Systemic Risk and Severe Economic Downturns: A Targeted and Sparse Analysis[1]

Massimiliano Caporin [a] Michele Costola [b] Jean-Charles Garibal [c, d] Bertrand Maillet [e, f, g, 2]

Abstract

Recent studies indicate that systemic risk has predictive power over severe economic downturns. We propose a novel methodology that employs sparsity and targeting approaches to optimally select and combine systemic risk measures to forecast the tail of a given economic variable. Out-of-sample analysis shows that the optimal combination of systemic risk metrics may vary over time, forecasting horizons and economic proxies. Moreover, a few systemic risk measures contain all the important information for capturing the relation between systemic risk and real economy; therefore, a rigid combination approach may not be optimal, and the flexible parsimonious extension we introduce leads to improvement in forecasting performance.

JEL classification: C45, C53, C58, G01, G11.

Keywords: Sparse PCA, Systemic Risk, Financial Crisis.

[1] The usual disclaimer applies.

[a] University of Padua – Padova, Italy.
[b] Ca' Foscari University of Venice – Venice, Italy.
[c] Grenoble Ecole de Management – Grenoble, France.
[d] LEO (University of Orleans) – Orleans, France.
[e] AIM QUANT Research Center (emlyon business school) – Paris, France.
[f] CEMOI (Univ. La Reunion) – Saint-Denis, France.
[g] Variances – Paris, France.
[2] Corresponding author: Bertrand Maillet, emlyon business school – Paris Campus, 15 boulevard Diderot, F-75012 Paris, France. *E-mail address:* bmaillet@em-lyon.com *Phone number:* +33 6 86 43 19 14.

1. Introduction

Since the financial crisis of 2008, systemic risk has attracted extensive research interest, resulting in the definition of new metrics aimed at measuring the stability of the financial system. Systemic risk presents several dimensions, including the degree of interconnectedness among financial institutions, illiquidity, financial leverage, and the likelihood of financial losses (Hansen, 2014). Given the endogenous and multidimensional nature of systemic risk, its measurement is a complex task (Allen et al., 2012; Morley, 2016). In that respect, a recent strand of literature advocates the use of a multipronged approach combining risk metrics into an index (Holló et al., 2012; Louzis & Vouldis, 2017). As Giglio et al. note (2016, p. 458), systemic risk measures can be thought of as "imperfectly measured versions of an unobservable systemic risk factor." Therefore, a dimension-reduction technique may improve the identification of the most relevant features of systemic risk. Among the methods belonging to the class of dimension-reduction methods, Factor Analysis (FA) and Principal Component Analysis (PCA), applied to a set of systemic risk measures, have recently received attention, i.e., Giglio et al., 2016 and Nucera et al., 2016. In particular, the first contribution shows that the composite index obtained from a PCA exhibits significant association with the real economic activity and is helpful in predicting future economic downturns.

We further investigate this relationship and make use of both sparsity and targeting approaches to propose a generalised method to predict downturn risk in proxies of the economic activity. The relationship between the unobservable systemic risk factor and a proxy for economic activity may only partially reflect the different dimensions characterising systemic risk. This suggests that only a specific subset of the systemic risk measures available may be required to unveil the connection between systemic risk and the real sector. Furthermore, the association between systemic risk and economic activity could change depending on the economic proxy considered; it could also be time-varying and associated with the forecasting horizon. The exclusion of redundant information resulting from a dimension-reduction approach may provide better identification of the systemic risk features related to given economic variables; therefore, the adoption of a more flexible approach might lead to an improvement in forecasting performance.

We introduce a TArgeted Sparse SYstemic Risk Index (TASSYRI). Our approach aims at adapting a parsimonious selection of variables to the features of the economic proxy downturns to be forecasted. Our methodology is in line with the focus of policymakers who closely follow key indicators and with recent literature that has explored the systemic risk impact on macroeconomic shocks. Consequently, we focus on the left tail (the downturns) and not on the central tendency of reference economic indicators.

The TASSYRI methodology builds on two steps. First, a sparse selection applied to a group of systemic risk measures provides a set of composite risk indexes featuring different degrees of sparsity, as governed by a smoothing parameter. We note that the classical PCA approach appears as a specific case where there is no sparsity; in that circumstance, all the systemic risk measures are included in the risk index, leading to the solution proposed in Giglio et al. (2016). The second stage involves targeting and leads to identification of the sparse index that best causes extreme variations in a given economic proxy. This is achieved through constrained optimisation of the smoothing parameter that drives the regularised Singular Value Decomposition (SVD) governing sparsity. This optimisation builds on the tail Granger causality test proposed by Hong et al. (2009). The index identified in this second step is the TASSYRI.

In the first part of the paper, we illustrate the construction of the TASSYRI and show how sparsity and targeting lead to the selection of a reduced set of systemic risk measures. We then perform several validity checks on the index construction procedure. The second part presents detailed out-of-sample analysis to test the predictive ability of our methodology in a quantile regression framework. We implement our tests on several measures of economic activity. Our empirical exercise extends that of Giglio et al. (2016), as we consider a wider set of systemic risk measures and economic proxies and longer forecast horizons (up to 12 months). Our empirical studies involve well-known systemic risk metrics computed on a dataset for 95 US financial institutions, similar to the data adopted in recent articles on systemic risk (e.g., Acharya et al., 2017; Brownlees & Engle, 2017).

Our results provide novel evidence. First, we find that indexes of systemic risk exhibit predictive ability over downturns for forecasting horizons of up to 12 months for the various proxies of economic activity we consider. Second, we show that the introduction of sparsity and targeting provides an improvement in forecasting ability with respect to the simple adoption of a set of control variables, or when considering alternative indexes based on dimension reduction. Further, empirical results confirm that the optimal combination of systemic risk metrics varies over time and, consequently, that the adoption of a rigid combination approach might not prove optimal. Consistent with the findings of Giglio et al. (2016), we do not find a specific combination of systemic risk measures that outperforms other approaches for all economic proxies and forecast horizons. In addition, empirical estimations reveal that the TASSYRI is most of the time the best approach for horizons over one month. Finally, our results indicate that only five systemic risk measures carry the most information on the latent systemic risk factor: The Value-at-Risk (VaR); the Conditional Value-at-Risk (CoVaR) and the Delta Conditional Value-at-Risk (ΔCoVaR) of Adrian and Brunnermeier (2016); the Marginal Expected Shortfall (MES) of Acharya et al. (2012); and the Conditional Expected Shortfall (CES) of Banulescu and Dumitrescu (2015).

The remainder of the article is organised as follows. Section 2 presents the TASSYRI construction methodology. Section 3 provides an illustration of the index construction and discusses methodological details. Section 4 focuses on the out-of-sample analysis of the

TASSYRI features using several proxies of real economic activity, contrasting our index with available alternative proposals. Section 5 presents conclusions. Extra definitions, descriptions, complementary results, and further supporting materials are provided in an online appendix.

2. Building an Aggregated Index of Systemic Risk Measures with Sparsity

In this section, we show how to build the TASSYRI, starting from a set of systemic risk measures, considering only the relevant ones based on regularised Singular Value Decomposition and targeting the TASSYRI sparsity to the reference economic variable and forecast horizon by using the extreme causality test of Hong et al. (2009).

2.1 About Systemic Risk Measures

Our approach relies on a set of 16 systemic risk measures.[3] We group them into micro- and macro-focused. In the first group, we include individual systemic risk measures evaluated at the firm level: The Component Expected Shortfall (CES, Banulescu and Dumitrescu, 2015), the Conditional Value-at-Risk (CoVaR), the Delta Conditional Value-at-Risk (ΔCoVaR) and the Conditional Expected Shortfall (CoES, Adrian & Brunnermeier, 2016), the Marginal Expected Shortfall (MES, Acharya et al., 2013; Brownlees & Engle, 2017), the SRISK (Acharya et al., 2012; Brownlees & Engle, 2017), the Amihud Illiquidity Measure (AIM, Amihud, 2002), and the Kyle Lambda measure (Kyle, 1985). We also include in this group the Volatility (Vol) and the Value-at-Risk (VaR), which allows us to account for the evolution of the return's dispersion and tail risk. Finally, we complement this selection by adding a

[3] For a survey on available systemic risk measures, see Bisias et al. (2012), Benoit et al. (2017), and Di Cesare and Rogantini Picco (2018).

(restricted) daily CATastrophic risk in the FINancial sector (CATFIN, Allen et al., 2012).[4] The second group includes measures that focus on a single aspect of systemic risk directly evaluated at the system level. In this category, among the measures available in the literature, we select the Spillover Index (SI, Diebold & Yilmaz, 2009), the Dynamic Causality Index (DCI, Billio et al., 2012), the Turbulence Index (TI, Kritzman & Li, 2010), the Absorption Ratio (AR, Kritzman et al., 2011), and a concentration measure, namely, the Herfindahl-Hirschman Index (HHI).[5]

For the evaluation of company-specific systemic risk measures, we select a database adopted in previous studies of systemic risk (e.g., Acharya et al., 2017; Brownlees & Engle, 2017) and focus on the largest 95 US financial institutions. We recover daily price data from Thomson Reuters Datastream covering January 2000 to December 2017.[6] Following Giglio et al. (2016), we first compute the individual measures for each of the 95 financial institutions at daily frequency. Whenever necessary, we make use of a rolling-sample estimation method and evaluate the measures using a one-year window. Next, we take an equally weighted average of the individual measures. To allow comparability across the aggregated measures, we transform all of them into z-ratios (i.e., we remove the mean and standardise each measure by its sample variance).

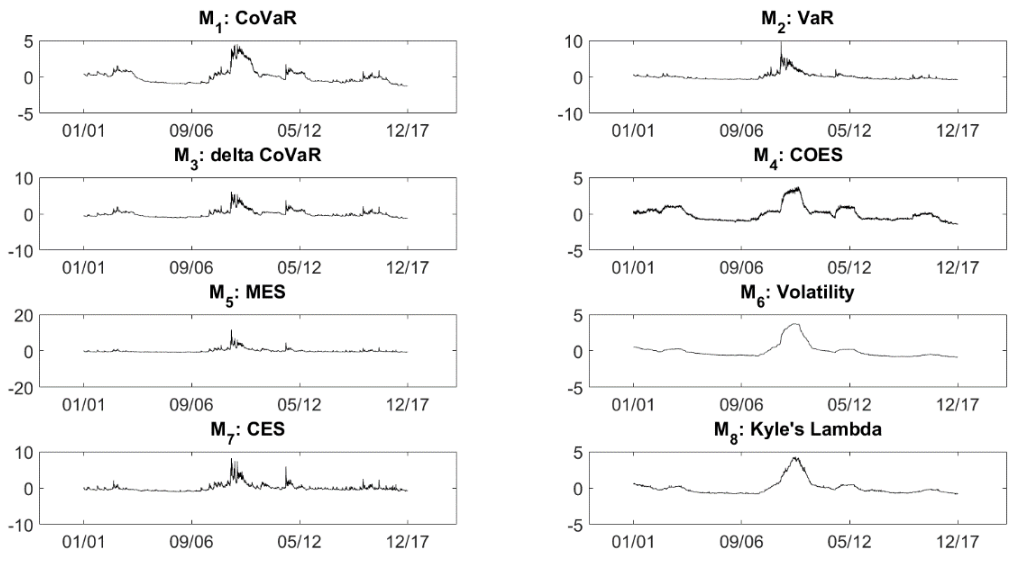
[4] Concerning the 'restricted' CATFIN, we build a constrained version of the CATFIN measure, following the same methodology as Allen et al. (2012) but based on our limited sample of 95 US financial institutions. Moreover, we work on a daily basis for coherence and time-consistency reasons with respect to the other 15 SRM. It appears that when aggregating the daily restricted CATFIN on a monthly and quarterly basis, we end up with a series that exhibits highly significant Pearson and Spearman correlations with the original monthly CATFIN. The correlations are 78% and 70%, respectively.

[5] We also performed preliminary tests using the GZ spread of Gilchrist and Zakrajsek (2012) and the Pástor and Stambaugh (2003) Illiquidity measure. The TASSYRI resulting from using the 16 systemic risk measures considered in the article and the one obtained with 18 measures (the 16 in the article and the 2 previously quoted ones) are similar (with a 20%-extreme Spearman correlation coefficient of 100%); in other words, these two additional indicators were not adding much to the analysis and were accordingly neglected.

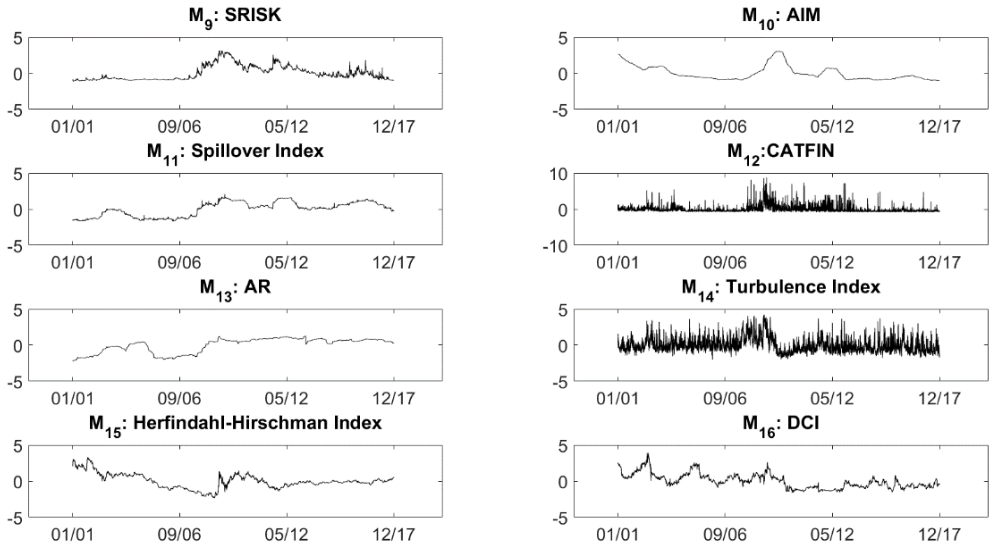
[6] Web Appendix A.1 details the computation of the systemic risk measures; Web Appendix A.2 presents the list of the 95 financial institutions in the sample; and Web Appendix A.3 shows the correlation matrix of the systemic risk measures.

Figure 1 shows the dynamics of the aggregated micro- and macro-focused systemic risk indicators; all are standardised over the full sample.[7] We note a significant increase in the level of all global systemic risk measures over the period 2007–2008, at the beginning of the financial crisis. Although common trends seem to emerge, there are some discrepancies between these measures, which confirm that systemic risk is a multidimensional phenomenon; each of the metrics captures one or several these dimensions. The analysis of correlations between the different risk measures (included in Web Appendix A.3) shows that the correlations differ across the measures, and that almost all are statistically significant at the 5% significance level. The empirical evidence confirms the presence of multicollinearity and indicates different informational content related to the measures. This further motivates the development of an indicator that integrates all dimensions of systemic risk using an information reduction technique.

Figure 1.
Dynamics of global systemic risk measures.



[7] When performing the out-of-sample evaluation, the standardisation of the measures will account only for the available information, thus excluding the presence of a look-ahead bias.



Notes: Datastream, daily data from 01/03/2001 to 12/29/2017; authors' computation. These measures estimated from rolling windows of one year. Presented here are monthly averages of z-scores of the 16 computed (aggregated) systemic risk measures, in the following order: M_1: CoVaR, M_2: VaR, M_3: ΔCoVaR, M_4: COES, M_5: MES, M_6: Volatility, M_7: CES, M_8: Kyle's Lambda, M_9: SRISK, M_{10}: AIM, M_{11}: Spillover Index, M_{12}: CATFIN, M_{13}: AR, M_{14}: Turbulence Index, M_{15}: Herfindahl-Hirschman Index, M_{16}: DCI.

2.2. Deriving a Composite Index Based on Regularised Singular Value Decomposition

The methodology adopted for the construction of the TASSYRI builds on the regularised Singular Value Decomposition (rSVD) proposed by Shen and Huang (2008). Application of the rSVD leads to a 'sparse' index and may allow identification of the most relevant features of systemic risk measures. Generally speaking, rSVD belongs to a class of Sparse Principal Component Analysis (SPCA) tools that generalise the classical PCA. While the PCA generates linear combinations of all input variables, SPCA methods select subsets of the input variables. The set of SPCA methods also include, besides that of Shen and Huang (2008), the SPCA of Zou et al. (2006). We stress that all those methods include as a special case the classical PCA; in fact, if all input measures in the SPCA or rSVD are identified as relevant, we end up with the classical PCA.[8]

[8] For the sake of completeness, we recall the principles of PCA and the full derivation of SPCA in Web Appendix B.

Shen and Huang (2008) propose to recover a principal component subject to sparsity building on a SVD, which has a methodological advantage compared to the approach of Zou et al. (2006). The latter requires a preliminary estimate of the dominant factor by classical PCA, and, in a second step, the first principal component is sparsified. On the contrary, rSVD does not require estimation of the target principal component. Moreover, as shown in Shen and Huang (2008), rSVD is not computationally intensive and leads to better identification of the null loadings.

In a classical SPCA of Zou et al. (2006), the sparse dominant factor is identified as an approximation of the first Principal Component, which is estimated in a preliminary step. The penalised regressions leading to the SPCA can be written as the following minimisation problem (Zou et al., 2006):

$$\min_{\widehat{\boldsymbol{\beta}}^s \in \mathbb{R}^p} \left[\left\| \boldsymbol{F} - \boldsymbol{M}\widehat{\boldsymbol{\beta}}^s \right\|_2 + \lambda P(\widehat{\boldsymbol{\beta}}^s) \right], \tag{1}$$

where the $(T \times p)$ matrix $\boldsymbol{M}$ contains the data—in our case the collection of standardised systemic risk indicators, where $\boldsymbol{F}$ is the $(T \times 1)$ first Principal Component (PC) of a classical PCA on $\boldsymbol{M}$ and represents here the target of the minimisation problem, $\widehat{\boldsymbol{\beta}}^s$ is the $(p \times 1)$ parameter vector we want to identify, $\|\cdot\|_a$ is the ℓ_a-norm, λ is a pre-specified tuning parameter, and $P(\widehat{\boldsymbol{\beta}}^s)$ represents a penalty function, depending on the coefficients in $\widehat{\boldsymbol{\beta}}^s$. The sparse principal component loadings are then obtained by normalisation of the estimated weights $\widehat{\boldsymbol{\beta}}^s$. Shen and Huang (2008) propose to replace the targeted minimisation term by resorting to a singular value decomposition (SVD). The SVD of the matrix $\boldsymbol{M}$ reads:

$$\boldsymbol{M} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{V}', \tag{2}$$

where $\boldsymbol{U} = \left[\boldsymbol{u_1} \dots \boldsymbol{u_p}\right]$ is a $(T \times p)$ matrix containing the left Singular Vectors, $\boldsymbol{D}$ is the diagonal $(T \times p)$ matrix of Singular Values, with $diag(\boldsymbol{D}) = \left[d_1 \dots d_p\right]$, where $\{d_1 \geq \dots \geq$

$d_p\}$ are the Singular Values, $diag(\cdot)$ is a matrix operator that extracts the diagonal of a matrix, and, finally, $\boldsymbol{V} = [\boldsymbol{v_1} \dots \boldsymbol{v_p}]$ is a $(T \times p)$ matrix which includes the right Singular Vectors of $\boldsymbol{M}$. The columns of the $(T \times p)$ matrix $\boldsymbol{Z} = \boldsymbol{UD}$ are the Principal Components and the columns of $\boldsymbol{V}$ are the corresponding Loadings. The first Principal Component $\boldsymbol{F}$ of a classical PCA is simply the first column of the $(T \times p)$ matrix $\boldsymbol{Z} = [\boldsymbol{z_1} \dots \boldsymbol{z_p}]$, such as $\boldsymbol{F} = \boldsymbol{z_1}$.

Shen and Huang (2008) thus focus on the following minimisation problem:

$$\min_{\widetilde{\boldsymbol{v}} \in \mathbb{R}^p} [\|\boldsymbol{M} - \widetilde{\boldsymbol{u}}\widetilde{\boldsymbol{v}}'\|_F^2 + \lambda\, P(\widetilde{\boldsymbol{v}})], \tag{3}$$

where $\widetilde{\boldsymbol{u}} = [\boldsymbol{u_1} \dots \boldsymbol{u_l}]$ and $\widetilde{\boldsymbol{v}} = \boldsymbol{D}_l[\boldsymbol{v_1} \dots \boldsymbol{v_l}]$ are the best rank-l approximations (with $l \leq p$) of the original $\boldsymbol{u}$ and $\boldsymbol{v}$, with $diag(\boldsymbol{D}_l) = [d_1, \dots, d_l]$. Therefore, the data matrix is approximated using the first l Singular Values (and the corresponding left and right Singular Vectors), with $\|\cdot\|_F^2$ the squared Frobenius norm, that is $\|\boldsymbol{M} - \widetilde{\boldsymbol{u}}\widetilde{\boldsymbol{v}}'\|_F^2 = tr\{(\boldsymbol{M} - \widetilde{\boldsymbol{u}}\widetilde{\boldsymbol{v}}')(\boldsymbol{M} - \widetilde{\boldsymbol{u}}\widetilde{\boldsymbol{v}}')'\} = \sum_{i=1}^{T}\sum_{j=1}^{p}(m_{i,j} - \tilde{u}_i\tilde{v}_j)^2$, where $tr(\cdot)$ is the trace operator. Finally, λ is a tuning parameter, selected *a priori*, and $P(\widetilde{\boldsymbol{v}})$ is the penalty function defined over the Loadings; in our analyses, we adopt a soft Least Absolute Shrinkage and Selection Operator (LASSO) penalty, i.e., we set $P(\widetilde{\boldsymbol{v}})$ to the ℓ_1-norm, such as: $P(\widetilde{\boldsymbol{v}}) \equiv \|\widetilde{\boldsymbol{v}}\|_1 = \boldsymbol{1}'|\widetilde{\boldsymbol{v}}| \leq \lambda$.

More precisely, for an integer $l \leq p$, the matrix $\boldsymbol{X}^{(l)} = \widetilde{\boldsymbol{u}}\widetilde{\boldsymbol{v}}'$ is such as $\boldsymbol{X}^{(l)} = \sum_{k=1}^{l} d_k \boldsymbol{u}_k \boldsymbol{v}_k'$, where $\boldsymbol{u}_k$, $\boldsymbol{v}_k$ and d_k are the k-th column of $\boldsymbol{U}$ and $\boldsymbol{V}$ and the k-th element of $diag(\boldsymbol{D})$, respectively. The optimal solution is chosen in such a way that it is the closest rank-l approximation of $\boldsymbol{M}$ that minimises the squared Frobenius distance between $\boldsymbol{M}$ and any arbitrary rank-l matrix $\boldsymbol{X}^{(l)}$, also accounting for the sparsity induced by the penalisation term. Finally, Shen and Huang (2008) show that the optimal $\widetilde{\boldsymbol{v}}^s$ for a LASSO penalty equals:

$$\widetilde{\boldsymbol{v}}^s = sign(\boldsymbol{X}'\widetilde{\boldsymbol{u}})(|\boldsymbol{X}'\widetilde{\boldsymbol{u}}| - \lambda)_+, \tag{4}$$

where $sign(\cdot)$ is the sign function, $|\cdot|$ the absolute value, and $(\cdot)_+ = max(0,\cdot)$ the max function.

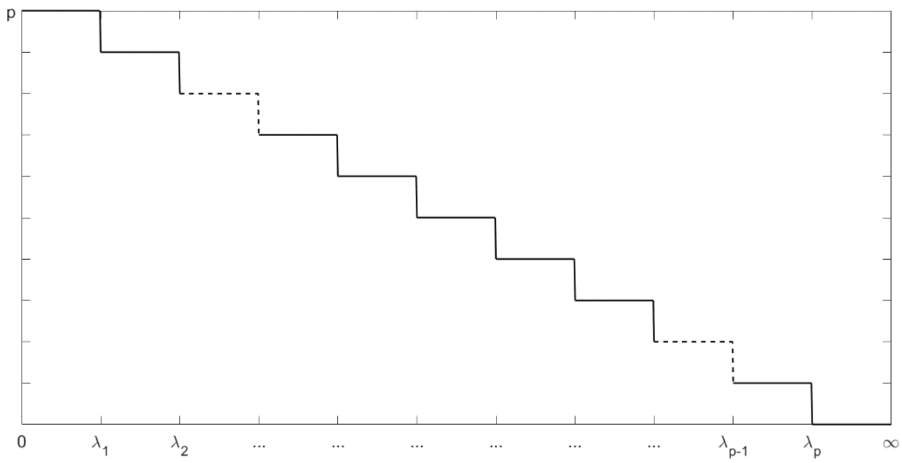
Then, we standardise the optimal parameters' vectors in (4), using the ℓ_2-norm obtaining $\widehat{\boldsymbol{x}}_i^s = \frac{\widehat{v}_i^s}{\|\widehat{v}_i^s\|_2}$, for $i = [1, 2, \dots, l]$; $\widehat{\boldsymbol{x}}_i^s$ is the i-th sparse loading vector and, by collecting the *l* $\widehat{\boldsymbol{x}}_i^s$ vectors into the $(T \times l)$ matrix $\widehat{\boldsymbol{x}}^s$, we obtain the first *l* Sparse Principal Components as $\widehat{\boldsymbol{Z}} = \boldsymbol{M}\widehat{\boldsymbol{x}}^s$.

By following the approach of Shen and Huang (2008), we are thus able to identify the first *l* sparse Principal Components; in the following, we consider $l = 1$ and $l = 2$, consistently with Giglio et al. (2016), who empirically consider the first two Principal Components, recovered from a classical PCA. When focusing on a single Sparse Principal Component, this will yield, by construction, our TASSYRI. Conversely, when $l = 2$, we will derive two indexes: one associated with the first component and the other matching the second component. We highlight that rSVD identifies sparse approximations for different Principal Components of the data, thus allowing us to analyse the impact of different orthogonal dimensions of systemic risk on the target economic variables.

We note that the solution of the minimisation problem requires knowledge of the tuning parameter λ, and thus, assuming for the time being that $l = 1$, the TASSYRI is also a function of λ. In the following, we denote the penalised Principal Component as $TASSYRI(\lambda)$. We already pointed out the sparsity induced by rSVD on the index. Beyond this principle, the use of rSVD leads to a volatility reduction that comes from the usual trade-off between bias and variance; therefore, the main factor from sparse SVD, by definition, has more stable time dynamics relative to classical PCA (since and if rSVD selects only a subset of the systemic risk measures). This property is desirable since the implementation of regulatory policies should not be based on noisy and erratic metrics of systemic risk.

A fundamental aspect of rSVD relates to the choice of the tuning parameter λ that controls the degree of sparsity. Unlike the extensive statistical literature on regularisation and penalisation, we do not identify the optimal λ by resorting, for instance, to cross-validation approaches. Instead, we identify the turning points over the λ support that lead to aggregated indexes characterised by different levels of sparsity. If we consider a rank-1 approximation, i.e., if we focus only on the first principal component, then when λ increases, the number of active elements (i.e., those different from zero) in $\widehat{\boldsymbol{v}}^{\boldsymbol{s}}$, and therefore the number of active elements (the underlying systemic risk measures) in the 'sparse' loading vector $\widehat{\boldsymbol{x}}^{s}$, approaches zero. This leads to a small number of columns of $\boldsymbol{M}$ in the estimated sparse principal component $\widehat{\boldsymbol{Z}}$. The degenerated limiting case is when $\lambda = \infty$, for which $\widehat{\boldsymbol{v}}^{\boldsymbol{s}}$ and $\widehat{\boldsymbol{x}}^{s}$ correspond to zero vectors (i.e., no systemic risk measure is selected). In the opposite case, when λ goes to zero, $\widehat{\boldsymbol{x}}^{s}$ tends to the loading vector characterising the first Principal Component of a traditional PCA, and the number of active elements then converges to its maximum value p. Therefore, there exist intermediate values of λ (between zero and plus infinity) that give a composite index $\widehat{\boldsymbol{Z}}$ based on different selections of the elements in $\boldsymbol{M}$, with a cardinality of $\widehat{\boldsymbol{x}}^{s}$ ranging from 1 to p. The values for λ can be identified by adopting a grid search, or by fixing the number of non-null elements in the vector $\widetilde{\boldsymbol{v}}^{\boldsymbol{s}}$ and searching for the minimum λ that satisfies the constraint. In our analyses, we follow the second approach. The step function in Figure 2 shows how the number of selected components (on the y-axis) changes according to λ (on the x-axis). Each turning point, from left to right, indicates when a new systemic risk measure exits from the composite index.

Figure 2.
Number of selected systemic risk measures in the rSVD according to the value of λ.



Notes: The total number of systemic risk measure is equal to p. The represented step function shows how the number of the selected components (y-axis) changes according to λ (x-axis), where each turning point indicates when a measure leaves the dominant principal component. When λ tends to infinity, the number of selected components approaches to zero while, on the opposite, when λ tends to infinity, all the components are selected as in the classical PCA.

Given p systemic risk measures, there will be a corresponding number of turning points for λ, each of which is associated with a specific selection of systemic measures and, consequently, a specific $TASSYRI(\lambda_i)$ for $i = [1, \ldots, p]$. These indexes might include from one systemic risk measure to the full set of the available risk measures. The latter case, with the $TASSYRI(\lambda_p)$, corresponds exactly to the seminal proposal by Giglio et al. (2016), i.e., the classical PCA. When $l = 2$, we will have two sets of TASSYRI, namely $TASSYRI_1(\lambda_i)$ and $TASSYRI_2(\lambda_i)$. The two sets will be separately identified and, consequently, the two TASSYRI might be associated with different values of the penalisation parameters and with different cardinality (i.e., TASSYRI_1 might include either a larger or a smaller number of systemic risk measures than TASSYRI_2).

2.3. Identifying the Optimal Index

The rSVD approach leads to p different TASSYRI associated with an increasing number of underlying systemic risk measures. Our objective at this stage is to identify the optimal combination of (a restricted small number of) systemic risk measures, according to a statistical or an economic selection rule. Since we aim to identify an index connected to economic fragility, we focus on causal links between systemic risk and real economic activity. In our analyses, following Giglio et al. (2016), we start by taking the Industrial Production Index (IPI) as the reference proxy for real economic activity.

Several empirical analyses have shown that the propagation of financial shocks to the real economy is non-linear. This implies that reaching extreme values in the aggregated index could help to explain systemic events, inducing slowdowns in economic activity. Consequently, to measure the association between a composite index of systemic risk measures and a real activity proxy, we select the Granger causality test of extremes proposed by Hong et al. (2009). Such a choice is consistent with the method adopted by Giglio et al. (2016); in fact, the authors first measured the predictive power of their aggregated index obtained from a classical PCA using a quantile regression setting. Then, they proceeded to test whether the lagged value of the index of systemic risk indicators explains extreme variations in industrial production by focusing on specifications that include or exclude the aggregated systemic risk index. Our proposal, however, differs in that we assess the extent to which positive or upper extreme values of the aggregated index of systemic risk Granger-cause the negative extreme movements in a measure of economic activity. In our setting, the Hong et al. (2009) test is used to identify the optimal TASSYRI across the p alternatives we compute, where each was obtained according to the methodology outlined in the previous subsection. In addition, the introduction of a delay factor allows us to identify the optimal index with a clear link to a predictive horizon.[9] We now summarise the test before describing the selection of the optimal TASSYRI.

Let $y_{1,t+h}$ be a proxy of the real economic activity at time $t+h$ and $Q_{1,t+h}(\alpha;\boldsymbol{\theta}_1)$ the quantile of order α of the conditional distribution of $y_{1,t+h}$, with $\boldsymbol{\theta}_1$ being a vector of parameters associated with the specification adopted for the dynamic of $y_{1,t+h}$ conditional quantiles. Note that h is the horizon factor that monitors the time distance between the economic variable and the aggregated systemic risk index. The horizon factor is introduced for flexibility purposes.

As in Giglio et al. (2016), we set α to 20%. This is based on a trade-off between the limited sample size (due to the use of monthly data for most proxies of real economic activity) and the need to have a significant number of observations in the left tail of the distribution of $y_{1,t+h}$ to perform the Hong et al. (2009) test. Let $Hit_{1,t+h}(\alpha;\boldsymbol{\theta}_1)$ be a dummy variable defined as follows:

$$Hit_{1,t+h}(\alpha;\boldsymbol{\theta}_1) = \mathbb{I}_{\{y_{1,t+h} \leq Q_{1,t+h}(\alpha;\boldsymbol{\theta}_1)\}} = \begin{cases} 1 & if\ y_{1,t+h} \leq Q_{1,t+h}(\alpha;\boldsymbol{\theta}_1) \\ 0 & otherwise. \end{cases} \quad (5)$$

This variable is equal to 1 when $y_{1,t+h}$ is extreme and negative (corresponding, for instance, to a severe contraction of economic activity). In the same manner, $y_{2,t} = -TASSYRI(\lambda_i)$, the opposite of the aggregated index of systemic risk obtained via the rSVD methodology described in the previous section. Moreover, $Q_{2,t}(\alpha;\boldsymbol{\theta}_2)$ is the quantile of order α of the conditional distribution of $y_{2,t}$, $\boldsymbol{\theta}_2$ is a vector of parameters, and $Hit_{2,t}(\alpha;\boldsymbol{\theta}_2)$ is the dummy variable defined as follows:

$$Hit_{2,t}(\alpha;\boldsymbol{\theta}_2) = \mathbb{I}_{\{y_{2,t} \leq Q_{2,t}(\alpha;\boldsymbol{\theta}_2)\}}, \quad (6)$$

[9] Coherently with Giglio et al. (2016), we consider the systemic risk index in levels since we are interested in the intensity of the risk.

where $\mathbb{I}_{\{.\}}$ is the indicator function.

Note that this variable is equal to 1 when the aggregated systemic index is extremely high (i.e., located on its right tale), indicating a systemic event. Under the null, the extreme levels of the aggregated index of systemic risk at time t have no predictive power with the negative extreme movements in the real economy proxy at time $t+h$.

The $TASSYRI(\lambda_i)$ indexes, as well as the underlying systemic risk measures, have daily frequency, while the proxy for real economic activity (here, the IPI) is available at monthly frequency. We therefore convert the $TASSYRI(\lambda_i)$ indexes to monthly frequency by averaging the daily data index values for each month.[10]

The test statistic proposed by Hong et al. (2009) depends on a weighted sum of the estimated cross-correlations between $Hit_{1,t+h}(\alpha;\widehat{\boldsymbol{\theta}}_1)$ and $Hit_{2,t}(\alpha;\widehat{\boldsymbol{\theta}}_2)$, where $\widehat{\boldsymbol{\theta}}_1$ and $\widehat{\boldsymbol{\theta}}_2$ are consistent estimators of the true parameter vectors. This weighted sum (making explicit the dependence on the $TASSYRI(\lambda_i)$) is precisely defined as follows:

$$Z[TASSYRI(\lambda_i)] = T\sum_{j=1}^{T-1}\kappa^2(j/d)\hat{\rho}^2(j), \qquad (7)$$

where the function $\kappa(\cdot)$ is the Daniell kernel[11], d is a positive integer, and $\hat{\rho}(j)$ is the cross-correlation of order j between $Hit_{1,t+h}(\alpha;\widehat{\boldsymbol{\theta}}_1)$ and $Hit_{2,t}(\alpha;\widehat{\boldsymbol{\theta}}_2)$.

We note that all possible lags are considered in the Z-statistic in Equation (7), but terms corresponding to the longest ones are multiplied by a weight that is decreasing with the lag. This is relevant in the current context, as the inclusion of a significant number of lags helps to

[10] See Section 3.2 for different frequency conversion criteria.
[11] We use the kernel function from Daniell (1946) that induces optimal properties for the causality test. See Hong et al. (2009) for further details.

capture the stronger (or weaker) inertia in the reaction of the economy to a systemic event.[12] As shown in Hong et al. (2009), under the null hypothesis of no causality in extreme movements, we have:

$$U[TASSYRI(\lambda_i), d] = \frac{Z[TASSYRI(\lambda_i)] - C_T(d)}{[D_T(d)]^{1/2}}, \quad (8)$$

which follows a standard normal distribution, with zero mean and unit variance, where:

$$C_T(d) = \sum_{j=1}^{T-1} (1 - j/T)\kappa^2(j/d),$$

and:

$$D_T(d) = 2\sum_{j=1}^{T-1} (1 - j/T)(1 - (j+1)/T)\,\kappa^4(j/d).$$

When the parameter d increases, the value of the kernel weights increases for values of j above 1. Consequently, larger values of d are associated with a much larger dependence on past lags. Given the monthly frequency of the data, we set d equal to 10 as a reasonable compromise, as it does not place too much weight on shorter lags and almost ignores the longer ones.

We propose to identify the optimal TASSYRI by means of the Hong et al. (2009) test outcomes. The optimal index is defined as the one that returns the highest value of the criterion function in the Hong et al. (2009) test statistics. Given our selection of different TASSYRI, that

[12] The Monte Carlo simulations carried out by Hong et al. (2009) show that the test has good properties at a finite distance. It is important here to note that the minimum sample size considered by the authors in the simulations is $T = 500$, and the minimum quantile is 5% (approximately 25 observations in the tails of distributions). In our case, if we use the monthly data for the changes in the IPI as the considered real economy proxy, we have 204 observations. With a 20% quantile, this leaves 40 observations. Such a figure is close to the test application conditions suggested by Hong et al. (2009) and corresponds to a relatively adequate amount of data in the tails of the distributions.

is, the $TASSYRI(\lambda_i)$ for $i = [1, \ldots, p]$, we evaluate the Hong et al. (2009) test for each i for a given economic proxy. Then, we adopt a simple criterion function, which concisely reads as follows:

$$TASSYRI(\lambda^*) = Argmax_{TASSYRI(\lambda_i)} \{U[TASSYRI(\lambda_i), d]\} \text{ for } i = [1, 2, \ldots, p]. \quad (9)$$

The optimal λ^* is one of the turning points identified in the first step (its role is explained in Figure 2), and it is matched with an aggregated index: the one that provides the highest Hong et al. (2009) test statistics. If two or more $TASSYRI(\lambda_i)$ indexes provide the same Hong et al. (2009) statistic, we should prefer the more parsimonious one—that is, the one associated with a smaller number of systemic risk measures. In fact, if the two indexes are equivalent in terms of quantile causality with respect to the real economy proxy, it simply means that the additional information included in the index with higher i (i.e., based on a larger set of systemic risk indexes) does not provide any statistical gain.

We now explain the implications of setting different values of h in Equation (5). If $h = 0$, we are monitoring, for a given sample, the cross-correlation between tail exceedances of both the TASSYRI and our target economic variable, where the TASSYRI is lagged with respect to the economic target, with a minimum lag of 1. Conversely, when $h > 0$, the TASSYRI we derive is tailored to a sort of predictive horizon for the target variable drawdown. In fact, the cross-correlations will measure the association between the exceedances of the economic variable and the TASSYRI, where the latter is lagged by a minimum of $h + 1$ period. We thus identify the optimal index with the goal to maximise the association between the actual systemic risk index and the tail exceedance of the economic target variable h periods ahead. This is a distinctive feature of our approach, which is not present in the framework of Giglio et al. (2016). The additional flexibility of our method allows for identification of the optimal TASSYRI on demand. In fact, the ability to choose both a target variable and a forecast horizon permits construction of the optimal combination of systemic risk measures that present the largest association with the downward movements of the target variables in h periods. This is in line with the common perception that systemic risk is a multidimensional phenomenon, the various facets of which could differently impact various proxies of economic activity with time-varying intensity. On the contrary, the use of standard PCA, without sparsity and without a horizon-based composition, is less flexible in some sense, as it relies on a unique combination of all systemic risk measures, irrespective of the target variable and of the predictive horizon.

3. An Illustrative Example: The Index of Systemic Risk Measures for US Financial Institutions

Focusing on the financial institutions mentioned in Section 2.1, we illustrate the final construction of the index based on the systemic risk measures using rSVD, identify the optimal index, and provide validation checks of the proposed methodology. We stress that the purpose of this section is to illustrate the construction of the TASSYRI; it is not focussed on the economic meaning of the index we identify. For that reason, we first simply set $h = 0$. The use of a predictive perspective with $h > 0$ within the construction of the TASSYRI and the economic and financial implications are discussed in the following section, where in-depth, out-of-sample analyses are developed.

3.1 Empirically Building the Index of Systemic Risk Measures

Table 1 shows the normalised Loadings $\widehat{\boldsymbol{x}}^{s}$ derived from the rSVD methodology for different values of the parameter λ (i.e., the turning points). We run all estimates using the systemic risk measures introduced in Section 2 and the full sample (from January 2000 to December 2017). For ease of reading, in Table 1 we report the indexes in columns, from the sparsest to the largest corresponding to a classical PCA, and in rows the measures from the most to the less relevant.

For the lowest value of λ (i.e., $\lambda = 0$) we consider that all measures are active in the dominant component since the penalisation does not operate. In that case, the rSVD corresponds to the first component of a classical PCA, as in Giglio et al. (2016). When $\lambda = .79$, the number of active global systemic risk measures in the approximated first Principal Component is equal to $i = 1$, and the selected index contains only the CoVaR (Adrian and Brunnermeier, 2016). At the other end, when λ increases from 0 to higher values, the constraint becomes tighter, and other systemic risk measures are discarded in the approximated first Principal Component. For illustration, when $\lambda = .75$, four measures are active in the index (namely: M_1: CoVaR, M_2: VaR, M_3: ΔCoVaR, M_4: COES).

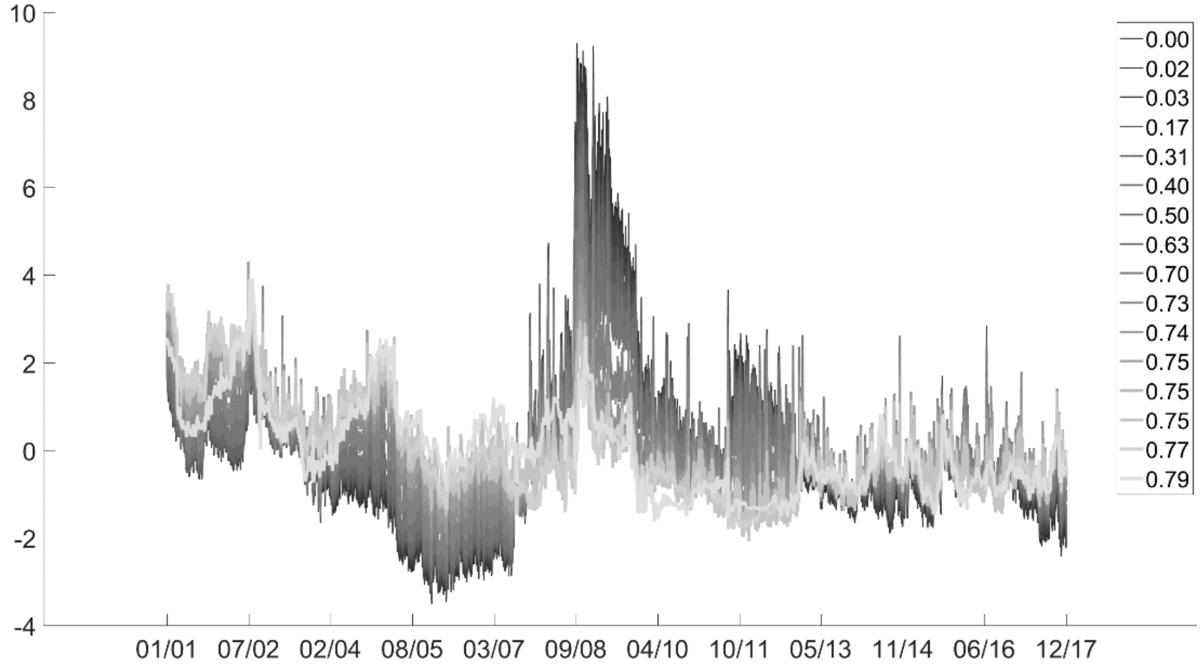
Table 1
Normalised Loadings $\hat{\boldsymbol{x}}^S$ (see Equation 5) of the various aggregated indexes $TASSYRI(\lambda_i)$.

λ_i	.79	.77	.75	.75	.75	.74	.73	.70	.63	.50	.40	.31	.17	.03	.02	.00
Id_i	Id_1	Id_2	Id_3	Id_4	Id_5	Id_6	Id_7	Id_8	Id_9	Id_{10}	Id_{11}	Id_{12}	Id_{13}	Id_{14}	Id_{15}	Id_{16}
M_1	1.00	.82	.70	.70	.69	.67	.64	.54	.43	.38	.36	.35	.34	.33	.32	.32
M_2	.00	.57	.61	.61	.61	.59	.58	.50	.42	.37	.36	.35	.33	.32	.32	.32
M_3	.00	.00	.38	.39	.39	.41	.42	.41	.38	.35	.34	.34	.33	.32	.32	.32
M_4	.00	.00	.00	.02	.02	.12	.17	.27	.31	.32	.32	.32	.31	.30	.30	.30
M_5	.00	.00	.00	.00	.00	.08	.15	.25	.31	.32	.32	.31	.31	.31	.30	.30
M_6	.00	.00	.00	.00	.00	.10	.15	.26	.31	.32	.32	.31	.31	.30	.30	.30
M_7	.00	.00	.00	.00	.00	.00	.09	.21	.29	.31	.31	.31	.31	.30	.30	.30
M_8	.00	.00	.00	.00	.00	.00	.00	.18	.27	.30	.31	.31	.30	.30	.30	.30
M_9	.00	.00	.00	.00	.00	.00	.00	.00	.21	.26	.28	.29	.29	.29	.29	.29
M_{10}	.00	.00	.00	.00	.00	.00	.00	.00	.00	.17	.21	.23	.24	.25	.25	.25
M_{11}	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.09	.14	.18	.20	.20	.21
M_{12}	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.07	.12	.16	.16	.16
M_{13}	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.08	.12	.12	.13
M_{14}	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.06	.06	.07
M_{15}	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.01	.01
M_{16}	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.01

Notes: M_1 to M_{16} represent the $i = [1,..,16]$ systemic risk measures and are in the following order: M_1: CoVaR, M_2: VaR, M_3: ΔCoVaR, M_4: COES, M_5: MES, M_6: Volatility, M_7: CES, M_8: Kyle's Lambda, M_9: SRISK, M_{10}: AIM, M_{11}: Spillover Index, M_{12}: CATFIN, M_{13}: AR, M_{14}: Turbulence Index, M_{15}: Herfindahl-Hirschman Index, M_{16}: DCI. Id_1 to Id_{16} represent the 16 TASSYRI according to the values of λ turning points reported in the first row, i.e., $Id_i = TASSYRI(\lambda_i)$.

Figure 3 shows the dynamics of all 16 aggregate indexes of systemic risk obtained through the analysis of the main sparse components. We obtained these indexes for values of the smoothing parameter λ ranging from $\lambda_{16} = 0$ to $\lambda_1 = .79$.

Figure 3.
Dynamics of the various $TASSYRI(\lambda_i)$, with λ_i being turning points reported in legend.



The dynamics of the indexes move between the two extremes: the CoVaR (Adrian and Brunnermeier, 2016) and the traditional first Principal Component. The addition of any extra systemic risk measure in the index increases its variability, even if a relevant fraction of the dynamic comes from the CoVaR. The latter is a consequence of the Loading vectors reported in Table 1, where we note that the weight of the CoVaR is always the largest across all 16 different aggregated indexes. Notably, the second systemic risk measure, in terms of Loadings size, is the VaR, and the ΔCoVaR and the COES are, respectively, third and the fourth in terms of relevance.

As illustrated in the previous section, the criterion for the selection of the optimal index is based on the rank provided by the U-statistic, which infers the existence of tail causality between aggregated indexes of systemic risk and our proxy for measuring the real economic activity.

The results of causality tests for the different competing TASSYRI (denoted below Id_1 to Id_{16} for the sake of simplicity) are summarised in Table 2. The optimal index derived from the rSVD methodology is the aggregated index Id_7, containing 7 systemic risk measures, namely, M_1: CoVaR, M_2: VaR, M_3: ΔCoVaR, M_4: COES, M_5: MES, M_6: Volatility and M_7: CES. This index provides the best rank (indicated by 1 in the last line of Table 2), followed by the other specifications (Id_8 and Id_3-Id_6). The index that includes all the systemic risk measures in the standard PCA proposed by Giglio et al. (2016) obtains a rank equal to 9 but other specifications (Id_{11}-Id_{15}) also have the same U-statistic, meaning that indexes with fewer measures have the same performance as that of the full index Id_{16}, corresponding to a classical PCA.

Table 2
Causality tests in extreme movements - Indexes of Systemic Risk Measures targeting the Industrial Production Index.

	PCA	SPCA														
λ_i	.79	.77	.75	.75	.75	.74	**.73**	.70	.63	.50	.40	.31	.17	.03	.02	.00
Id_i	Id_1	Id_2	Id_3	Id_4	Id_5	Id_6	**Id_7**	Id_8	Id_9	Id_{10}	Id_{11}	Id_{12}	Id_{13}	Id_{14}	Id_{15}	Id_{16}
$U(10)$	3.27	8.34	9.73	9.73	9.73	9.73	**10.56**	10.03	8.80	8.80	8.40	8.40	8.40	8.40	8.40	8.40
Rank	16	15	3	3	3	3	**1**	2	7	7	9	9	9	9	9	9

Notes: Datastream, daily data from 01/03/2001 to 12/29/2017; authors' computation. The Table shows the values of the $U(10)$ statistic of Hong et al. (2009) for inference on causality from monthly values of each aggregated index to the monthly change in the Industrial Production Index (IPI). In bold are highlighted the highest statistics for the best index. The threshold for significance at nominal risk level of 5% is 1.96. Id_1 to Id_{16} represent the 16 Indexes according to the values of λ turning points reported in the first row, i.e., $Id_i = TASSYRI(\lambda_i)$.

3.2 Validation of the Index of Systemic Risk Measures Construction

To validate the proposed approach, we perform a battery of tests to check whether the selection of the optimal index is driven by the informational content of the measures selected or only by the methodology itself. We consider three criteria related to 1) alternative penalisation methods; 2) alternative criteria for selecting the dimension-reduction approach; and 3) alternative frequency conversion rules for computing monthly indexes.

Alternative Penalisation Methods

Regarding the choice of the penalisation methodology, we further examine a variation of the penalty, in which LASSO, RIDGE, or Elastic-net penalties replace the rSVD. The four penalised regressions provide almost identical results and thus similarly lead to the construction of the same aggregated optimal index. Results are included in Web Appendix D.1.

Alternative Criteria for Selecting the Dimension Reduction Approach

We also consider alternative criteria to validate whether rSVD is preferred with respect to the classical PCA proposed by Giglio et al. (2016). In this respect, we consider the Kaiser criterion, which is a common choice in a PCA framework, as well as three information criteria (AIC, BIC, and SIC) and another indicator of the strength of the link between the aggregate systemic risk index and a measure of economic activity, namely, the Linear Granger Causality test (Granger, 1969), in addition to the tail Granger Causality test (Hong et al., 2009) we previously used. The findings are reported in Web Appendix D.2 and clearly show that: 1) the full PCA-related index is never selected as the optimal one; 2) the SPCA index is preferred in each considered case; 3) as predicted, the optimal index, depending on the chosen criterion, is in some cases more or less restrictive with respect to previously identified optimal Id_7: sometimes only 1 and 2 variables are selected whilst often additional measures are highlighted, but with a maximum of 10 variables.

Alternative Frequency Conversion Rules

Finally, we check for differences in alternative aggregation rules for the monthly TASSYRI computed with daily systemic risk measures. More specifically, since the target variables have monthly frequency, the issue of averaging arises in a mixed-frequency setting.

Accordingly, we may select: 1) averaging first using the monthly SRM (averaged before entering into the rSVD); 2) averaging then by aggregating the daily TASSYRI to a monthly frequency (as for the actual version of the TASSYRI); or 3) using the last value of daily SRM for computing the last TASSYRI at daily frequency to target the economic variable one month ahead. The three methods lead to similar results (with coefficients of extreme correlation between the variants of the various resulting indexes close to 100%); therefore, we can conclude that the averaging process has almost no effect on the proposed outcome.[13]

4. Systemic Risk and Economic Activity: TASSYRI and the Prediction of Downturns

We now focus on the out-of-sample predictive ability of our methodology over severe economic downturns and consider a set of economic target variables and a number of relevant control variables. From a methodological standpoint, we associate severe variation with the lower tails of the shocks hitting the target variable and thus adopt a quantile regression framework to predict a lower quantile of the target variable, using the TASSYRI lagged level as a predictor.

Consistently with Giglio et al. (2016), we consider macroeconomic shocks[14] on a set of alternative measures of economic activity, such as: 1) the Industrial Production Index (IPI); 2) the Chicago Fed National Activity Index (CFNAI); 3) the Aruoba-Diebold-Scotti Business Conditions Index (ADSI) introduced by Aruoba et al. (2009); 4) the New Privately Owned Housing Units Started (HOUST); 5) the Advance Retail Sales Excluding Food Services (RSXFS); 6) the Total Nonfarm Payrolls–All Employees (PAYEMS); 7) the Total Capacity

[13] Robustness checks on this point are available upon request to the authors. We are indebted to an anonymous referee for making the suggestion to further study, outside the IPI and CFNAI, the various proxies listed here.

[14] Macroeconomic shocks for a given economic proxy of real economic activity are obtained by removing the predictable component deriving from the own past values of the analysed variable; see Stock and Watson (2012) and Giglio et al. (2016).

Utilisation (TCU); 8) Total Vehicle Sales (TOTALSA); 9) Manufacturers' Value of Shipments on Durable Goods (AMDMVS); 10) Manufacturers' New Orders on Durable Goods (DGORDER); 11) Manufacturers–Inventories to Sales Ratio (IRSA); 12) Retail Sales on Furniture and Home Furnishings Stores (RSFHFS); and 13) the unemployment rate (UNRATE). We link the movements in the TASSYRI with the changes in the previously listed variables.

We also include a set of control variables: 1) the Economic Policy Uncertainty Index (EPU) and the Equity Market-related Economic Uncertainty Index (EMEUI) by Baker et al. (2016); 2) two Term Spreads[15]; 3) the TED spread; 4) the Default Spread (BAA10Y); and 5) the Consumer Sentiment of the University of Michigan (CSUM). Consistently with the target variables reported above, we include levels of the control variables in the analysis.

We compare the information contained in two TASSYRI (associated with the first and second regularised Principal Components, $TASSYRI_1$ and $TASSYRI_2$, respectively)[16] to three alternatives proposed by Giglio et al. (2016), namely, the first and second Principal Components coming from the Principal Components Quantile Regressions ($PCQR_1$ and $PCQR_2$, respectively) and the Partial Quantile Regression (PQR), which is an adaptation to the quantile regression framework of the Three-Pass Regression Filter (3PRF) proposed by Kelly and Pruitt (2015). We suppress the dependence on the tuning parameter in both TASSYRI and analyse the sparsity of the index in a specific subsection. Moreover, consistently with our discussion at the end of Section 2, we now focus on optimal TASSYRI identified on an adaptive basis for each target variable (and not only considering the IPI, as previously) and for predictive horizons ranging from 1 to 12 steps ahead (i.e., in Equation 5 we set h to a value from 1 to 12, coherently

[15] We consider the 10-Year Treasury Constant Maturity minus 3-Month Treasury Constant Maturity and the 10-Year Treasury Constant Maturity minus 2-Year Treasury Constant Maturity.

[16] This choice is equivalent to Giglio et al. (2016) that considers two Principal Component Quantile Regression indexes, i.e., $PCQR_1$ and $PCQR_2$.

with the predictive horizon defined below). We stress that for each target variable and predictive horizon, we have two TASSYRI, as we sparsify the first and second Principal Components by resorting to rSVD.

Accordingly, the out-of-sample analysis is conceived in a quantile regression setting with alternative horizons h, from 1 to 12 months:[17]

$$\mathrm{Q}^{i}_{1,t+h}(\alpha;\boldsymbol{\theta}_1) = \omega + \vartheta_j X^{j}_{t} + \zeta_w w_t, \tag{10}$$

where $\mathrm{Q}^{i}_{1,t+h}(\alpha;\boldsymbol{\theta}_1)$ is the conditional α-quantile of the monthly change in the i-th measure of economic activity, denoted y^{i}_{t+h} at horizon h, X^{j}_{t} is the level of the j-th predictor at time t (i.e., one among TASSYRI_1, TASSYRI_2, PCQR_1, PCQR_2, or PQR), and ϑ_j is the associated regression coefficient. Moreover, in model (10), we allow for the impact of control variables. However, as we are performing a quantile regression at a monthly frequency with limited sample size, we summarise the informative content of control variables by making use of the standard PCA, and we introduce in the model the first principal component extracted from a set of control variables, w_t; ζ_w the associated coefficient. As previously discussed, we set α equal to 20% following Giglio et al. (2016). Note that the indexes TASSYRI_1 and TASSYRI_2 are tailored to the target variable y^{i}_{t+h}; thus, for different i and different h, distinct TASSYRI will be used. Conversely, PCQR_1, PCQR_2, and PQR are independent of the target variable and predictive horizon and are invariant over i and h.

At a given horizon and for each measure of economic activity, we obtain a series of the predicted out-of-sample quantiles according to the selected predictors (e.g., the TASSYRI and control variables and the PCQR/PQR indexes and control variables). For each case, we compute the out-of-sample quantile pseudo-R^2 as in Koenker and Machado (1999). We denote it as R^2_{ij},

[17] Giglio et al. (2016) focus in their studies on the case where *h=1*.

where i identifies the target variable and j the predictor. We evaluate the pseudo-R^2 using the following equation:

$$R_{ij}^2 = 1 - \frac{\frac{1}{T}\sum_t\left[\rho_\alpha\left(y_{t+h}^i - \widehat{\omega} - \widehat{\vartheta}_j X_t^j - \hat{\zeta}_w w_t\right)\right]}{\frac{1}{T}\sum_t\left[\rho_\alpha\left(y_{t+h}^i - \widehat{\omega} - \hat{\zeta}_w w_t\right)\right]}, \quad (11)$$

where $\rho_\alpha(u) = u \times \left(\alpha - \mathbb{I}_{\{u<0\}}\right)$, with $\mathbb{I}_{\{.\}}$ the indicator function, is the tick loss function of Koenker and Bassett (1978). The pseudo-R^2 builds on a ratio between two losses: at the numerator, the loss of the proposed model taking into account the systemic risk index under evaluation, and at the denominator, the loss of the baseline model, which is the reference benchmark to be improved.

Note that parameters with hats in Equation (11) identify estimated quantities, and we stress that the pseudo-R^2 is evaluated out-of-sample. Also note that unlike the usual approach adopted for evaluating the quantile regression pseudo-R^2 as in Giglio et al. (2016), we do not consider the unconditional quantile as the benchmark model but rather the model with only the control variables plus a constant (the baseline case).[18] In that respect, we compute the predicted out-of-sample quantiles of the corresponding baseline model for each measure of economic activity considered at a given horizon. Therefore, the loss at the numerator of Equation (11) refers to the use of conditioning information coming from the predictor and controls relative to the loss at the denominator coming from controls only. Consistently with the quantile regression literature, the out-of-sample R_{ij}^2 is negative when the information coming from the denominator provides a better forecast with respect to the numerator (i.e., the numerator is larger than the denominator). Following Diebold and Mariano (1995), the statistical significance for the out-of-sample estimates is then obtained by evaluating the quantile forecast losses, $\rho_\alpha\big(y_{t+h}^i -$

[18] We are indebted to one of the anonymous referees for this suggestion.

$\hat{\omega} - \hat{\vartheta}_j X_t^j - \hat{\zeta}_w w_t)$ and $\rho_\alpha(y_{t+h}^i - \hat{\omega} - \hat{\zeta}_w w_t)$, as in Giglio et al. (2016). We implement the modified test proposed by Harvey et al. (1997), which overcomes the issues that might emerge in small samples, especially with somewhat long forecast horizons (as in our case, in which the largest horizon is one year). The testing approach we adopt thus allows us to identify in which cases the various systemic risk indexes (namely: $TASSYR_1$, $TASSYR_2$, $PCQR_1$, $PCQR_2$, and PQR) are an improvement over the baseline case, which contains only the control variables.

We run our forecasting exercise following the usual back-testing process to deal with the look-ahead bias. We choose an expanding window approach, where the predictors ($TASSYR_1$, $TASSYR_2$, $PCQR_1$, $PCQR_2$, and PQR) are built recursively in each time frame, using all the information available at a given point in time.[19] This implies that the degree of sparsity is time-varying, and that the selection of the systemic risk measure also vary over time. Accordingly, the TASSYRI are dynamically built in each window following the procedure presented in Section 2.3 and are based on the economic target variable to be forecast, i.e., the TASSYRI are adapted to the proxy of economic activity we consider and the prediction horizon h. The out-of-sample period begins in January 2006, after having selected the initial window with a length of 60 observations. For the sake of comparison and clarity, we summarise here the results by reporting the percentage of significant R^2 over the total for: 1) each considered proxy of real economic activity and thus aggregating on the forecasting horizons; and 2) on a given forecasting horizon and thus aggregating on the proxies of economic activity. Details of the results are reported in Web Appendix E.

Table 3 reports the percentage of significant R^2 for all the economic target variables on all the forecast horizons with a 10% significance level. We see first that overall (first column), the $TASSYRI_1$ and the $TASSYRI_2$ provide the highest percentage of significant R^2 (28%),

[19] The same applies to the predictors that are based on the set of systemic risk measures built using a rolling approach (Billio et al., 2012; Giglio et al., 2016).

followed by the $PCQR_2$ (21%), $PCQR_2$ (16%), and PQR (4%). Regarding the Industrial Production Index (IPI), $TASSYRI_2$ and $PCQR_1$ reach the highest percentage (25%) of the significant R^2, followed by $TASSYRI_1$ (17%), $PCQR_2$, and PQR (8%).[20] The $PCQR_1$ gives the highest percentage of significant R^2 (33%) for the Chicago Fed National Activity Index (CFNAI), followed by the $TASSYRI_1$ (17%).[21] The $TASSYRI_1$ gets the highest percentage of significant R^2 (58%) for the Aruoba-Diebold-Scotti Business Conditions Index (ADSI), followed by the $PCQR_1$ (42%). Again, the $TASSYRI_1$ has the best performance (42%) for the New Privately Owned Housing Units Started variable (HOUST), followed by $PCQR_2$ (33%), PQR (25%), $PCQR_1$, and $TASSYRI_2$ (17%). Regarding the Advance Retail Sales Excluding Food Services variable (RSXFS), the $TASSYRI_2$ leads with the highest percentage of significant R^2 (42%), followed by $TASSYRI_1$ (33%). Focusing on the Total Vehicle Sales figures (TOTALSA), the Manufacturers' Value of Shipments on Durable Goods index (AMDMVS) and the Retail Sales on Furniture and Home Furnishings Stores (RSFHFS), the $TASSYRI_1$ exhibits the highest percentage (25%) of the significant R^2 followed by the $TASSYRI_2$. On the other side, the $TASSYRI_2$ provides the highest R^2 for the Manufacturers' New Orders on Durable Goods variable (DGORDER) and the Unemployment Rate (UNRATE). Finally, the results for the Total Capacity Utilisation indicator (TCU) are equal for $TASSYRI_1$, $TASSYRI_2$, and $PCQR_1$ (17%).

Conversely, however, it is the $PCQR_2$ that gives the higher significant percentage of R^2 for predictions on the Total Nonfarm Payrolls–All Employees figures (PAYEMS), with a value of 50%, followed by the $TASSYRI_1$ (25%), $PCQR_1$ (17%), and PQR (0%). Finally, the $PCQR_1$ and the $PCQR_2$ show the highest percentage of significant R^2 (75%) for the Manufacturers–

[20] We have also considered the Vintage series for IPI. Results are included in Web Appendix E (Table E.18–E.20), and it happens that using the Vintage series led to similar results.
[21] The results for the CFNAI subcomponents are included in Web Appendix E (Table E.3–E.6).

Inventories to Sales Ratio (IRSA), followed by $TASSYRI_2$ and $TASSYRI_1$ (67% and 58%, respectively).

Second, as already shown in the empirical analysis of Giglio et al. (2016), there is no index or systemic risk measure that performs better than the baseline case (i.e., with only control variables) across all target variables. For instance, Giglio et al. (2016) find weak to no forecasting ability of $PCQR_1$, $PCQR_2$, and PQR for the Chicago Fed National Activity Index (CFNAI) and its subcomponents, while some systemic risk measures, such as the CATastrophic risk in the FINancial sector (CATFIN), for instance, perform well in several instances.[22]

Third, by extending the forecasting horizons up to 12 periods, we show that when averaging the results on horizons, the $TASSYR_1$ allows for superior forecasts for the Aruoba-Diebold-Scotti Business Conditions Index (ADSI), the New Privately Owned Housing Units Started variable (HOUST), the Total Vehicle Sales figures (TOTALSA), the Manufacturers' Value of Shipments on Durable Goods index (AMDMVS), and the Retail Sales on Furniture and Home Furnishings Stores index (RSFHFS). The $TASSYR_2$ instead provides better forecasts for the Advance Retail Sales Excluding Food Services variable (RSXFS), the Manufacturers' New Orders on Durable Goods variable (DGORDER), and the Unemployment Rate (UNRATE). Finally, summarising the results, it happens that $TASSYR_1$ is ranked first in terms of percentage of significant R^2 (among the 5 alternatives) for 6 of the 13 target variables considered, followed by $TASSYR_2$ (4), $PCQR_1$ (3) and $PCQR_2$ (2), and, finally, PQR (0).

Table 4 presents the overall 20th percentile shock forecasts in each horizon for a given Index of Systemic Risk Measures. Interestingly, the PQR provides the best forecast at horizon 1, consistently with the findings of Giglio et al. (2016). Moving to the longer horizons, the $TASSYR_1$ and $TASSYR_2$ show the best performance in every case except for horizons 10 and

[22] See Table 3 on page 464 in Giglio et al. (2016).

11, where the $PCQR_1$ provides a better performance. It is worth noting that in those cases, the improvement of $TASSYR_1$ versus $PCQR_1$ is on average larger than 30%, while is larger than 39% for $TASSYR_2$ against $PCQR_2$. Therefore, the combination of targeting and sparsity represents a valuable tool that can improve the forecasting ability of the systemic risk index with respect to the full case represented by methods combining all of the available risk measures.

Table 3.
The table reports the fraction of cases over the forecast horizons (1 to 12 months) where the aggregated index over the rows provides a statistically significant R^2 with a 10% significance level compared to the baseline quantile model including only the controls.

	Overall	*IPI*	*CFNAI*	*ADS*	*HOUST*	*RSXFS*	*PAYEMS*	*TCU*	*TOTALSA*	*AMDMVS*	*DGORDER*	*IRSA*	*RSFHFS*	*UNRATE*
TASSYRI$_1$	**28%**	17%	17%	**58%**	**25%**	33%	17%	17%	**25%**	**42%**	8%	58%	**42%**	8%
TASSYRI$_2$	**28%**	**25%**	8%	25%	0%	**42%**	25%	17%	17%	33%	**17%**	67%	33%	**50%**
PCQR$_1$	21%	**25%**	**33%**	42%	17%	0%	17%	17%	0%	25%	0%	**75%**	8%	17%
PCQR$_2$	16%	8%	0%	17%	0%	8%	**50%**	0%	0%	0%	0%	**75%**	8%	42%
PQR	4%	8%	0%	0%	0%	0%	0%	8%	0%	0%	0%	17%	0%	17%

Notes: We focus on the 20th percentile shock forecasts for changes in the target variables reported over columns. The *Overall* column aggregates the results across the target variables. Over rows we have: the Index of Systemic Risk Measures replicating the first Principal Component (TASSYRI$_1$) and the two first Principal Components (TASSYRI$_2$); the first and second Principal Components from the Principal Components Quantile Regressions (denoted PCQR$_1$ and PCQR$_2$, respectively); and the Partial Quantile Regression (PQR) proposed in Giglio et al. (2016). The target variables are: i) the Industrial Production Index (IPI); ii) the Chicago Fed National Activity Index (CFNAI); iii) the Aruoba-Diebold-Scotti Business Conditions Index (ADSI); iv) the New Privately Owned Housing Units Started (HOUST); v) the Advance Retail Sales Excluding Food Services (RSXFS); vi) the Total Nonfarm Payrolls–All Employees (PAYEMS); vii) the Total Capacity Utilisation (TCU); viii) Total Vehicle Sales (TOTALSA); ix) Manufacturers' Value of Shipments on Durable Goods (AMDMVS); x) Manufacturers' New Orders on Durable Goods (DGORDER); xi) Manufacturers–Inventories to Sales Ratio (IRSA); xii) Retail Sales on Furniture and Home Furnishings Stores (RSFHFS); and xiii) the unemployment rate (UNRATE).

Table 4.
The table reports the fraction of cases over the target variables where the aggregated index over the rows provides a statistically significant R^2 with a 10% significance level compared to the baseline quantile model including only the controls.

	h=1	*h=2*	*h=3*	*h=4*	*h=5*	*h=6*	*h=7*	*h=8*	*h=9*	*h=10*	*h=11*	*h=12*
TASSYRI$_1$	8%	**8%**	17%	**17%**	8%	**25%**	17%	58%	58%	42%	50%	**58%**
TASSYRI$_2$	8%	**8%**	**25%**	8%	8%	0%	**42%**	**67%**	**75%**	33%	42%	42%
PCQR$_1$	8%	0%	8%	8%	0%	17%	8%	25%	33%	**67%**	**58%**	42%
PCQR$_2$	8%	0%	8%	8%	8%	17%	17%	25%	25%	33%	42%	17%
PQR	**25%**	0%	0%	0%	8%	0%	0%	8%	0%	0%	0%	8%

Notes: We focus on the 20th percentile shock forecasts for changes in the target variables. The table corresponds to the split over forecast horizons of the results in the *Overall* column of Table 3. Over rows we have: the Index of Systemic Risk Measures replicating the first Principal Component (TASSYRI$_1$) and the two first Principal Components (TASSYRI$_2$); the first and second Principal Components from the Principal Components Quantile Regressions (PCQR$_1$ and PCQR$_2$); and the Partial Quantile Regression (PQR).

As a robustness check, we have also changed the initial bandwidth for the starting date in the forecasting exercise by selecting January 2004, January 2005, and January 2007. The results are included in Web Appendix E and are qualitatively similar to those reported here. We also run a set of additional robustness analyses[23] by replacing the pseudo-R^2 with other loss functions (Lopez, 1999; Caporin, 2008). The results confirm the previous evidence and are in line with the findings of Giglio et al. (2016), showing that an optimal aggregated index with the best predictive performance over all forecast horizons and all target variables does not exist.

Comparing systemic risk indexes and the control variables case

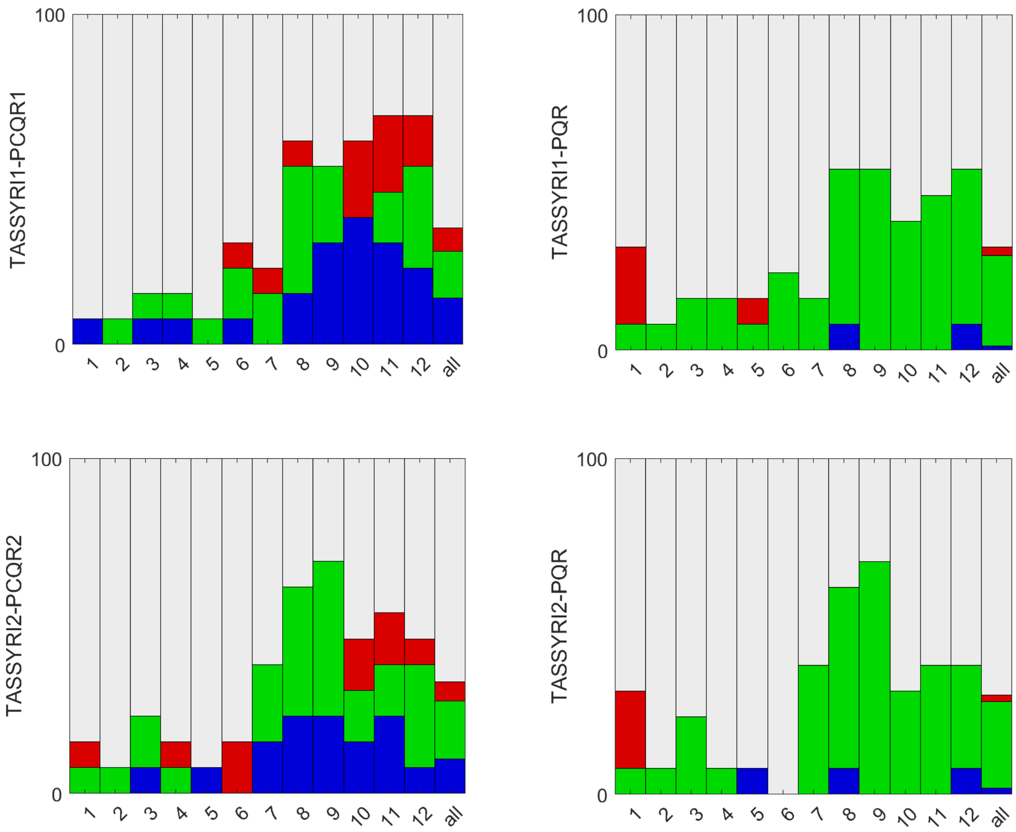
The analyses reported above, despite being informative on the predictive advantage of a tailored TASSYRI (designed for each target economic variable), do not represent a proper horse-race against alternative aggregated indexes, namely PCQR and PQR. This is a consequence of the choice of introducing in the analyses a set of control variables that allows for definition of a reference specification not dependent on aggregated systemic risk indexes. We believe that before moving to a direct comparison between aggregated indexes, we should consider if a direct comparison is appropriate. In fact, a horse-race evaluation between two aggregated indexes is informative if the indexes forecasting ability is, for different target variables and forecast horizons, superior to the prediction obtained by using only control variables. Thus, we report in Figure 4 a comparison between selected pairs of aggregated indexes, namely $TASSYRI_1$ vs. $PCQR_1$ (i.e., SPCA vs. classical PCA), $TASSYRI_1$ vs. PQR, $TASSYRI_2$ vs. $PCQR_2$, and $TASSYRI_2$ vs. PQR.[24] For each pair, still focusing on the statistical significance of the R^2 (with respect to the baseline case with control variables only), we track the following:

[23] All results are available upon request.

[24] We exclude the comparison between aggregated indexes associated with different principal factors, i.e., $TASSYRI_1$ vs. $TASSYRI_2$, $PCQR_1$ vs. $PCQR_2$, $TASSYRI_1$ vs. $PCQR_2$, and $ISMR_2$ vs. $PCQR_1$. For the same reason, we also do not report the comparisons between PQR, $PCQR_1$, and $PCQR_2$.

1) the number of cases in which both indexes do not improve the forecasting performance over the baseline model with only controls (i.e., the R^2 related to the indexes are both not statistically significant); 2) the cases where both indexes do improve with respect to the baseline case (i.e., both R^2 are statistically significant); 3) the number of cases in which a TASSYRI improves over the baseline and the PCQR/PQR does not improve; and 4) the opposite of case 3, i.e., the number of cases in which the PCQR/PQR related index improves over the baseline and TASSYRI does not improve.

Figure 4.
Pairwise R^2 comparison among the TASSYRI indexes and the PCQR or PQR indexes at each horizon and over all horizons, aggregative outcomes with respect to all the considered proxies of real economic activity.



Notes: The comparison includes four cases: 1) both indexes do not provide a forecast gain with respect to the benchmark (grey); 2) both indexes do provide a forecast gain with respect to the benchmark (blue); 3) the TASSYRI provides a forecast gain with respect to the benchmark, while neither the PCQR nor PQR do (green); and iv) the PCQR or PQR provides a forecast gain with respect to the benchmark, while the TASSYRI does not (red). The forecast gain corresponds to a R^2 value at a 10% significance level when contrasting the aggregate index to a baseline specification including only control variables.

The number of cases where the models do not provide any forecast improvement with respect to the adoption of a model based only on control variables (case 1) above) is generally large. In this scenario, the composite indexes could have an R^2 equivalent or statistically inferior to that of the baseline case. As shown in Appendix F (Figure F.1), the former realises: the two compared indexes are providing, in almost all cases, predictive performances equivalent (and not inferior) to those of the specification with only the controls. In addition, we observe different patterns over the prediction horizons. In fact, for shorter horizons, we note only minor improvements to the baseline case. Notably, and consistent with the evidence in Giglio et al. (2016), the PQR provides some better results (see right panels of Figure 4 for $h=1$). The behaviour changes when focusing on the longer horizons, where the percentage of cases in which the aggregated indexes outperform the baseline specification clearly increases. We note that $TASSYRI_1$ and $TASSYRI_2$ do dominate occasionally over $PCQR_1$ and $PCQR_2$, respectively. However, $PCRQ_1$ and $PCQR_2$ also outperform, in some cases, the corresponding TASSYRI; this happens, in particular, when focusing only on the first Principal Component ($TASSYRI_1$ vs $PCQR_1$) at horizons 9, 10, and 11. Conversely, $TASSYRI_1$ and $TASSYRI_2$ do provide sensibly better performances than PQR (right panels of Figure 4 for $h>6$), but the fraction of cases where the aggregated indexes dominate the model with only controls peaks around 50%.

On the one hand, the different behaviour of TASSYRI, PCQR, and PQR depends on the features of our approach, features that are adaptive and related to specific target variables and specific various predictive horizons. This flexibility, overall, leads to better performance relative to fixed indexes. On the other hand, the fact that the indexes outperform the baseline model only in a fraction of the total number of cases limits the insight one might obtain from a proper horse-race comparison between models. In fact, an informative horse-race should focus only on the cases where both aggregated indexes over-perform the baseline specification.

However, as our evidence shows, those cases represent only a fraction of all the combinations of horizons and target variables we study.[25]

The role of sparsity

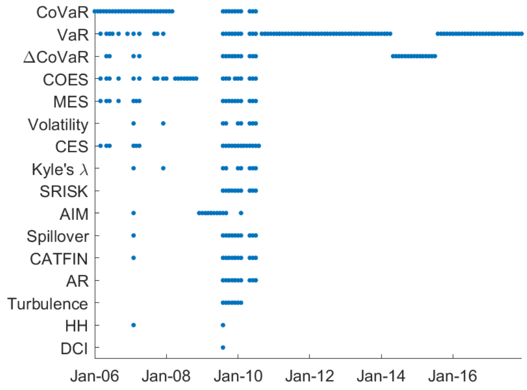
To highlight the positive impact of sparsity combined with targeting as a more flexible approach than the classical PCA, we provide graphical evaluation of the systemic risk measures selected for the out-of-sample construction of TASSYRI_1. We consider three proxies of real economic activity and three horizons (i.e., values of *h*), leading to a different level of sparsification in the associated TASSYRI_1 indexes, reminding that the various TASSYRI are target-variable specific, i.e., their composition depends on the variable we adopt to track the economic activity. In our example, we consider the Retail Sales on Furniture and Home Furnishings Stores index (RSFHFS) at horizon $h=12$, the New Privately Owned Housing Units Started variable (HOUST) at horizon $h=4$, and the Total Nonfarm Payrolls–All Employees figures (PAYEMS) at horizon $h=9$. Figure 5 reports the TASSYRI_1 composition for the period considered. It appears in these situations that RSFHFS forecast (respectively, the HOUST, the PAYEMS) goes with a high-level of sparsity (respectively, mid-level of sparsity, low-level of sparsity) since few systemic measures are selected (respectively, a few, many). The different level of sparsification can also be viewed in Table 5, where we report the average, the standard deviation, and the average turnover[26] of the selected systemic risk measures. In fact, we observe that the sparsity levels (row 1) are negatively related to the statistics (rows 2–4).

---

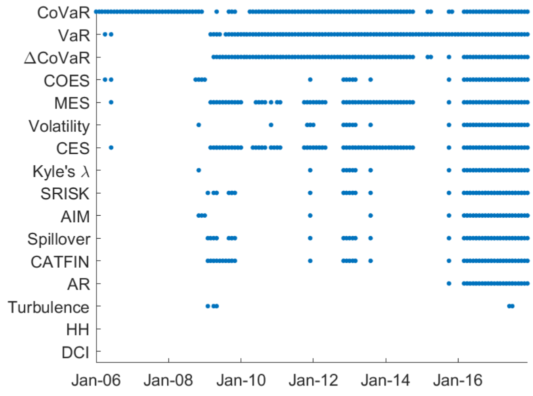
[25] For the sake of completeness, we report in Appendix F, Figure F.2 the direct comparison between TASSYRI and PCQR/PQR, irrespective of the fact the models under/over-perform the baseline case with only controls. We stress that the use of the R^2 is in this case rather inappropriate as this index is designed for the comparison of nested models, but the statistical test adopted by Giglio et al. (2016) could still be used. The figure shows that the TASSYRI has performances somewhat inferior to PCQR and somewhat superior to PQR. Nevertheless, the number of cases where the indexes are statistically equivalent dominates.

[26] The turnover is defined as the percentage of change in the selection of the systemic risk measures for the TASSYRI_1 with respect to the previous period.

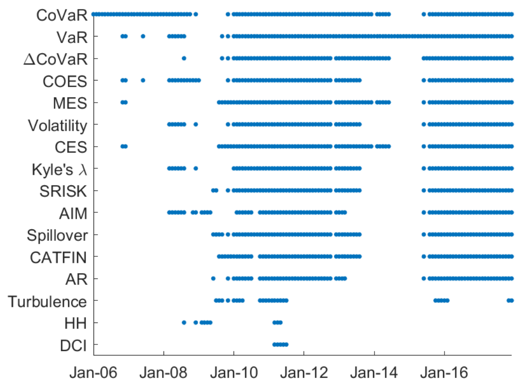
Figure 5.
The selected systemic measures in each time frame for the $TASSYRI_1$ according to the proxy of real economic activity at the specified horizon. The index is built recursively to perform the out-of-sample forecasting evaluation.



a) RSFHFS (high-level of sparsity)



b) HOUST (mid-level of sparsity)



c) PAYEMS (low-level of sparsity)

Notes: Panel a) The Retail Sales on Furniture and Home Furnishings Stores index (RSFHFS) at horizon h=12 (high-level of sparsity); Panel b) The New Privately Owned Housing Units Started variable (HOUST) at horizon h=4 (mid-level of sparsity); and Panel c) The Total Nonfarm Payrolls – All Employees figures (PAYEMS) at horizon h= 9 (low-level of sparsity).

Table 5.
Descriptive statistics of the $TASSYRI_1$ selected systemic risk measures over the period.

	(a)	(b)	(c)
	RSFHFS	HOUST	PAYEMS
Sparsity	High	Medium	Low
Average selected	2.11	5.14	8.01
Std Dev selected	3.16	4.30	5.54
Average Turnover	5.94%	6.73%	7.65%

Notes: For the selected systemic risk measures, the table reports the average (average selected measures), the standard deviation (Std Dev selected measures), and the average turnover. Column (a): The Retail Sales on Furniture and Home Furnishings Stores index (RSFHFS) at horizon *h=12* (high-level of sparsity); Column (b): The New Privately Owned Housing Units Started variable (HOUST) at horizon h=4 (mid-level of sparsity); and Column (c): The Total Nonfarm Payrolls–All Employees figures (PAYEMS) at horizon *h=9* (low-level of sparsity).

Finally, to monitor the strength of the relation between the degree of sparsity and the forecast performances of the TASSYRI, we compare the out-of-sample quantile forecasts for the three indexes considered in the graphical illustration. The closest competitor of TASSYRI_1 is, in these three cases, the PCQR_1, i.e., the first Principal Component of the classical PCA taking into account all the systemic risk measures. Table 6 reports the out-of-sample quantile forecast pseudo-R^2 relative to the baseline quantile model using only control variables.[27]

Table 6.

Overall 20th percentile shock forecasts for the RSFHFS at horizon 12 (high-level of sparsity), the HOUST at horizon 4 (medium-level of sparsity), and the PAYEMS at horizon 9 (low-level of sparsity).

	(a)	(b)	(c)
	RSFHFS	HOUST	PAYEMS
Sparsity	High	Medium	Low
TASSYRI_1	0.039**	0.038**	0.033***
TASSYRI_2	0.058**	0.012	0.05*
PCQR_1	0.023	0.0054	0.029
PCQR_2	0.020	0.022	0.066**
PQR	0.003	-0.0082	0.028

Notes: The table reports the out-of-sample quantile forecast R^2 (in percentage) relative to the baseline quantile model including only the controls for the Indexes of Systemic Risk Measures replicating the first Principal Component (TASSYRI_1), the first two Principal Components (TASSYRI_2), the first and second Principal Components from the Principal Components Quantile Regressions (PCQR_1 and PCQR_2), and the Partial Quantile Regression (PQR). Column (a): The Retail Sales on Furniture and Home Furnishings Stores index (RSFHFS) at horizon $h = 12$ (high-level of sparsity); Column (b): The New Privately Owned Housing Units Started variable (HOUST) at horizon $h = 4$ (medium-level of sparsity); and Column (c): The Total Nonfarm Payrolls–All Employees figures (PAYEMS) at horizon $h = 9$ (low-level of sparsity). Statistical significance at the 10%, 5%, and 1% levels are denoted by $*$, $**$, and $***$, respectively.

The TASSYRI_1 provides a significant R^2 for all three proxies of real economic activity, while the PCQR_1 does not exhibit a significant R^2. This indicates that TASSYRI_1 improves over the baseline case and PCQR_1 does not. These examples highlight how sparsity combined with the flexibility of our approach, where the TASSYRI is tailored to the target economic variable and forecast horizon, can improve the prediction of those economic variables. In this respect, the TASSYRI_1 overcomes the disadvantage of PCQR_1, where the Principal Component

[27] The full results for all the proxies of real economic activity are included in Web Appendix E (Table E1–E17).

is a linear combination of all the considered systemic risk measures and the index is unique for all target variables and forecast horizons. As stated by Giglio et al. (2016), the index might be understood as an imperfect version of "an unobservable systemic risk factor." A cross-sectional dimension reduction can enhance the forecasting power of single systemic risk measures by aggregating the relevant informative content. We take our proposed analysis one step further, as the unobservable risk factor might be associated with only a subset of the available systemic risk measures. The inclusion of irrelevant information drivers in the set of systemic risk measures leads to an increase in noise, with an impact on the forecast performance of the aggregate measure. Moreover, as the systemic risk factor might impact in different ways the various measures of economic activities and with different intensity over time, the introduction of tailored systemic risk indexes acts as a clear advantage. Finally, our empirical evidence shows that the ability of those imperfect measures of systemic risk may also vary over time, and the adoption of a rigid combination approach may not prove optimal.

On the Systemic Risk Measures in the TASSYRI

Finally, we consider the systemic risk measures that are included in the TASSYRI within the out-of-sample analysis in order to check if there exists a set of systemic risk measures that are more selected than others during the period under study. In this respect, Figure 6 shows the frequency of the selected systemic risk measures for each proxy of real economic activity across the 12 forecast horizons.[28]

Interestingly, a set of five measures emerge (red coloured bars in Figure 6) that are most often selected by the sparse and targeting approach for TASSYRI construction. Those measures

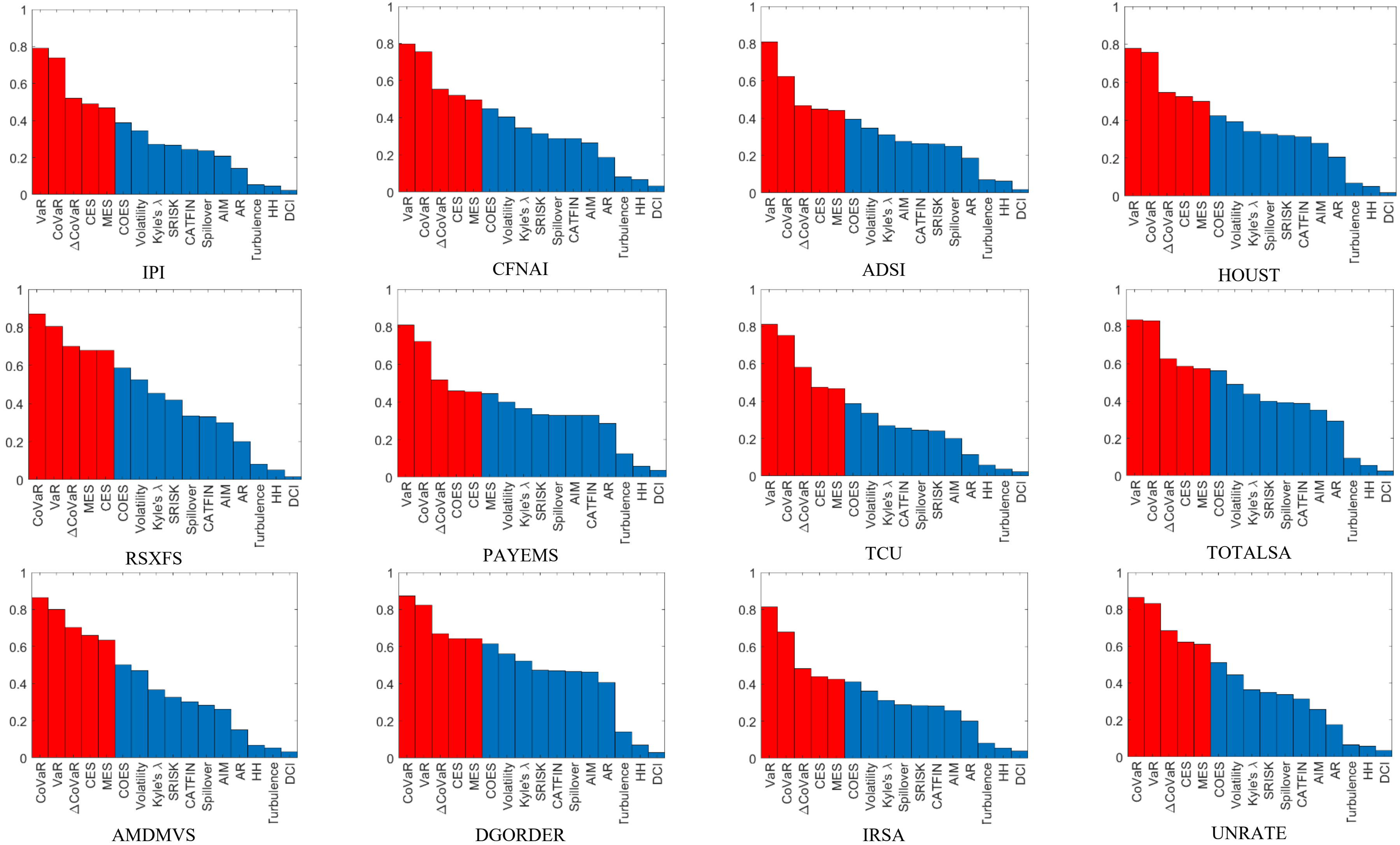
[28] Results for a given proxy of real economic activity at a given time horizon are available upon request.

are the VaR, the CoVaR, and the ΔCoVaR (Adrian and Brunnermeier, 2016); the MES (Acharya et al., 2012); and the CES (Banulescu & Dumitrescu, 2015).

It is worth noting that all those five systemic risk measures represent the risk dimension of financial losses and are derived by two central measures that play a pivotal role in risk management: The Value-at-Risk (VaR) and the Expected Shortfall (ES). Furthermore, the frequency of the selected systemic risk measures shows a similar behaviour also for the less selected systemic risk measures. Those are the Dynamic Causality Index (Billio et al., 2012), the Amihud Illiquidity Measure (Amihud, 2002), and the Turbulence Index (Kritzman and Li, 2010). The evidence we provide gives clear indication that the most relevant information related to the systemic risk is included in a small subset of the systemic risk measures. This further supports our proposal for a sparse and targeted systemic risk index. For instance, the VaR and CoVaR are selected approximately 80% of the time. The only exception is for the PAYEMS where the COES (Adrian and Brunnermeier, 2016) is selected instead of the MES.[29]

[29] We select five measures by looking at the patterns in Figure 6, where we note in several cases a drop in the frequency of most selected measures between the fifth and the sixth measures.

Figure 6.
The frequency of the selected systemic risk measures for the considered proxy of economic variables.



Notes: The red bar colour highlights the five most selected systemic risk measures for all the considered horizons in the period.

The target variables are: i) the Industrial Production Index (IPI); ii) the Chicago Fed National Activity Index (CFNAI); iii) the Aruoba-Diebold-Scotti Business Conditions Index (ADSI); iv) the New Privately Owned Housing Units Started (HOUST); v) the Advance Retail Sales Excluding Food Services (RSXFS); vi) the Total Nonfarm Payrolls–All Employees (PAYEMS); vii) the Total Capacity Utilisation (TCU); viii) Total Vehicle Sales (TOTALSA); ix) Manufacturers' Value of Shipments on Durable Goods (AMDMVS); x) Manufacturers' New Orders on Durable Goods (DGORDER); xi) Manufacturers–Inventories to Sales Ratio (IRSA); and xii) the unemployment rate (UNRATE).

Conclusion

This article introduces a novel approach to identify an aggregate measure of systemic risk, focusing on the left tail of an economic proxy at a given forecasting horizon. Our proposal complements and extends the literature on systemic risk and its non-linear association with economic downturns. The possibility of identifying an optimal subset of systemic risk metrics closely connected with a proxy of economic activity is a distinctive feature of our approach. The methodology allows us to account for the time-varying systemic risk-real economy nexus, which may further depend on the forecasting horizon. Our findings show that the ideal combination of systemic risk metrics may be time-varying; hence the adoption of a rigid combination approach might not be optimal. Sparsity and targeting result in superior forecasting ability with respect to the full information case, especially for forecasting horizons ranging above one month and up to one year. As with Giglio et al. (2016), we find no evidence that a specific combination of systemic risk measures dominates all others. Moreover, our analyses clearly highlight that most of the relevant systemic information is captured by five systemic risk measures, namely: the VaR, the CoVaR, and the ΔCoVaR (Adrian and Brunnermeier, 2016); the MES (Acharya et al., 2012); and the CES (Banulescu and Dumitrescu, 2015). Finally, regular monitoring of the index of systemic risk measures can serve as a complementary guideline to assess market conditions and the dynamics of the overall systemic risk as an early warning indicator (Billio et al., 2016; Engle, 2018) targeted to specific economic variables of interest.

Among possible future extensions of this work, we mention first the application of our methodology to other datasets, which would allow us to assess the differences across markets in response to some crises (Engle et al., 2015; Dungey et al., 2020). Second, extending the analysis to non-market value systemic risk measures as well as to other indicators tracking additional systemic dimensions would also be of interest; as examples, one might consider other illiquidity measures (Belkhir et al., 2020) or interconnectedness measures (Giudici et al., 2020; Bonaccolto et al., 2021).

Acknowledgments

We thank the Associate Editor, the General Editor Geert Bekaert, two anonymous referees, as well as Sylvain Benoit, Régis Breton, Laurent Clerc, Gilbert Colletaz, Olivier Darné, Alexis Direr, John Duca, Sylvain Friederich, Eric Girardin, Christophe Hurlin, Juan Angel Jimenez-Martin, Manfred Kremer, Catherine Lubochinsky, Valérie Mignon, Alfonso Novales, Daniel Pena, Carlos Vladimir Rodriguez Caballero, Esther Ruiz, Bernd Schwaab, Roger Stein, Bertrand Tavin, Anne-Gaël Vaubourg and Maria Helena Veiga, for constructive comments, positive feedbacks and encouragements when preparing this work, as well as Patrick Kouontchou and Sessi Tokpavi for previous fruitful collaborations on this topic. We also thank the participants to the IFABS Asia conference 2017 in Ningbo (China), to the 8th International Research Meeting in Business and Management at IPAG Business School in Nice (France), to the seminars at the Complutense University of Madrid (Spain) and to Carlos III University of Madrid (Spain), to the IFC2019 in Paris (France), to the IRMC2019 in Milan (Italy), to the FFM conference 2019 in Venice (Italy), to the SoFiE conference 2019 in Shanghai (China), to the AFFI conference 2019 in Québec (Canada) and to the CFE conference 2020 (online). All authors thank the GRI in Financial Services (www.globalriskinstitute.org), as well as the Louis Bachelier Institute for support. Michele Costola acknowledges financial support from the Marie Skłodowska-Curie Actions, European Union, Seventh Framework Program HORIZON 2020 under REA grant agreement n. 707070. He also gratefully acknowledges research support from the Research Center SAFE, funded by the State of Hessen initiative for research LOEWE. 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 and from project PRIN2017 "HiDEA: Advanced Econometrics for High-frequency Data", 2017RSMPZZ. Extra materials linked to this article can be found at: www.systemic-risk-hub.org.

Bibliography

Acharya, V., Brownless, C., Engle, R., Farazmand, F., Richardson, M., 2013. Measuring systemic risk, in Roggi-Altman (dir.), Managing and Measuring Risk: Emerging Global Standards and Regulation after the Financial Crisis, World Scientific Series in Finance, 65-98.

Acharya, V., Engle, R., Richardson, M., 2012. Capital shortfall: a new approach to ranking and regulating systemic risks. American Economic Review 102 (3), 59-64.

Acharya, V., Pedersen, L., Philippon, T., Richardson, M., 2017. Measuring systemic risk. Review of Financial Studies 30 (1), 2-47.

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

Allen, L., Bali, T., Tang, Y., 2012. Does systemic risk in the financial sector predict future economic downturns? Review of Financial Studies 25 (10), 3000-3036.

Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time-series effects. Journal of Financial Markets 5 (1), 31-56.

Aruoba, S.B., Diebold, F.X., Scotti, C., 2009. Real-time measurement of business conditions. Journal of Business and Economic Statistics 27 (4), 417-427.

Baker, S. R., Bloom, N., Davis, S., 2016. Measuring economic policy uncertainty. The Quarterly Journal of Economics 131 (4), 1593-1636.

Banulescu, G.-D., Dumitrescu, E.-I., 2015. Which are the SIFIs? A component expected shortfall approach to systemic risk. Journal of Banking and Finance 50 (C), 575-588.
Belkhir, M., Saad, M., Samet, A., 2020. Stock extreme illiquidity and the cost of capital. Journal of Banking and Finance 112, 105281, 1-21.
Ben Bouheni, F., Hasnaoui, A., 2017. Cyclical behavior of the financial stability of eurozone commercial banks. Economic Modelling 67, 392-408.
Benoit, S., Colliard, J.-E., Hurlin, C., Pérignon, C., 2017. Where the risks lie: a survey on systemic risk. Review of Finance 21 (1), 109-152.
Billio, M., Casarin, R., Costola, M., & Pasqualini, A. 2016. An entropy-based early warning indicator for systemic risk. Journal of International Financial Markets, Institutions and Money 45, 42-59.
Billio, M., Getmansky, M., Lo, A., Pelizzon, L., 2012. Econometric measures of connectedness and systemic risk in the finance and insurance sectors. Journal of Financial Economics 104 (3), 535-559.
Black, L., Correa, R., Huang, X., Zou, H., 2016. The systemic risk of European banks during the financial and sovereign debt crises. Journal of Banking and Finance 63, 107-125.
Bonaccolto, G., Caporin, M., Maillet, B., 2021. Dynamic large financial networks *via* conditional expected shortfalls. Forthcoming in the European Journal of Operations Research.
Brownlees, C., Engle, R., 2017. SRISK: a conditional capital shortfall index for systemic risk measurement. Review of Financial Studies 30 (1), 48-79.
Caporin, M., 2008. Evaluating value-at-risk measures in the presence of long memory conditional volatility. The Journal of Risk 10 (3), 79 -110.
Daniell, P. J., 1946. Contribution to Discussion in the Symposium on Autocorrelation in Time Series, Journal of the Royal Statistical Society, Supplement 8, 88-90.
Di Cesare, A., Rogantini Picco, A., 2018. A survey of systemic risk indicators. Banca d'Italia Occasional Paper n°438.
Diebold, F., Mariano, R., 1995. Comparing predictive accuracy. Journal of Business and Economic Statistics 13, 253-265.
Diebold, F., Yilmaz, K., 2009. Measuring financial asset return and volatility spillovers, with application to global equity markets. Economic Journal 119 (534), 158-171.
Dungey, M., Flavin, Th., Lagoa-Varela D., 2020. Are banking shocks contagious? Evidence from the Eurozone. Journal of Banking and Finance 112, 105386, 1-17.
Engle, R., 2018. Systemic risk 10 years later. Annual Review of Financial Economics, 10, 125-152.
Engle, R., Jondeau, E., & Rockinger, M. (2015). Systemic risk in Europe. Review of Finance, 19(1), 145-190.
Giglio, S., Kelly, B., Pruitt, S., 2016. Systemic risk and the macroeconomy: an empirical evaluation. Journal of Financial Economics 119 (3), 457-471.
Granger, C., 1969. Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37, 424-438.
Gilchrist, S., Zakrajsek, E., 2012. Credit spreads and business cycle fluctuations. American Economic Review 102 (4), 1692-1720.
Giudici, P., Sarlin, P., Spelta, A., 2020. The interconnected nature of financial systems: direct and common exposures. Journal of Banking and Finance 112, 105149, 1-15.
Hansen, L. P. (2014). 1. Challenges in Identifying and Measuring Systemic Risk". Risk Topography, edited by Markus Brunnermeier and Arvind Krishnamurthy, Chicago: University of Chicago Press, 2014, 15-30.
Harvey, D., Leybourne, S., Newbold, P., 1997. Testing the equality of prediction mean squared errors. International Journal of Forecasting 13 (2), 281-291.

Hastie, T., Tibshirani, R., Wainwright, M., 2015. Statistical learning with sparsity: the lasso and generalizations, in Chapman & Hall/CRC Monographs on Statistics & Applied Probability 143.
Hiemstra, C., Jones, J., 1994. Testing for linear and nonlinear granger causality in the stock price-volume relation. Journal of Finance 49 (5), 1639-1664.
Holló, D., Kremer, M., Lo Duca, M., 2012. CISS - A composite indicator of systemic stress in the financial system. ECB working paper 1426, 51 pages.
Hong, Y., Liu, Y., Wang, S., 2009. Granger causality in risk and detection of extreme risk spill-over between financial markets. Journal of Econometrics 150 (2), 271-287.
Kelly, B., Pruitt, S., 2015. The three-pass regression filter: a new approach to forecasting using many predictors. Journal of Econometrics 186 (2), 294-316.
Koenker, R., Bassett, G., 1978. Regression quantiles. Econometrica: Journal of the Econometric Society, 33-50.
Koenker, R., Machado, J. A., 1999. Goodness of fit and related inference processes for quantile regression. Journal of the American Statistical Association 94 (448), 1296-1310.
Kritzman, M., Li, Y., 2010. Skulls, financial turbulence, and risk management. Financial Analysts Journal 66 (5), 30-41.
Kritzman, M., Li, Y., Page, S., Rigobon, R., 2011. Principal components as a measure of systemic risk. Journal of Portfolio Management 37 (4), 112-126.
Kyle, A., 1985. Continuous auctions and insider trading. Econometrica, 53 (6), 1315-1336.
Lopez, J., 1999. Regulatory evaluation of value-at-risk models. The Journal of Risk 1, 37-64.
Louzis, D., Vouldis, A., 2017. A methodology for constructing a financial systemic stress index: an application to Greece. Economic Modelling 29 (4), 1228-1241.
Morley, J. (2016). Macro-finance linkages. Journal of Economic Surveys, 30(4), 698-711.
Nucera, F., Schwaab, B., Koopman, S., Lucas, A., 2016. The information in systemic risk rankings. Journal of Empirical Finance 38, 461-475.
Pástor, L., Stambaugh, R., 2003. Liquidity risk and expected stock returns. Journal of Political Economy 111, 642-685.
Petrella, L., Laporta, A. G., Merlo, L., 2019. Cross-country assessment of systemic risk in the European stock market: evidence from a CoVaR analysis. Social Indicators Research 146 (1-2), 169-186.
Tibshirani, R., 1996. Regression shrinkage and selection *via* the lasso. Journal of the Royal Statistical Society (B) 58 (1), 267-288.
Zou, H., Hastie, T., 2005. Regularization and variable selection *via* the elastic net. Journal of the Royal Statistical Society (B) 67, 301-320.
Zou, H., Hastie, T., Tibshirani, R., 2006. Sparse principal component analysis. Journal of Computational and Graphical Statistics 15 (2), 265-286.