. Fast signal changes because of a short window size might lead to misleading company rankings. If the size of the window is too small, and therefore it has less memory, there is a higher chance of days without a signal because of zero-value rolling statistics, which will lead to the company being classified as green. In addition, a short window size will make TALIS^3 more dependent on unusual returns and will therefore lead to drastic changes in rankings. In contrast, if the window size is longer, the signal will be “smoothed out” and company rankings will be more stable.

5.3. TALIS^3 versus ΔCoVaR. A new dimension to risk?

We are interested in analyzing the extent to which TALIS^3 measures a different dimension of systemic risk that is not already captured by other indicators. TALIS^3 combines two dimensions of risk: ΔCoVaR which measures the marginal contribution of the institution to systemic risk; and the magnitude of company losses, (i.e. realized losses larger than the company VaR levels) which can be thought of as the amount of capital that the government would have to provide to bail out the financial institution. Greater company losses when in distress would imply a larger expected amount of capital to be provided by the government in the event of a bail out and therefore would increase the systemic risk of the company. This argument is in line with Acharya, et al. (2010) who state that an important part of the systemic risk of a company is its expected capital shortfall which contains information about size, leverage and interconnectedness, and thereby provides a market-based measure of a firm’s fragility.

Figures 14-16 show daily TALIS^3 and ΔCoVaR rankings for Citi Group, AIG and LEH from January 2, 2007 to September 15, 2008. On the x-axis we find the dates and, on the y-axis the rank of the company among the 73 companies available during the first part of the sample. In order to produce a more consistent ranking comparison avoiding sudden rank fluctuations we are going to smooth the signal released by both TALIS^3 and ΔCoVaR. For every day, we first compute the previous 60-day average of ΔCoVaR and TALIS^3 which is then used to calculate the daily position in terms of their systemic risk contribution among the 73 companies in the sample before Lehmann’s bankruptcy, and among the 58 companies remaining in the sample after Lehmann’s bankruptcy. Companies in the top positions, the most systemic, are those with rankings close to 1.

Figure 14.- TALIS^3 versus ΔCoVaR: Citi Group rankings out of 73. From January 2007 to September 2008.

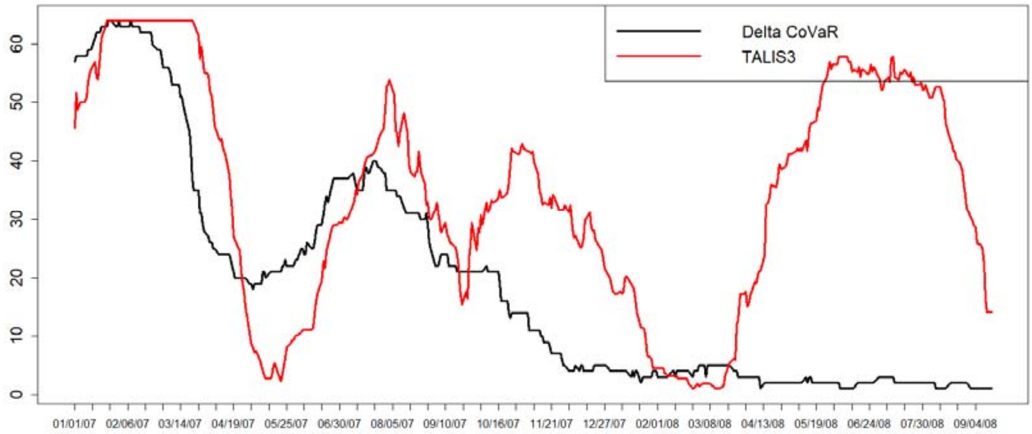
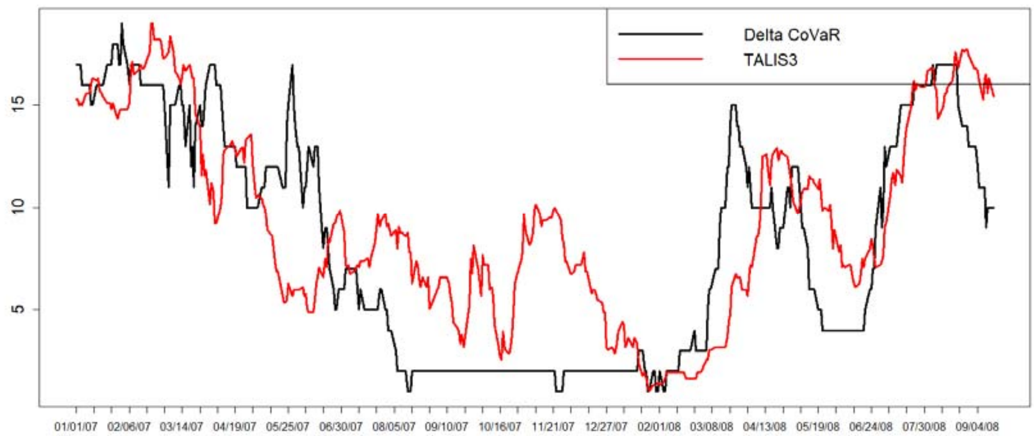
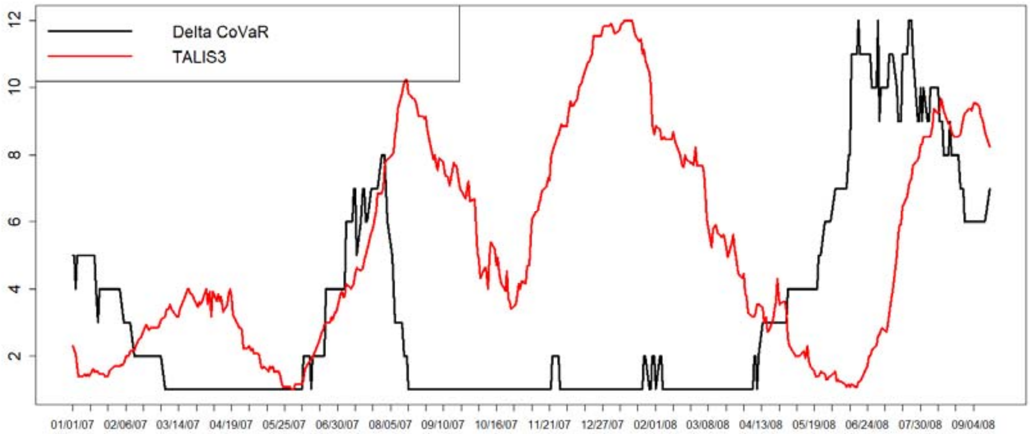


Figure 15.- TALIS^3 versus ΔCoVaR: AIG rankings out of 73. From January 2007 to September 2008.



In Figure 14 we can observe TALIS^3 and ΔCoVaR rankings for Citi Group. Both systemic risk measures show similar behavior, but for the April 2008-July 2008 period the TALIS^3 ranking first improves in March 2008, then plunges in July 2008. In Figure 15, AIG is ranked in the top 20 during this period. We find the largest difference between the two rankings from August 2007 to January 2008. Finally, Figure 16 shows the TALIS^3 and ΔCoVaR rankings for Lehman Brothers. From August 2007 to March 2008, the difference between both rankings is significant.

Figure 16- TALIS[3] versus ΔCoVaR: Lehman Brothers rankings out of 73. From January 2007 to September 2008



.

A proper way of interpreting TALIS[3] is like a signal extraction approach developed to identify turnings points in the financial company's state. The signal is "on" if one of the two indicators (ΔCoVaR or capital shortfall) crosses the threshold, changing the company's category, and "off" otherwise, keeping the company in the same category. In this sense, policy makers should be concerned when any of the two risks signals is "on". To illustrate this, we can follow the TALIS[3] and ΔCoVaR ranking of LEH during the period August 2007 to February 2008 shown in Figure 16. During that period, ΔCoVaR ranks LEH in the top 3 most systemic companies. Nonetheless, the TALIS[3] ranking of LEH swings from the top 12 to the top 3. While the ΔCoVaR ranking of LEH does not change (top 3), between October 2007 and February 2008 the TALIS[3] labels LEH as green over several days (see Figure 6 above) and this is what explains the drop in the contribution of LEH to systemic risk according to TALIS[3] during that period. The signal that produces the change in the TALIS[3] ranking comes from a decreasing capital shortfall for LEH, and not from ΔCoVaR. Equity market prices are increasing in response to new resources or expected future gains and Lehman's stock rebounded in the fourth quarter of 2007, as global equity markets reached new highs and fixed-income asset prices staged a temporary recovery. According to TALIS[3], after February 2008 LEH start rising to the top of the systemic risk contributors, as an indicator of the distress that led LEH to finally collapse in September 2008. It seems that TALIS[3], combining multiple measures of the company's contribution to systemic risk, amplifies a signal that might be easier for policy maker to interpret. TALIS[3] can be taken as a simple and effective way to produce a robust measure of systemic risk.

TALIS[3] might also have some key implications in terms of capital procyclicality. Ideally, the increase in regulatory capital should take place during the run-up phase to the crisis and not at the outbreak of the crisis. ΔCoVaR alone is unlikely to comply with this property. Indeed, during a crisis market tail risk increases very quickly and so would a ΔCoVaR-based capital charge. Gehrig and Iannino (2017) suggest that systemic measures such as ΔCoVaR and SRISK exhibit procyclicality, suggesting a growing reduction in resilience when we would expect a long-run decline in the trajectories of these measures, reflecting an overall enhancement in the safety and soundness of the financial system because the resiliency is enhanced. ΔCoVaR is ignoring the reaction of the central bank, neglecting the potential measures of the regulatory authorities (micro- and macro prudential regulations) and neglecting the potential reactions of financial institutions (for example, withdrawal of interbank exposures and raising capital) that $TALIS^3$ takes into account while monitoring the magnitude of losses beyond the VaR of the company. By construction, $TALIS^3$ is less procyclical than ΔCoVaR because it will issue quasi-real time early signals when either of the two indicators (ΔCoVaR or capital shortfall) crosses the threshold. For example, as $TALIS^3$ does not only depend on ΔCoVaR, but also on capital shortfalls, a company can change from red to green because the magnitude of its losses beyond VaR moves below its historical threshold. This change in the $TALIS^3$'s ranking might lead to policy makers reducing the burden of a large capital requirement for the company, which will contribute to improving its balance sheet.

6. Conclusions

We introduced a criterion for evaluating the systemic risk contribution of companies accounting for their distress and their impact on the market. Building on our estimates, we derive a new color-based systemic indicator, the TrAffic LIght System for Systemic Stress ($TALIS^3$), inspired by the loss functions frequently adopted in backtesting analyses of the market and systemic risk. We are thus able to identify more precisely the companies and periods of distress. This new indicator provides a company ranking system identifying various levels of systemic risk and providing a color-based code similar to the Basel Committee's TrAffic LIght adopted to classify financial institution market risk violations. Our measure of aggregate systemic risk is shown to be especially valuable for predicting the lower conditional quantiles of financial markets suggesting the potential of aggregate $TALIS^3$ to be used as an intuitive, effective and powerful early warning system for financial crises, which could be used for the monetary authorities while designing their macroprudential policy strategies to accomplish financial stability.

$TALIS^3$ is a promising avenue for future work. It is a novel technique that extracts and processes signals that are generated by underlying risk metrics, providing a visual and automatic classification of financial companies, according to their marginal contribution to the systemic risk. In this sense, $TALIS^3$ permits to manage two individual signals in order to provide more information than the original one individually does. Although, in this paper, TALIS^3 monitors the evolution of a company's VaR and ΔCoVaR, TALIS^3 could be easily adapted to extract signals from any other risk metrics as MES or SRISK, which could potentially add to our understanding of how best to measure and forecast systemic risk.

References

Acharya, V., Pedersen, L.H., Philippon, T., and Richardson, M., 2010. "Measuring Systemic Risk", Working paper, New York University.

Adrian, T., and Brunnermeier, M.K., 2016. "CoVaR", American Economic Review, 106 (7), 1705-1741.

Aikman, D., Kiley, M., Lee, S.J., and Palumbo, M.G., 2017. "Mapping heat in the US financial system", Journal of Banking and Finance, 81, 36-64.

Allen, L., Bali, T.G., and Tang, Y., 2010. "Does systemic risk in the financial sector predict future economic downturns?", Working paper, SSRN.

Alexander, C.O., 2001. "Orthogonal GARCH in Mastering Risk", Volume 2 (C.O. Alexander, Ed.), Financial Times-Prentice Hall.

Alexander, C.O. and Chibumba, A., 1996. "Multivariate Orthogonal Factor GARCH", University of Sussex, Discussion Papers in Mathematics.

Bangdiwala, S.I., and Shankar, V., 2013. "The agreement chart", BMC Medical Research Methodology, 13, 97. http://doi.org/10.1186/1471-2288-13-97.

Bangdiwala, S.I., 1988. "The Agreement Chart", Department of Biostatistics, University of North Carolina at Chapel Hill, Institute of Statistics Mimeo Series No. 1859, http://www.stat.ncsu.edu/information/library/mimeo.archive/ISMS_1988_1859.pdf

Barone-Adesi G., Bourgoin, F., and Giannopoulos, K., 1998. "Don't Look Back", Risk, 11, August, 100-104.

Barone-Adesi G. and Giannopoulos, K., 1996. "A Simplified Approach to the Conditional Estimation of Value at Risk", Futures and Options World, October, 68-72.

Barone-Adesi G., Giannopoulos, K. and Vosper, L., 1999. "VaR without Correlations for non-linear Portfolios", Journal of Futures Markets, 19, August, 583-602.

Benoit, S., Colliard, J.E., Hurlin, C., and Pérignon, C., 2017. "Where the risks lie: A survey on systemic risk", Review of Finance, 21(1), 109-152.

Billio, M., Getmansky, M., Lo, A.W., and Pelizzon, L., 2012. "Econometric measures of systemic risk in the finance and insurance sectors", Journal of Financial Economics 104, 535-559.

Bisias, D., Flood, M., Lo, A. W., and Valavanis, S., 2012. "A survey of systemic risk analytics, Annual Review of Financial Economics", 4(1), 255-296.

Brownlees, C., and Engle, R., 2011. "Volatility, Correlation and Tails for Systemic Risk Measurement", Working paper, New York University of CoVaR.

Caporin, M., 2008. "Evaluating value-at-risk measures in the presence of long memory conditional volatility", Journal of Risk, 10(3), 79-110

Caporin, M., and McAleer, M., 2012. "Do we really need both BEKK and DCC?" A tale of two multivariate GARCH models", Journal of Economic Surveys, 26(4), 736-751.

Chao, S.K., Härdle, W.K., and Wang, W., 2012. "Quantile regression in risk calibration".

Cicchetti, D., and Allison, T., 1971. "A new procedure for assessing reliability of scoring EEG sleep recordings", The American Journal of EEG Technology, 11, 101-109.

Engle, R.F., 2002. "Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroscedasticity models", Journal of Business and Economic Statistics, 20, 339-350.

Engle, R.F., and Kroner, K.F., 1995. "Multivariate simultaneous generalized ARCH", Econometric theory, 11(1), 122-150.

Exchange, C.B.O., 2009. "The CBOE volatility index-VIX". White Paper, 1-23.

Fleiss, J.L. and Cohen, J., 1973. "The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability", Educational and Psychological Measurement, 33, 613-619.

Gehrig, T., and Iannino, M.C., 2017. "Did the Basel Process of Capital Regulation Enhance the Resiliency of European Banks?", CEPR Discussion Paper No. DP11920, March. Available at SSRN: https://ssrn.com/abstract=2937771

Girardi, G., and Ergün, A.T., 2013. "Systemic risk measurement: Multivariate GARCH estimation of CoVaR", Journal of Banking and Finance, 37, 3169-3180.

Glosten, L.R., Jagannathan, R., and Runkle, D.E., 1993. "Relationship between the expected value and the volatility of the nominal excess return on stocks", The Journal of Finance, 48(5), 1779-1801.

Huang, X., Zhou, H., and Zhu, H., 2009. "A framework for assessing the systemic risk of major financial institutions", Journal of Banking and Finance, 33, 2036–2049.

Kliesen, K.L., Owyang, M.T., and Vermann, E.K., 2012. "Disentangling diverse measures: A survey of financial stress indexes". Federal Reserve Bank of St. Louis Review, 94(5), 369-397.

Koenker, R., and Machado, J.A. (1999). “Goodness of fit and related inference processes for quantile regression”. Journal of the american statistical association, 94(448), 1296-1310.

Kroner, K.F., and Ng, V.K., 1998. “Modeling asymmetric comovements of asset returns”. The review of financial studies, 11(4), 817-844.

Lopez, J.A., 1999. “Methods for evaluating value-at-risk estimates”. Economic review, 2, 3-17.

Mainik, G., and Schaanning, E., 2014. “On dependence consistency of CoVaR and some other systemic risk measures”. Statistics & Risk Modeling, 31(1), 49-77.

Manamperi, N., 2015. “A comparative analysis on US Financial stress indicators”. International journal of Economics and Financial issues, 5(2), 613-623.

Oet, M. V., Dooley, J. M., and Ong, S. J., 2015. “The financial stress index: Identification of systemic risk conditions”, Risks, 3(3), 420-444.

Oet, M-V., Eiben, R., Bianco, T., Gramlich, D., and Ong, S-J., 2011. “Financial stress index: Identification of systemic risk conditions”. Federal Reserve Bank of Cleveland Working Paper, (11-30), 1-73.

Segoviano, M.A., and Goodhart, C.A.E., 2009. “Banking stability measures”, International Monetary Fund, 627.

Silva, W., Kimura, H., and Sobreiro, V.A., 2017. “An analysis of the literature on systemic financial risk: A survey”, Journal of Financial Stability, 28, 91-114.

Silvennoinen, A., and Teräsvirta, T., 2009. “Multivariate GARCH models”. In Handbook of financial time series (pp. 201-229). Springer, Berlin, Heidelberg.

Zhou, C., 2010. “Are banks too big to fail? Measuring systemic importance of financial institutions”, International Journal of Central Banking, 6, 205–250.

WEB APPENDICES

APPENDIX A. Multivariate GJR models and Filtered Historical Simulation

This Appendix provides details on the implementation of the Multivariate GJR-GARCH models we consider, the Dynamic Conditional Correlation of Engle (2002), the BEKK model of Engle and Kroner (1995), and the OGARCH model of Alexander (2001) and Alexander and Chibumba (1996). For each model, we report the model equations, the steps for estimation, and the steps for the generation of future paths with Filtered Historical Simulation. Note that we adopt in all cases a common rolling approach for model estimation (window size: 1000 observations), forecasting, with the window size, and forecast horizon defined in the main paper. Furthermore, the paper includes also details on the number of replications used when adopting a simulation approach to recover the future evolution of the system and company returns.

In the following subsections, we describe the implementation of the three models. Our discussion is detailed in the case of the DCC model of Engle (2002) where we also indicate the steps for the CoVaR evaluation.

A.1. Dynamic Conditional Correlation (DCC)

Partly borrowing from Girardi and Ergün (2013), GE thereafter, we adopt a 3-steps strategy to compute the expected Conditional-Value-at-Risk (CoVaR). We also recall that we embrace the GE definition of CoVaR, see Section 2 of the paper. Our procedure mimics the GE approach in the first step, and deviates from step 2 onward. In details, GE estimate a bivariate DCC model to compute the conditional mean and variance, and they were assuming either a Gaussian or a Skewed-t distribution for the returns. They were also using the distributional assumption the computation of the CoVaR. In our step 1, we also estimate a DCC model. However, we do not make a distributional assumption. On the contrary, we recover the conditional correlations, we compute the uncorrelated returns, and then we used a resampling framework to generate the future evolution of the innovations. Our proposal falls within the Filtered Historical Simulation (FHS) approach of Barone-Adesi and Giannopoulos (1996), Barone-Adesi et al. (1998), and Barone-Adesi et al. (1999). Moreover, as we do not make any distributional assumption, we estimate model parameters by Quasi Maximum Likelihood.

We now describe in detail the procedure we follow to estimate the CoVaR for the system and the *j*-th financial institution for time T+1. Note that we repeat the steps across all financial institutions and for all rolling samples.

Step 1 – Model estimation and filtering. We first estimate a univariate AR(1) model on the returns of the system and of the *j-th* company to account for possible serial dependence (which, if present, will be of minor relevance):

$$R_t^i = \mu_t^i + \varepsilon_{i,t}, \quad \varepsilon_{i,t} = \sigma_{i,t} z_{i,t}; \quad i = s, j \tag{A1}$$

where $\mu_t^i = \phi_0 + \phi_1 R_{t-1}^i$, and $z_{i,t}$ is assumed to be an i.i.d. sequence with zero mean and unit variance. For the conditional variances we adopt the standard GJR- GARCH (1,1) specification:

$$\sigma_{i,t}^2 = \omega_0^i + \alpha_0^i \varepsilon_{i,t-1}^2 + \gamma_0^i \varepsilon_{i,t-1}^2 I(\varepsilon_{i,t-1} < 0) + \beta_0^i \sigma_{i,t-1}^2 \tag{A2}$$

where $I(.)$ is an indicator function taking a unit value if the condition in parentheses is true.

Given the estimated conditional variances, we recover the variance standardized innovations, $z_{i,t} = \frac{R_{i,t} - \mu_{i,t}}{\sigma_{i,t}}$. Then, following Engle (2002), we model the conditional correlation of Z_t,, the bivariate vector of the system and company *j* innovations, as follows:

$$\Gamma_t = diag(Q_t)^{-1/2} Q_t diag(Q_t)^{-1/2},$$

$$Q_t = (1 - \alpha_1 - \beta_1)\overline{Q} + \alpha_1 (Z_{t-1} Z'_{t-1}) + \beta_1 Q_{t-1} \tag{A3}$$

where $\overline{Q}$ is the unconditional correlation matrix of Z_t, $diag(Q_t)$ is a matrix operator that creates a diagonal matrix with the element on the main diagonal of Q_t, and Γ_t is the conditional correlation matrix, that comes from a standardization of the Q_t matrices. We highlight that the dynamic in equation (A3) corresponds to a DCC-type recurrence adapted to the variance standardized residuals z_t.

Finally, given the fitted conditional correlations, we recover uncorrelated historical shocks as:

$$Z_t^D = \Gamma_t^{-1/2} Z_t \; . \tag{A4}$$

where $\Gamma_t = \left(\Gamma_t^{1/2}\right)^T \Gamma_t^{1/2}$.

Step 2 – Resampling and forecasting. We then proceed with the FHS step based on the uncorrelated shocks $\left\{Z_t^D\right\}_{t=1}^T$, from which we draw random samples, called $Z_{l,T+1}^D,\ for\ l = 1,\dots,B$ (B=10000). From each draw, we simulated returns for t = $T+1$ as follows:

$$\begin{aligned} E_{l,T+1} &= D_{T+1}\Gamma_{T+1}^{1/2}Z_{l,T+1}^D = D_{T+1}Z_{l,T+1} \\ R_{l,T+1} &= \mathrm{M}_{T+1} + E_{l,T+1} \end{aligned} \quad \text{(A5)}$$

where M_{T+1} is the vector of the predicted mean returns for the system s and the *j-th* financial institution obtained from the AR(1) model, D_{T+1} is the diagonal matrix containing the predicted conditional volatilities of the system and the *j-th* financial institution as provided by the GJR-GARCH model, and $\Gamma_{T+1}^{1/2}$ is the predicted conditional correlation for time T+1. We continue this simulation for K days (out-of-sample period), and repeat it for $l = 1, \dots, 10000$ simulated shocks.

Note that the predicted mean, variances and correlations do not depend on the index for the random sample as we restrict ourselves to the one-step-ahead forecast. If the forecast horizon will be greater than 1 period, then all the generated paths for the mean, variance and correlation will also depend on the random sample index.

Step 3 – Conditional Value-at-Risk evaluation. Given a collection of B bootstrap simulations of the company and the systems 1-day ahead returns, we proceed as follows for the evaluation of the CoVaR:

3.A. We compute the *Value-at-Risk* for company j at the α confidence level as $VaR_{\alpha,T+1}^j = inf\left\{x \in \left\{R_{l,T+1}^j\right\}_{l=1}^B : \hat{F}_{R_{T+1}^j}(x) \geq 1-\alpha\right\}$, where $\hat{F}_{R_{T+1}^j}(x)$ is the empirical CDF of the $T+1$ company returns estimated by using the B simulated values for R_{T+1}^j;

3.B. We compute the CoVaR of the system s on the *distress* state of institution j, i.e. when $R_{T+1}^j \leq VaR_{\alpha,T+1}^j$, is computed as $CoVaR_{\alpha,T+1}^{s|j} = inf\left\{x \in \mathcal{A}: \hat{F}_{R_{T+1}^s \in \mathcal{A}}(x) \geq 1-\alpha\right\}$, where the set $\mathcal{A}$ includes the returns of the system when the company is in distress, i.e. $\mathcal{A} = \left\{R_{l,T+1}^s \in \left\{R_{l,T+1}^s, R_{l,T+1}^j\right\}_{l=1}^B : R_{l,T+1}^j \leq VaR_{\alpha,T+1}^j\right\}$ and $\hat{F}_{R_{T+1}^s \in \mathcal{A}}(x)$ is the empirical CDF of the system conditional of the company being in distress and evaluated only with the returns belonging to the set $\mathcal{A}$;

3.C. We compute the CoVaR of the system s on the *benchmark* state of institution *j*, i.e. when $m_{T+1}^{j} - s_{T+1}^{j} \leq R_{T+1}^{j} \leq m_{T+1}^{j} + s_{T+1}^{j}$, where m_{T+1}^{j} and s_{T+1}^{j} are, respectively, the sample mean and sample standard deviations of the 1-day ahead simulated returns of the company. The CoVaR equals $CoVaR_{\alpha,T+1}^{s|b^j} = inf\left\{x \in \mathcal{B}: \hat{F}_{R_{T+1}^{s} \in \mathcal{B}}(x) \geq 1 - \alpha\right\}$, where the set $\mathcal{B}$ includes the returns of the system when the company is in a benchmark state, i.e. $\mathcal{B} = \left\{R_{l,T+1}^{s} \in \left\{R_{l,T+1}^{s}, R_{l,T+1}^{j}\right\}_{l=1}^{B}: m_{T+1}^{j} - s_{T+1}^{j} \leq R_{T+1}^{j} \leq m_{T+1}^{j} + s_{T+1}^{j}\right\}$ and $\hat{F}_{R_{T+1}^{s} \in \mathcal{B}}(x)$ is the empirical CDF of the system conditional of the company being in a benchmark state and evaluated only with the returns belonging to the set $\mathcal{B}$;

3.D. Finally, we forecast the systemic risk contribution of a particular institution *j*, at time *T* with a horizon 1, as $\Delta CoVaR_{\alpha,T+1}^{s|j} = CoVaR_{\alpha,T+1}^{s|j} - CoVaR_{\alpha,T+1}^{s|b^j}$.

A.2. BEKK

We focus on a variation of the BEKK model of Engle and Kroner (1995), namely, the model introduced by Kroner and Ng (1998) that extends the BEKK with the introduction of an asymmetric component. The Asymmetric BEKK (ABEKK) specification we consider is:

$$\Sigma_t = CC' + A\varepsilon_{t-1}\varepsilon_{t-1}'A' + B\Sigma_{t-1}B' + G\eta_{t-1}\eta'_{t-1}G' \quad \text{(A6)}$$

where the innovations are obtained as in equation (A1), C is a lower triangular matrix, and A, B and *G* are square matrices (not necessarily symmetric) and $\eta_t = \left[\eta_{st}, \eta_{jt}\right]'$ is a 2 by 1 vector such as $\eta_{it} = \max\left\{0, -\varepsilon_{it}\right\}$, $i = s, j$. The conditional covariance matrices are positive definite by construction, and the conditional variances are positive, regardless of the parameter signs. As suggested in Engle and Kroner (1995), to avoid observationally equivalent parametrizations, we set the upper left elements of *A*, *B* and *G* to be strictly positive.

Compared to the DCC case, Step 3 is equivalent, and the differences are in Step 1 and Step 2. In Step 1, we estimate the ABEKK model by Quasi maximum likelihood and then recover the variance standardized and uncorrelated historical shocks as $Z_t^D = \Sigma_t^{-1/2}\varepsilon_t$, i.e. by removing in a single stage both the conditional heteroscedasticity and the dynamic correlation (see Caporin and McAleer, 2012, on the presence of dynamic correlation within a BEKK framework). Then, when generating the future returns paths in Step 2, we modify (A5) as follows:

$$E_{l,T+1} = \Sigma_{T+1}^{1/2} Z_{l,T+1}^{D}$$

$$R_{l,T+1} = M_{T+1} + E_{l,T+1}$$

where we account for the different model adopted for the conditional covariances.

A.1.3. OGARCH

Finally, we adopt the OGARCH specification of Alexander (2000). In this model, similarly to the BEKK case, the only differences with respect to the DCC steps of section A.1.1. are in Step 1 and Step 2.

In the OGARCH model, the observed time series are first linearly transformed to a set of uncorrelated time series by using principal component analysis. Step 1 proceeds as follows. Given the innovations (residuals) of (A1), we compute their unconditional covariance $\sum$. We decompose the covariance exploiting the role of volatilities and correlations, $\Sigma = V\Gamma^{O}V$, where Γ^{O} is the unconditional correlation matrix and V a diagonal matrix of unconditional volatilities. We further decompose Γ^{O} as follows, $\Gamma^{O} = D\Delta D'$, where D is the matrix of eigenvectors and Δ the diagonal matrix of ordered eigenvalues.

Building on the previous matrices, we compute the principal components on the variance standardized innovations as follows

$$\eta_t = D'V^{-1}\varepsilon_t \, .$$

Note that the principal components are orthogonal and with covariance matrix equal to Δ. On the principal components we then fit a GARCH model under targeting, as the unconditional variance is known (and equal to Δ_i):

$$\sigma^2_{(k,t)} = \left(1-\alpha^k_0 - \beta^k_0 - \gamma^k_0/2\right)\Delta_k + \alpha^k_0\eta^2_{(k,t-1)} + \gamma\eta^2_{(k,t-1)}I\left(\eta_{(k,t-1)}<0\right) + \beta^k_0\sigma^2_{(k,t)}. \quad \text{(A7)}$$

Given the estimated conditional variances of the components, we recover the variance standardized and uncorrelated innovations as

$$Z^D_{k,t} = \sigma^{-1}_{k,t}\eta_{k,t-1}.$$

When focusing on the generation of future paths, i.e. Step 2, we use the following equations in place of(A5):

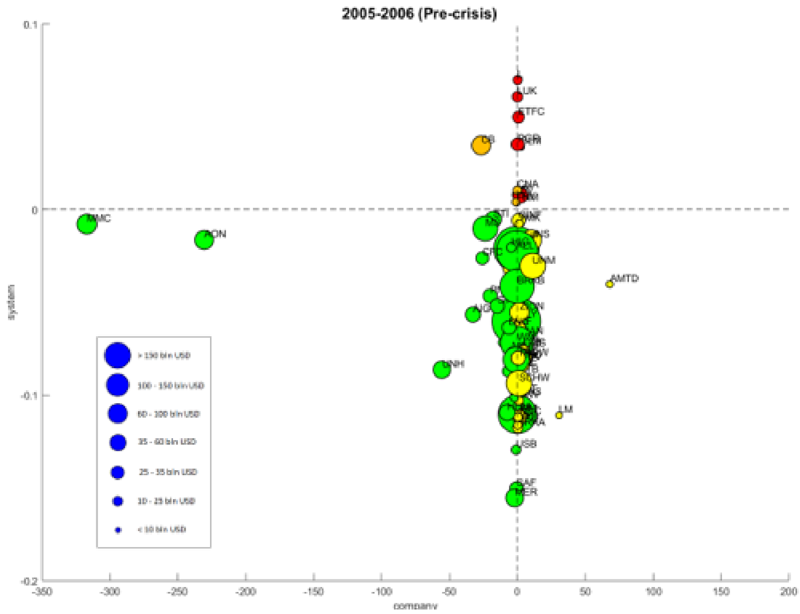
$$E_{l,T+1} = VD\Upsilon_{T+1}Z^{D}_{l,T+1}$$

$$R_{l,T+1} = M_{T+1} + E_{l,T+1}$$

where Υ_{T+1} is a diagonal matrix containing the predicted volatilities for the principal components.

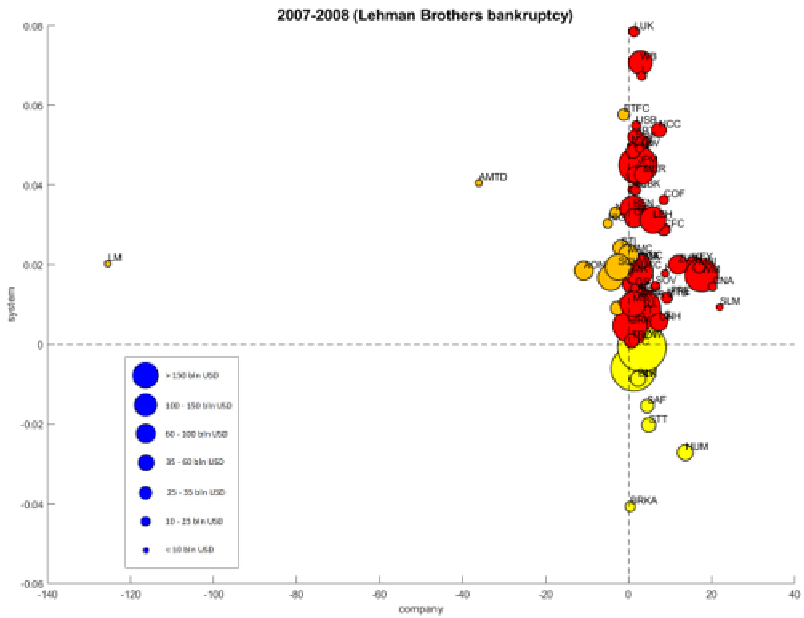
APPENDIX B. Additional Figures of TALIS[3] at work

Figure B.1.- Scatter plot: January 01, 2005 to December 31, 2006



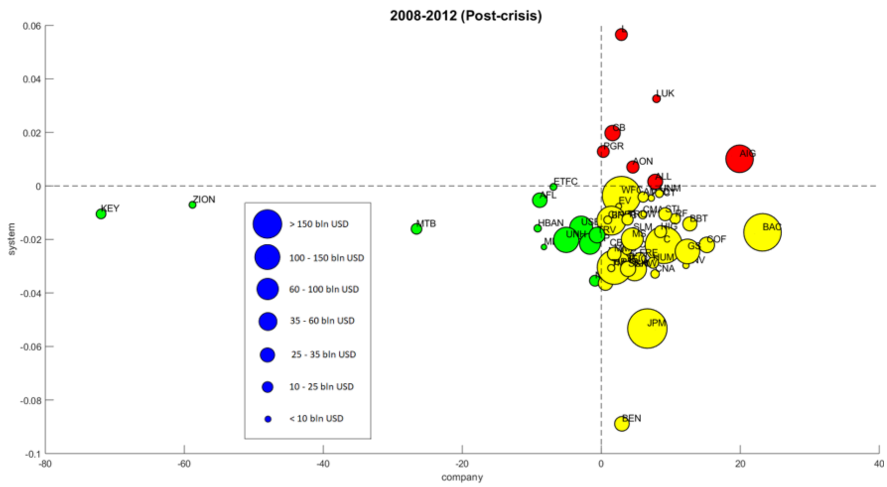
Note: Vertical (horizontal) axis represents $\left[MD^{s|j}_{t+1,60} - \bar{z}^{s|j}_{t+1,60}\right]$ ($\left[MD^{j}_{t+1,60} - \bar{z}^{j}_{t+1,60}\right]$), i.e. the system's (company's) capital shortfall in stress states that exceeds the system's (company's) threshold based on the own system's (company's) history of capital shortfalls (high stress state (+), low stress state (-)). Northeast (+ system, + company), red; Northwest (+, -), amber; Southwest (-, -), green; Southeast (- , +), yellow. The size of the circles is based on the company capitalization in the analyzed period.

Figure B.2.- Scatter plot: January 01, 2007 to September 15, 2008



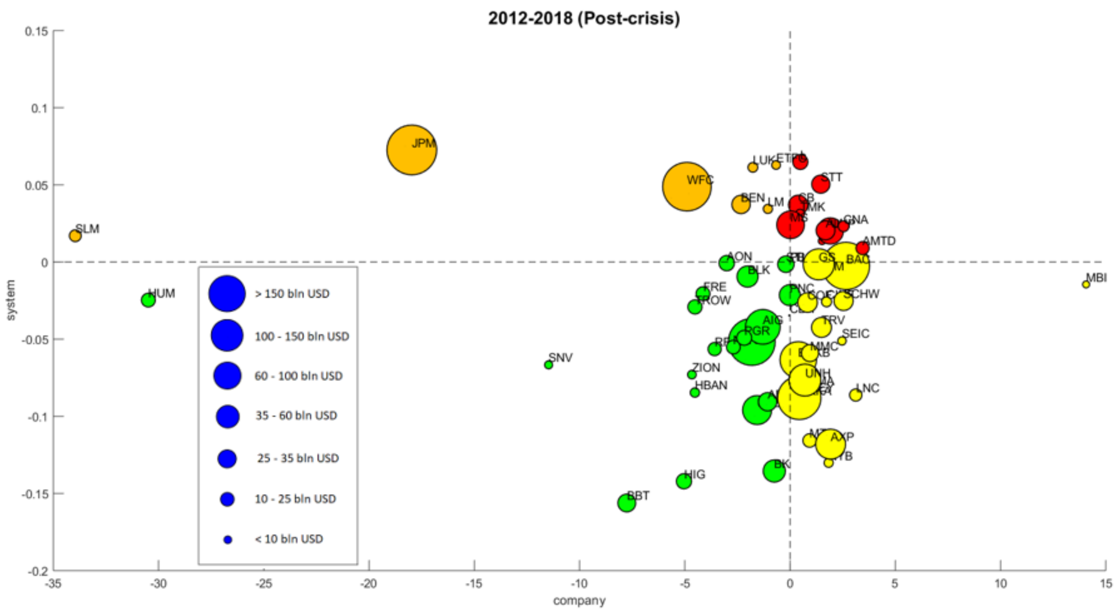
Note: Vertical (horizontal) axis represents $\left[MD_{t+1,60}^{s|j} - \bar{z}_{t+1,60}^{s|j}\right]$ $\left(\left[MD_{t+1,60}^{j} - \bar{z}_{t+1,60}^{j}\right]\right)$, i.e. the system's (company's) capital shortfall in stress states that exceeds the system's (company's) threshold based on the own system's (company's) history of capital shortfalls (high stress state (+), low stress state (-)). Northeast (+ system, + company), red; Northwest (+, -), amber; Southwest (-, -), green; Southeast (- , +), yellow. The size of the circles is based on the company capitalization in the analyzed period.

Figure B.3.- Scatter plot: September 15, 2008 to July 25, 2012



Note: Vertical (horizontal) axis represents $\left[MD_{t+1,60}^{s|j} - \bar{z}_{t+1,60}^{s|j}\right]$ $\left(\left[MD_{t+1,60}^{j} - \bar{z}_{t+1,60}^{j}\right]\right)$, i.e. the system's (company's) capital shortfall in stress states that exceeds the system's (company's) threshold based on the own system's (company's) history of capital shortfalls (high stress state (+), low stress state (-)). Northeast (+ system, + company), red; Northwest (+, -), amber; Southwest (-, -), green; Southeast (- , +), yellow. The size of the circles is based on the company capitalization in the analyzed period.

Figure B.4.- Scatter plot: July 25, 2012 to the end of the sample.



Note: Vertical (horizontal) axis represents $\left[MD^{s|j}_{t+1,60} - \bar{z}^{s|j}_{t+1,60}\right]$ $\left(\left[MD^{j}_{t+1,60} - \bar{z}^{j}_{t+1,60}\right]\right)$, i.e. the system's (company's) capital shortfall in stress states that exceeds the system's (company's) threshold based on the own system's (company's) history of capital shortfalls (high stress state (+), low stress state (-)). Northeast (+ system, + company), red; Northwest (+, -), amber; Southwest (-, -), green; Southeast (- , +), yellow. The size of the circles is based on the company capitalization in the analyzed period.

APPENDIX C. Additional Tables of Robustness Analysis

Table C.1a. - Summary statistics for $\Delta CoVaR^{s|j}_{5\%,t}$ for all 73 US financial institutions (before Lehman Brothers bankruptcy) and for institutions by industry group. (BEKK)

	Mean	Std TS	Std CS	Max	Min
Overall	-2.00	1.52	0.04	0.76	-9.85
Depositories	-2.12	1.59	0.04	1.09	-10.17
Others	-1.91	1.48	0.04	0.52	-9.79
Insurance	-1.77	1.36	0.04	0.40	-9.07
Brokers-Dealers	-2.45	1.81	0.05	1.10	-11.07

Note: Sample 01/03/2000 - 09/15/2008

Table C.1b - Summary statistics for $\Delta CoVaR^{s|j}_{5\%,t}$ for all 73 US financial institutions (before Lehman Brothers bankruptcy) and for institutions by industry group. (OGARCH)

	Mean	Std TS	Std CS	Max	Min
Overall	-1.99	1.51	0.04	0.89	-9.18
Depositories	-2.16	1.71	0.05	0.88	-10.61
Others	-1.93	1.49	0.04	1.25	-9.41
Insurance	-1.73	1.30	0.04	1.00	-7.35
Brokers-Dealers	-2.29	1.48	0.04	-0.19	-8.69

Note: Sample 01/03/2000 - 09/15/2008

Table C.2a. - Summary statistics for $\Delta CoVaR^{s|j}_{5\%,t}$ for all 58 US financial institutions (after Lehman Brothers bankruptcy) and for institutions by industry group. (BEKK)

	Mean	Std TS	Std CS	Max	Min
Overall	-3.18	2.93	0.06	1.28	-23.58
Depositories	-3.23	2.84	0.06	0.86	-22.45
Others	-3.18	2.99	0.06	1.18	-23.70
Insurance	-3.02	2.88	0.06	2.11	-23.49
Brokers-Dealers	-3.58	3.37	0.07	0.00	-28.34

Note: Sample 06/16/2008 - 01/15/2018

Table C.2b. - Summary statistics for $\Delta CoVaR^{s|j}_{5\%,t}$ for all 58 US financial institutions (after Lehman Brothers bankruptcy) and for institutions by industry group. (OGARCH)

	Mean	Std TS	Std CS	Max	Min
Overall	-3.33	3.02	0.06	0.93	-22.93
Depositories	-3.31	3.10	0.06	1.24	-23.91
Others	-3.33	2.89	0.06	0.50	-21.55
Insurance	-3.27	3.01	0.06	1.27	-22.72
Brokers-Dealers	-3.63	2.98	0.06	-0.78	-22.94

Note: Sample 06/16/2008 - 01/15/2018

APPENDIX.D. Additional Tables of Further robustness checks

Table D.1.- Test of correspondence. Size of the rolling window: $m_1 = 20$ and $m_2 = 60$.

	Bangd-U	Bangd-W	K-U	K-U_ASD	K-W	K-W_ASD
Pre-cris	0.22	0.50	0.23	0.06	0.30	0.09
Cris 1	0.27	0.76	0.07	0.06	0.17	0.13
Cris 2	0.57	0.69	0.29	0.11	0.25	0.14
Post-Cris	0.11	0.48	0.10	0.08	0.34	0.10

Note. Bang-U and Bang-W are the unweighted and weighted Bangdiwala statistics, respectively, K-U and K-U_ASD are the unweighted Kappa statistics and its asymptotic standard deviation, respectively and K-W and K-W_ASD are the weighted Kappa statistics and its asymptotic standard deviation, respectively. Bangdiwala's agreement statistic measures the degree of agreement above that expected by chance alone, between TALIS[3]'s rankings based on a different size of window. It has a maximum of 1 when agreement is perfect, 0 when agreement is no better than chance, and negative values when agreement is worse than chance. Other values can be roughly interpreted as: < 0.20- Poor, > 0.40 - Fair, > 0.60 – Moderate, > 0.80 - Good, 1 - Very good.

Table D.2.- Test of correspondence size of the rolling window: $m_1 = 60$ and $m_2 = 120$.

	Bangd-U	Bangd-W	K-U	K-U_ASD	K-W	K-W_ASD
P-cris	0.19	0.47	0.20	0.06	0.21	0.09
Cris 1	0.11	0.78	0.04	0.04	0.24	0.09
Cris 2	0.55	0.72	0.29	0.09	0.26	0.15
Pos-Cris	0.20	0.63	0.21	0.07	0.53	0.08

Note: see note table A.4.1

Table D.3.- Test of correspondence size of the rolling window: $m_1 = 20$ and $m_2 = 120$.

	Bangd-U	Bangd-W	K-U	K-U_ASD	K-W	K-W_ASD
P-cris	0.14	0.43	0.01	0.07	0.10	0.11
Cris 1	0.46	0.84	0.04	0.04	0.24	0.09
Cris 2	0.40	0.70	0.09	0.09	0.00	0.14
Pos-Cris	0.09	0.37	0.03	0.07	0.04	0.11

Note: see note table A.4.1

Table D.4.- Test of correspondence size of the rolling window: *m* = 20, 60, 120. Bearn Stearns (BSC)

	Bangd-U	Bangd-W	K-U	K-U_ASD	K-W	K-W_ASD
60-20	0.14	0.57	0.04	0.02	0.16	0.03
60-120	0.13	0.56	0.12	0.02	0.27	0.03
120-20	0.16	0.52	0.09	0.02	0.13	0.03

Note. Bang-U and Bang-W are the unweighted and weighted Bangdiwala statistics, respectively, K-U and K-U_ASD are the unweighted Kappa statistics and its asymptotic standard deviation, respectively and K-W and K-W_ASD are the weighted Kappa statistics and its asymptotic standard deviation, respectively. Bangdiwala's agreement statistic measures the degree of agreement above that expected by chance alone, between TALIS[3]'s rankings based on a different size of window. It has a maximum of 1 when agreement is perfect, 0 when agreement is no better than chance, and negative values when agreement is worse than chance. Other values can be roughly interpreted as: < 0.20- Poor, > 0.40 - Fair, > 0.60 – Moderate, > 0.80 - Good, 1 - Very good

Table D.5.- Test of correspondence size of the rolling window: *m* = 20, 60, 120. Federal National Mortgage Assn (FNM)

	Bangd-U	Bangd-W	K-U	K-U_ASD	K-W	K-W_ASD
60-20	0.19	0.56	0.07	0.02	0.11	0.03
60-120	0.17	0.65	0.17	0.02	0.41	0.03
120-20	0.22	0.61	0.13	0.02	0.28	0.03

Note: see note table A.4.4

Table D.6.- Test of correspondence size of the rolling window: *m* = 20, 60, 120. Lehman Brothers (LEH)

	Bangd-U	Bangd-W	K-U	K-U_ASD	K-W	K-W_ASD
60-20	0.13	0.48	0.02	0.02	0.13	0.03
60-120	0.10	0.48	0.08	0.02	0.19	0.03
120-20	0.12	0.47	0.03	0.02	0.13	0.03

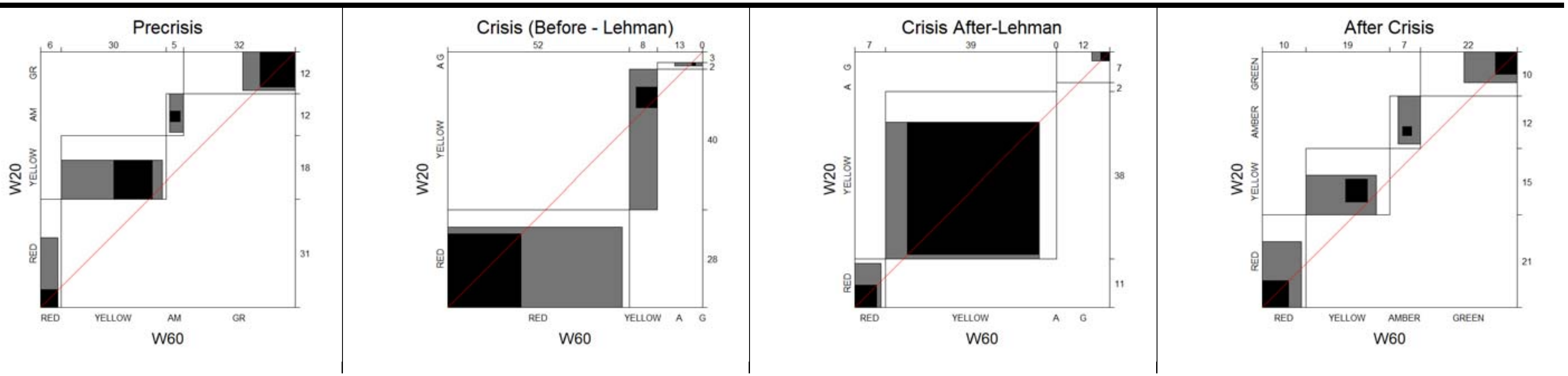
Note: see note table A.4.4

Table D.7.- Test of correspondence size of the rolling window: *m* = 20, 60, 120. Merrill Lynch (MER)

	Bangd-U	Bangd-W	K-U	K-U_ASD	K-W	K-W_ASD
60-20	0.11	0.44	-0.02	0.02	0.06	0.03
60-120	0.11	0.48	0.10	0.02	0.20	0.03
120-20	0.06	0.26	-0.04	0.02	-0.10	0.02

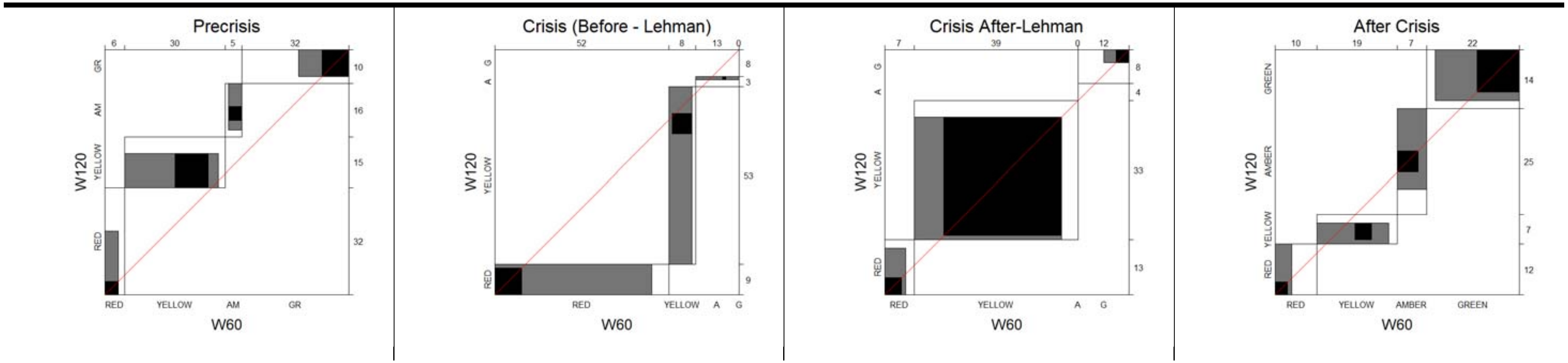
Note: see note table A.4.4

Figure D.1.- Bangdiwala's (1988) Observer Agreement Charts changing the Size of the rolling window: *m= 20* vs *m=60*



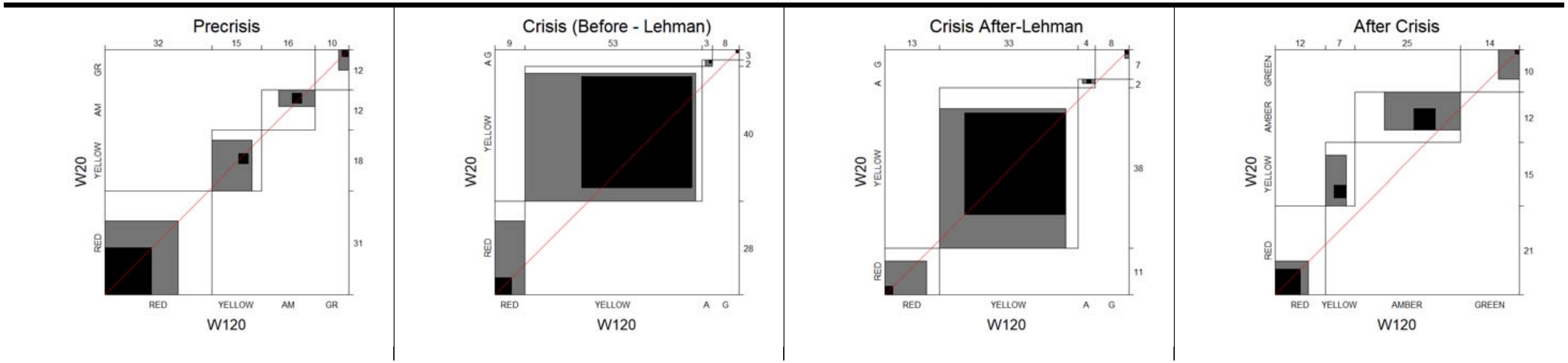
Note: the table shows Bangdiwala's (1988) agreement charts to compare TALIS^3(m) rankings with two different size of the rolling window m = *20* and m = *60* used to compute $MD^s_{t+1,m}\left(MD^j_{t+1,m}\right)$, the daily magnitude of the capital shortfall for the system (company), computed as the average capital shortfall during the last m days, and $\overline{z}^s_{t+1,m}\left(\overline{z}^j_{t+1,m}\right)$, the system's (company) threshold, defined in (11), which is the Maximum of two arguments: the median of the full expected losses history and the median of the last *m-day* window. The agreement between TALIS^3 is represented by the size and level of coincidence of the superposed white, grey and black rectangles inside the charts for each category, in our case, colors: red, yellow, amber (A) and green (G). To illustrate the information provided by these charts, consider the second panel that represents the agreement chart between TALIS^3*(m)* rankings based on m=*20* and m= *60* during 2007-2008 period. The size of the large square of this chart is *N* by *N*, where N is the, the total number of companies available in the sample during this first part of the sample. The horizontal (vertical) axis represents the number of companies ranked as red, yellow, A or G by TALIS^3 based on the 60 (20) day-rolling window. Inside this, we can find a set of four rectangles, one for each color. Looking at the first set of rectangles on the lower left vertex starting from (0, 0), which refers to the red color category, we can see a white big rectangle of size 52 by 28 (marginal totals). It reports that TALIS^3 (60) (20) ranked 52 (28) companies as red. The black square placed inside of side 21 by 21, reports that both TALIS^3(20) and TALIS^3(60) ranked the same 21 companies as red. Finally, the grey rectangle size considers partial agreement counting it differently with decreasing weights for cells further away from the diagonal cells. For example, in addition to the strict agreement of 21 companies ranked as red, we found that 29 out of 52 companies ranked as red by TALIS^3 (60), were ranked as yellow, 1 as amber and 1 as green by TALIS^3 (20). 2 out of 28 companies ranked as red by TALIS^3(20) were ranked as yellow and 5 as amber by TALIS^3(60). None of these 28 companies was ranked as green by TALIS^3(60). An additional matrix of weights must be considered. Two commonly-used patters are those based on equal spacing weights (Cicchetti and Alison, 1971) (CA) that are defined by $1-|i-j|\,/(r-1)$, where r is the number of columns/rows, and Fleiss and Cohen (1973)'s weights (FC), defined by $1-|i-j|^2\,/(r-1)^2$. Fleiss-Cohen weights attach greater importance to near disagreements. In the case of perfect agreement, black, grey and white rectangles are exactly equal. Lesser agreement is visualized by comparing the area of the blackened squares to the area of the rectangles, while model bias is visualized by examining the 'path of rectangles' and how it deviates from the 45° diagonal line within the larger $N \times N$ square. A detailed description of these agreement plots can be found in Bangdiwala and Shankar (2013). Left panel refers to the pre-crisis period (from January 2005 to December 2007), left-middle panel is for the crisis before Lehman demise (January 2007 to September 15, 2008), right-middle panel refers to the crisis after Lehman bankruptcy (from September 16, 2008 to July 24, 2011) and the right panel is for the after crisis period (from July 25, 2011 to February 2018).

Figure D.2.- Bangdiwala's (1988) Observer Agreement Charts. Size of the rolling window: *m* = 120 vs *m* =60



Note: See note in Figure D.1

Figure D.3.- Bangdiwala's (1988) Observer Agreement Charts changing. Size of the rolling window: m= 20 vs m=120



Note: See note in Figure D.1

Figure D.4.- Bangdiwala's (1988) Observer Agreement Charts. Size of the rolling window: m = 20, 60, 120. BSC

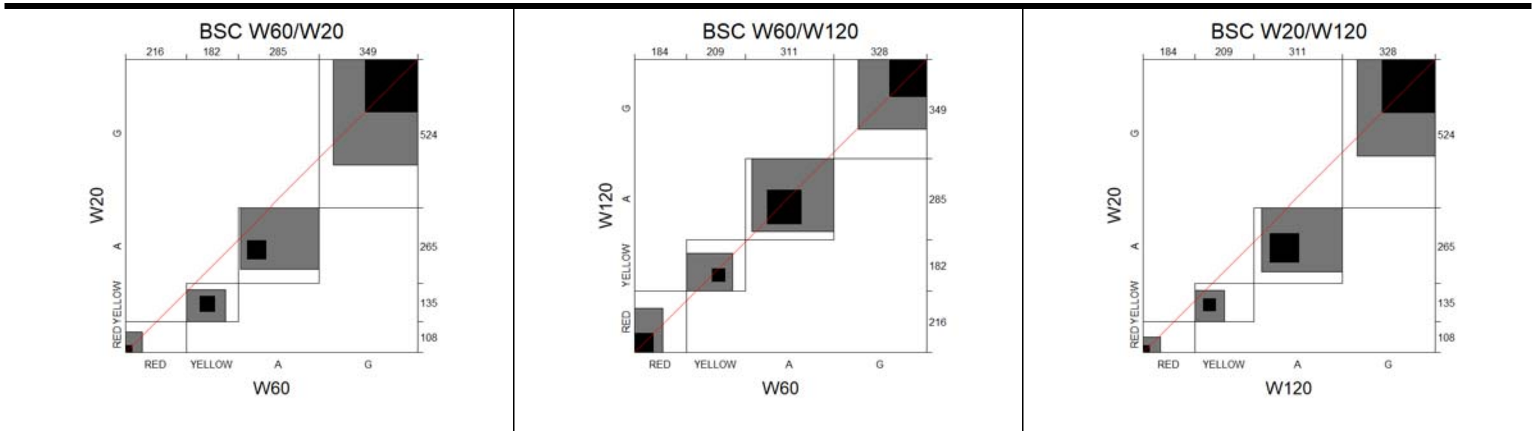


Figure D.5- Bangdiwala's (1988) Observer Agreement Charts. Size of the rolling window: m = 20, 60, 120. FNM

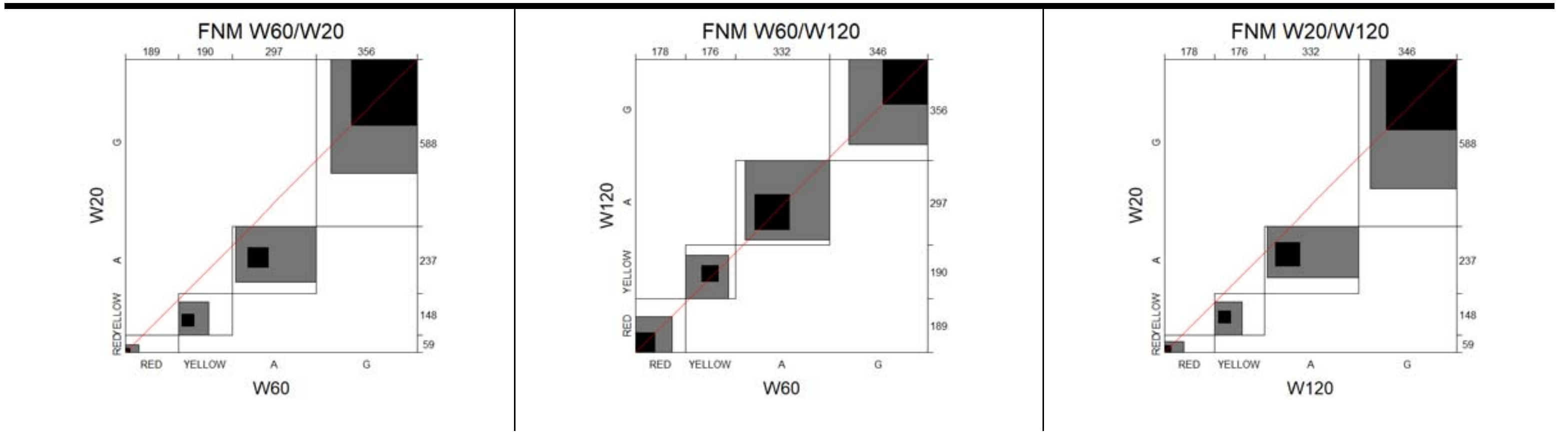


Figure D.6.- Bangdiwala's (1988) Observer Agreement Charts. Size of the rolling window: *m* = 20, 60, 120. LEH

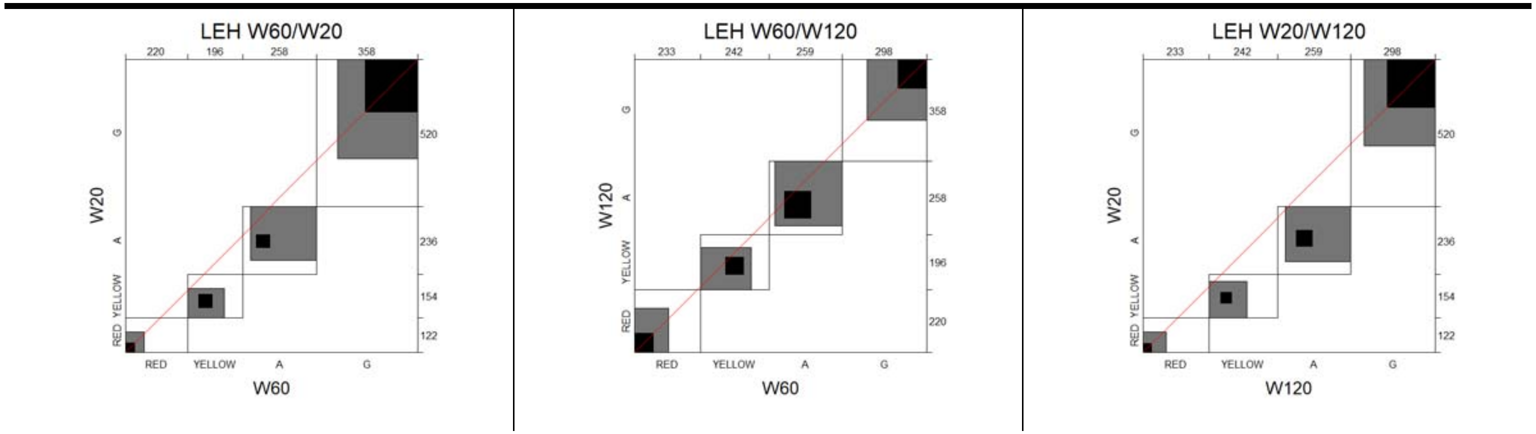


Figure D.7.- Bangdiwala's (1988) Observer Agreement Charts. Size of the rolling window: *m*= 20, 60, 120. MER