

Systemic risk measures: the simpler the better?

María Rodríguez-Moreno and Juan Ignacio Peña¹

Introduction

The financial system plays a fundamental role in the global economy as the middleman between agents who need to borrow and agents who are willing to lend or invest. As a consequence, it is naturally linked to all economic sectors and, therefore, if the financial system does not work properly, its problems have a strong impact on the real economy. We can see this in the deteriorating fundamental macro variables of the United States and Europe, including GDP growth rate, unemployment rates and government deficits.

Many of these problems come from events related to systemic risk, extending from the banking sector to the real economy. Thus, it is important to understand the available measures of systemic risk and their relationships with one another in order to show later an indicator of their relative usefulness. This article aims to shed some light on these pressing issues by means of an empirical analysis of the most widely used and best known measures of systemic risk employed by investors and regulators worldwide.

In this paper, we concentrate on what are widely acknowledged to be the most important systemic actors: the biggest banks in the two biggest economic areas (Western Europe and the United States).² Using daily data, we construct and compare a comprehensive set of measures for these major actors. Specifically, we compute six different sets of systemic risk measures for a sample of the 20 biggest European and 13 biggest US banks from January 2004 to November 2009.³ The six measures are based on (i) principal components of the bank's credit default swaps (CDS); (ii) interbank interest rates; (iii) structural credit risk models (Merton (1973)); (iv) collateralised debt obligation (CDO) indices and their tranches; (v) multivariate densities computed from CDS spreads; and (vi) co-risk measures. Unlike previous studies, we do not focus on just one market (stock market) to compute our measures, but use data from several markets: stock, credit and derivatives markets. Then, we run a "horse race" to evaluate the usefulness of these measures on the basis of three different criteria: (i) Granger Causality tests; (ii) Gonzalo and Granger metric; and (iii) McFadden R-squared. Finally, we run some additional checks to verify the robustness of the "horse race" results.

Our results imply that the best-performing measures of systemic risk are based on simple indicators obtained from credit derivatives and interbank rates. Measures based on complex models or convoluted statistical procedures do not perform particularly well in our sample. It seems that "model risk" is an issue when developing appropriate measures of systemic risk. Therefore, the implication for investors and regulators looking for reliable systemic risk indicators is to stick to simple, robust indicators based on credit derivatives and on interest rates data from the market.

¹ Department of Business Administration, Universidad Carlos III de Madrid.

² Billio et al (2010) find that banks may be more central to systemic risk than non-bank financial institutions that engage in banking functions.

³ In a recent study, the International Monetary Fund (2009) posits that smaller institutions may also contribute to systemic risk if they are closely interconnected. However, systemic risk should be most readily observed in the largest banks.

Systemic risk measures

In this section, we briefly summarise the systemic risk measures proposed in the literature, which are based on market information.

- **Principal Component Analysis (PCA):** CDS are credit derivatives that provide insurance against the risk of default of a certain company and, hence, their spreads measure the risk that is faced by bondholders of the reference entity. We perform a PCA on a pool of CDS spreads and consider the first principal component as a systemic risk measure because it is the common factor that is driving the CDS spreads, and the nature of such spreads should be reflecting the common risk of default.
- **LIBOR spreads:** this group of measures involves the use of LIBOR⁴ as the reference interest rate relative to either the Overnight Interest Swap (OIS) or treasury bills (TBILLS). These two proxies of systemic risk are similar, but important conceptual differences exist between them. The LIBOR-OIS spread reflects liquidity and default risk over the following months while the LIBOR-TBILL spread captures not only liquidity and default risk but also the additional fact that, during periods of turmoil, investors lend against treasury bills (the best form of collateral), measuring the “flight to quality” effect.
- **Structural model:** the third group of measures is based on the framework of Merton’s model (1973). The basic reference is Lehar (2005), although other authors use similar approaches. Lehar (2005) proposes a systemic risk measure based on the probability of default of a given proportion of banks in a given financial system. The probability of default is linked to the relationship between the banks’ asset value and their liabilities. In summary, the procedure for estimating this variable consists of recovering the bank’s asset portfolio and correlations through Merton’s model and an Exponential Weighted Moving Average (EWMA) model, respectively. Then a simulation is carried out to infer the banks’ future asset portfolio and compare it with their liabilities according to different criteria, in order to construct two systemic risk indexes: SIV and SIN, which refer to a systemic risk index based on the expected value of bank’s asset portfolio and a systemic risk index based on the expected number of defaulted banks, respectively.
- **CDO indexes and tranches:** some authors used CDOs to estimate systemic risk measures. We base this group of measures on Bhansali, Gingrich and Longstaff’s (2008) paper. They extract the idiosyncratic, sector-wide and economy-wide, or systemic risk components from US (CDX) and European (iTraxx) prices of indexed credit derivatives by means of a linearised three-jump model. In summary, they calibrate the model to both the CDO indexes and their tranches. They then break down the CDO indexes into three different spreads.
- **Multivariate Density:** Another set of measures assesses systemic risk by recovering the multivariate density of an analysed portfolio. Within this line, we follow Segoviano and Goodhart (2009), who propose a set of banking stability measures based on distress dependence, which is estimated by the Banking System Multivariate Density (BSMD). BSMD is the key element for measuring banking stability and is estimated by means of Consistent Information Multivariate Density Optimizing (CIMDO) methodology (see Segoviano (2006)). Once the BSMD is recovered, the authors propose two measures for common distress in the banking

⁴ We use the LIBOR for the main currencies under study (ie USD LIBOR and EURIBOR).

system: the joint probability of distress (JPoD), which represents the probability of all banks in the portfolio becoming distressed, and the Banking Stability Index (BSI), which represents the expected number of institutions to become distressed, conditional on the fact that at least one institution has become distressed. However, the estimation of the BSMD becomes harder as we increase the number of banks under analysis. To overcome this problem, we analyse this measure using “reduced portfolios” according to three criteria: (i) level of CDS spread; (ii) level of liabilities; and (iii) level of the liabilities over market value ratio. That is, for each period of time, we choose three banks which are at the top of each classification and estimate the corresponding BSMD.

- “Co-risk management” tools: the last set of systemic risk measures are based on the traditional risk management tools such as value-at-risk (VaR) and expected shortfall (ES). Adrian and Brunnermeier (2009) propose estimating institution i 's co-value-at-risk ($CoVaR_i$) as the whole system's (ie, portfolio's) VaR_s conditioned on institution i being in distress (ie being at its unconditional VaR_i level). On the basis of $CoVaR$, they calculate the marginal contribution of institution i to the overall systemic risk as the difference between $CoVaR$ and the unconditional whole system's VaR , which is denoted as Delta Co-Value-at-Risk ($\Delta CoVaR_i$). Therefore, $\Delta CoVaR_i$ allows us to determine how much an institution adds to overall systemic risk. Then, we add the $\Delta CoVaR_i$ of each institution using two different criteria, equally weighted, and weighted by market capitalisation. Additionally, we apply the “co-risk” methodology to the ES through the quantile regression.

Horse race

In this section, we choose the most informative variables within each group, by regressing the measures against the influential events that have marked the main episodes of the crisis. Then we rank the systemic risk measures according to their performance in the Granger causality test, the Gonzalo and Granger metric and the McFadden R-squared.

To choose the most informative variables about systemic risk, we use the Influential Events Variable (IEV). This is a dummy variable which equals 1 on a date when there is important systemic news during the financial crisis and is equal to 0 otherwise. This variable covers a wide variety of events such as: bankruptcies (financial and non-financial firms), substantial firms' losses, stock market downturns, takeovers among financial institutions, and nationalisations or bail-outs among others. Then, we run logistic regressions, using each systemic risk measure as an explanatory variable, and choose the systemic risk measures with the highest average McFadden R-squared. Table 1 summarises the selected systemic risk measures for those groups in which there is more than one proposed measure.

In order to rank the selected variables across groups within each economic area, we employ three criteria: (i) Granger causality test; (ii) Gonzalo and Granger metric; and (iii) McFadden R-squared. The first criterion gives information on the very important issue of which market provides the leading indicator. Having reliable early indicators is of crucial importance for market agents and especially for regulators.⁵ The second and third criteria relate, from

⁵ As documented in Sorkin (2010), on 5 June 2007, none other than the Fed chairman Ben Bernanke said, “At this point the troubles in the subprime sector seem unlikely to seriously spill over to the broader economy or the financial system.” To be fair, not only regulators appear strikingly short-sighted, but academics as well deserve a mention. For instance, in a paper on measures of systemic risk published in the *Journal of Financial Economics* (available online, edition 3 July 2007), Bartram, Brown and Hund (2007) said that “the distress of central bankers, regulators and politicians about the events (systemic risk) we study could be overstated and

different perspectives, to each measure with an unobservable component, that is, the underlying *systemic risk trend* in the economy. In the Gonzalo and Granger metric, we compare, by pairs of measures, the contribution of each measure to the unobserved component, while by means of the McFadden R-squared, we measure the relationship between each measure and the IEV which is a proxy for the main systemic events of the current crisis.

- (i) Granger causality test: the first classification is based on the Granger causality test (Granger (1969)). This test intuitively examines whether past changes in one variable, X_t , help to explain contemporary changes in another variable, Y_t . If not, we conclude that X_t does not Grange-cause Y_t . We carry out the Granger causality test by pairs of measures within each economic area. To rank the measures, we give a score of +1 to measure X if X Granger causes another measure, Y, and we give a score of -1 to measure X if X is caused in the Granger sense by Y. The best measure receives the highest positive score and the worst measure the highest negative score.⁶
- (ii) Gonzalo and Granger metric: the second classification is based on the Gonzalo and Granger (1995) (GG) metric. This analysis allows us to determine, by pairs of measures, the relative contribution of each measure to the unobserved factor that is the driving force in the cointegration relationship. In our framework we define that unobserved factor as the *systemic risk common trend* in the economy. To rank the estimated measures we make use of the fact that the GG metric is bounded between 0 and 1. We assign a score of +1 to measure X if X contributes more to the systemic risk common trend factor (ie its GG metric is larger than 0.5) than the measure Y, which is assigned -1, and we assign a score of -1 to measure X and a score of +1 to measure Y otherwise.
- (iii) McFadden R-squared: previously, we made use of the logistic and multinomial regressions against the Influential Event Variable (IEV) to rule out those variables which provide less information about the systemic events within each group of measures. In this case, we rank the systemic risk measures by pairs, assigning a score of +1 to the measure with the highest R-squared and -1 to the one with the lowest R-squared.

Table 2 summarises the ranking scores according to the specified criterion. For the European market, the best indicator is the LIBOR-OIS spread followed by the first principal component of the single-name CDS whereas the least reliable indicator is the $\Delta CoES_i$. For the US market, the best indicator is the first principal component of the single-name CDS followed by the LIBOR-TBILL spread, whereas the least reliable indicator is the $\Delta CoVarR_i$.

Additionally, we conduct robustness tests to show the consistency of the results. It can be argued that assigning the same weight to all the events, irrespective of their potential severity and impact on the financial system, may bias the results. To deal with this possible problem, we employ an alternative definition of the dependent variable: instead of using only two values (ie 1 or 0), we decompose the IEV variable into four different levels, in accordance with the event's severity. The multinomial Logit regression outcomes are not materially

that current policy responses to financial crises could be adequate to handle major macroeconomic events." Spot on.

⁶ This ranking procedure is related to the well-known Condorcet voting method. The Marquis de Condorcet, a prominent reformer who became a secretary of the revolutionary French National Assembly in 1791, suggested dividing elections into a series of one-on-one contests, so that every candidate is directly compared with every other. If there is a candidate who wins every such match, it is clear who should be the overall winner of the tournament. However to avoid some of the problems of the Condorcet approach we also allow for negative as well as positive scores.

different from the ones obtained using a standard Logit regression. On the other hand, it can also be argued that the Granger causality test is designed to handle pairs of variables, and could produce misleading results when the true relationship involves three or more variables. To deal with this criticism, we run a VAR equation in which we consider the six selected systemic risk measures by economic area and establish a ranking score according to the significant coefficients. Those scores are broadly in agreement with the Granger causality test in both portfolios.

Conclusion

In this paper, we compute six different sets of systemic risk measures for a sample of European and US banks from January 2004 to November 2009. The six measures are based on (i) principal components of the banks' credit default swaps; (ii) interbank interest rates; (iii) structural credit risk models; (iv) collateralised debt obligation indexes and their tranches; (v) multivariate densities computed from CDS spreads; and (vi) co-risk (CoVaR) measures. We then compare them using three different criteria: (i) Granger causality tests; (ii) the Gonzalo and Granger metric; and (iii) their relation to an index of systemic events. We find that for the European market, the best indicator is the LIBOR-OIS spread followed by the principal component of the single-name CDS, whereas the least reliable indicator is the equally weighted Delta Co-Expected Shortfall. For the US market, the best indicator is the first principal component of the single-name CDS followed by the LIBOR-TBILL spread, whereas the least reliable indicator is the equally weighted Delta-Co-Value-at-Risk.

Therefore, our results imply that the best-performing measures of systemic risk are based on simple indicators obtained from credit derivatives and interbank rates. On the other hand, indicators relying on complex statistical procedures or questionable assumptions do not perform particularly well in our sample. The implications for investors and regulators are straightforward: look for simple, robust indicators based directly on liquid market prices of credit-sensitive instruments; beware of overcomplicated modelling based on dubious assumptions; and avoid inferences from the prices of financial products traded in thin markets.

Table 1
Selected measures by group

	LIBOR spreads	Structural models	Multivariate densities	“Co-Risk” tools
European Portfolio	LIBOR-OIS	SIN	BSI Reduced Portfolio Liabilities/MV	$\Delta CoES$ Equally Weighted
US Portfolio	LIBOR-TBILL	SIV	BSI Reduced Portfolio Liabilities/MV	$\Delta CoVaR$ Equally Weighted

This table contains the selected measures by group based on the average McFadden R-squared.

Table 2
Horse Race

Panel A: European Portfolio

	Granger causality test	GG metric	McFadden R-squared	Final score
LIBOR-OIS	0	3	5	8
PCA	3	3	1	7
CDO	0	2	-3	-1
BSI	-1	-3	3	-1
SIN	-1	0	-5	-6
$\Delta CoES$	-1	-5	-1	-7

Panel B: US Portfolio

	Granger causality test	GG metric	McFadden R-squared	Final score
PCA	3	4	-3	4
LIBOR-TBILL	1	-3	5	3
CDO	0	-1	3	2
SIV	2	4	-5	1
BSI	-3	1	-1	-3
$\Delta CoVaR$	-3	-5	1	-7

This table reports the ranking scores of the systemic risk measures among three classifications: Granger causality test, Gonzalo and Granger metric and McFadden R-squared. We also report the final score, which is the sum of scores among classifications. Panel A refers to the European portfolio and Panel B to the US portfolio.

References

Adrian, T and M Brunnermeier (2009): “CoVar: a method for macroprudential regulation”, *Federal Reserve Bank of New York Staff Report*, no 348.

Bhansali, V, R Gingrich and F Longstaff (2008): “Systemic credit risk: what is the market telling us?”, *Financial Analysts Journal*, no 64, pp 16–24.

Billio M, M Getmansky, A Lo and L Pelizzon (2010): “Econometric measures of systemic risk in the finance and insurance sectors”, *NBER Working Paper*, no 16223.

Granger, C (1969): “Investigating causal relations by econometric models and cross-spectral methods”, *Econometrica*, 37, pp 424–38.

Gonzalo, J and C Granger (1995): “Estimation of common long-memory components in cointegrated systems”, *Journal of Business and Economic Statistics*, no 13, pp 27–35.

Lehar, A (2005): “Measuring systemic risk: a risk management approach”, *Journal of Banking and Finance*, no 29, pp 2577–603.

Merton, R (1973): “A rational theory of option pricing”, *Bell Journal of Economics and Science*, no 4, pp 140–83.

Sorkin, A (2010): *Too Big to Fail*, Penguin Books.