

Kremer, Manfred; Chavleishvili, Sulkhan

Conference Paper

Measuring Systemic Financial Stress and its Impact on the Macroeconomy

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2021: Climate Economics

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Kremer, Manfred; Chavleishvili, Sulkhan (2021) : Measuring Systemic Financial Stress and its Impact on the Macroeconomy, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2021: Climate Economics, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/242346>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Measuring Systemic Financial Stress and its Impact on the Macroeconomy*

Sulkhan Chavleishvili
European Central Bank

Manfred Kremer
European Central Bank

Version: February 9, 2021

Abstract

This paper proposes a general statistical framework for systemic financial stress indexes. Several existing index designs can be represented as special cases. We introduce a daily variant of the ECB's CISS for the euro area and the US. The CISS aggregates a representative set of stress indicators using their time-varying cross-correlations as systemic weights, like portfolio risk is computed from the risk characteristics of individual assets. A bootstrap algorithm delivers test statistics. A linear VAR suggests that systemic stress is the main cause behind the Great Recession, while its contribution to the Covid-19 crisis appears limited. A quantile VAR features stronger real effects of financial stress in the worst states of the economy.

JEL classification: C14, C31, C43, C53, E44, G01.

Keywords: Financial crisis; Financial stress index; Macro-financial linkages; Quantile vector autoregression; Systemic risk.

*Chavleishvili: Financial Research Division, European Central Bank, Frankfurt am Main, Germany, e-mail: sulkhanchavleishvili@gmail.com; Kremer: Financial Research Division, European Central Bank, Frankfurt am Main, Germany, e-mail: manfred.kremer@ecb.europa.eu. We thank Thomas Kostka, Oliver Grothe and Vincenzo Quadrini for useful comments and suggestions. We also thank Paul Konietschke for outstanding research assistance. The views expressed in this paper are those of the authors and do not necessarily represent those of the European Central Bank.

1. Introduction

Finance and growth are twin sisters. Finance helps overcome frictions in the real sector arising from transaction and information costs, thereby supporting economic agents' savings and investment decisions and thus capital accumulation and growth (Levine (2005), Beck (2014)). A prosperous real economy, in turn, provides profitable opportunities for the financial system to invest in and develop. However, this mutual dependence between finance and growth holds true for good and bad times alike. The global financial crisis (GFC) reminded us that financial development may at times become the root cause of a deep financial and economic crisis. This ambiguous role of finance reflects the fact that the financial sector itself is prone to market failures induced by, e.g., externalities, information asymmetries, incomplete markets or limited human cognitive abilities (Bisias, Flood, Lo, and Valavanis (2012)). When such financial frictions intensify and predominate, giving rise to widespread financial stress, they tend to have severe repercussions for the real economy. A better understanding of these important macro-financial linkages requires—apart from suitable theoretical models—meaningful empirical measures of financial stress caused by severe financial frictions.

A great many of indicators exist which measure stress in individual market segments. Each of these indicators captures certain market- and instrument-specific stress symptoms like increased market volatility or wider default and liquidity risk premia, which in turn reflect stress reactions like increased uncertainty, higher risk aversion, or run-like phenomena such as flight-to-safety and flight-to-liquidity behaviours. For instance, volatility measures derived from options prices provide information about market participants' degree of risk aversion and uncertainty (Bekaert and Hoerova (2014)); the most well-known of such volatility measures, the VIX, is often taken as a general “fear gauge” and thus as an indicator of financial stress (Carr (2017)). Such standard indicators form the backbone of every financial stability surveillance toolkit. However, the sheer amount of individual stress measures complicates the task of inferring whether stress observed in one particular market segment is either idiosyncratic or instead a more systemwide phenomenon. Sometimes you “can't see the wood for the trees”.

Financial stress indexes are one way to synthesise such scattered information. A financial stress index (FSI) quantifies the aggregate stress level by compressing a certain number of individual stress indicators into a single statistic. This paper proposes a general statistical framework for *systemic* financial stress indexes that is firmly rooted in standard definitions of systemic risk. Systemic risk can be characterised as the risk that financial instability becomes so widespread that it severely disrupts the provision of financial services to the

broader economy with significant adverse effects on growth and employment (de Bandt and Hartmann (2000) and Freixas, Laeven, and Peydró (2015)). We interpret systemic stress as an *ex post* measure of systemic risk, i.e. a measure of the degree to which systemic risk has materialised in each point in time. A systemic FSI can thus be seen as a coincident financial (in)stability indicator. The statistical framework defines systemic stress as a state of the financial system in which a representative set of stress indicators is considered extremely high and strongly co-dependent. The composite indicator results from a matrix association index that combines two matrices quantifying the extremeness and the co-dependence hypotheses.

We furthermore propose an enhanced variant of the ECB’s Composite Indicator of Systemic Stress (CISS, pronounced /kis/, originally developed by Hólló, Kremer, and Lo Duca (2012)) as a nonparametric operationalisation of the framework. This paper thus provides a rigorous statistical foundation of the CISS. In the empirical application, we introduce new daily CISS series for the US and the euro area. Both indicators aggregate 15 components capturing stress symptoms in money, bond, equity and foreign exchange markets. All raw input variables are first transformed into relative ranks by applying the probability integral transform. System-wide stress is then computed as the average cross-product of all pairs of transformed indicators (measuring extremeness) weighted by its time-varying rank correlation (measuring co-dependence) in the same way as portfolio risk is computed from the risk of individual assets in standard finance theory. The CISS accordingly puts more weight on situations in which stress becomes widespread and thus systemic. The correlations can also capture externalities like contagion or spillover effects from one part of the financial system to the financial system at large, a feature that any systemic risk measure should pay attention to (Freixas et al. (2015)). From a statistical point of view, the various steps in the design of the CISS are geared towards delivering a composite indicator which does not suffer from look-ahead bias, is sufficiently robust to outliers, largely unaffected by differing distributional properties of the underlying raw data, and is easy to update. We further propose a bootstrap algorithm to test whether the CISS exceeds a level that can be considered as normal.

The final part of the paper addresses macro-financial linkages. It is a well-known stylised fact that systemic financial crises produce large losses in output and employment. Any meaningful measure of systemic financial stress should therefore be able to replicate this fact. The present paper runs linear and quantile vector autoregressions (VAR) based on euro area data for the CISS, the Purchasing Manager’s Index (PMI) and annual real GDP growth. The linear VAR indeed confirms the CISS as a statistically and economically significant driver of economic activity. Simulation results suggest a dominant role of financial stress shocks in explaining the deep recession observed during the GFC. This is different from the (still

ongoing) Covid-19 crisis, in which financial stress shocks, despite the huge initial surge in the CISS, only play a minor role compared to aggregate output shocks. Linear VARs, however, can only capture conditional mean effects between the model variables, while the relationship between financial stress and economic activity may be much stronger during crisis times. We address this potential asymmetry by reestimating the model using the Quantile VAR (QVAR) methodology recently developed by Chavleishvili and Manganeli (2019). We indeed find stronger effects of the CISS on both measures of economic activity in the lower tails of their distributions, while the model displays no further significant asymmetries.

The paper is organised as follows. The next section presents a selective literature survey and outlines how the paper contributes to the various fields covered in the survey. Section 3 describes the general statistical framework for estimating a systemic stress index. Section 4 presents and discusses the CISS as a practical implementation of the framework using data for the US and the euro area. Section 5 suggests a statistical inference tool to address the joint extremeness-codependence hypothesis underlying the CISS. Section 6 studies empirical linkages between the CISS and measures of economic activity applying a linear VAR and a nonlinear QVAR. Section 7 concludes, and several Appendices contain supplementary information.

2. Related Literature

The paper contributes to several strands of literature. First, our work speaks to the literature on financial crisis indicators. Crisis indicators aim to identify periods of extreme stress in either one or several systemically important segments of a country’s financial system. The most well-known and widely-used indicators assign ordinal values to crisis and non-crisis periods based on qualitative information and events. For instance, Laeven and Valencia (2008, 2013, 2018) create annual crisis dummies capturing systemic distress in the banking system if they find significant signs of distress accompanied by significant banking policy interventions. Reinhart and Rogoff (2009) develop a composite crisis index by summing up the values of five crisis dummies identifying severe distress in the banking system, in currency markets, in domestic and external debt markets, and in inflation conditions, thereby measuring the severity of each crisis episode in a specific country. Romer and Romer (2017a,b) go one step further in differentiating crisis severity. Studying the semiannual OECD Economic Outlook as a real-time source of information, they identify and categorise financial stress events into five groups, from pure credit disruptions, minor crisis, moderate crisis, major crisis up to extreme crisis. Severity of stress within each group is further differentiated into three subcategories (minus, regular, and plus events), such that they obtain a “measure of

financial distress” with values ranging from 0 for non-distress periods to a maximum of 15 for an extreme crisis-plus. In contrast to qualitative ordinal crisis indicators, FSIs apply a statistical approach to measure the severity of financial distress on a continuous scale. The relationship between FSIs and qualitative crisis indicators resembles that between composite real-time business cycle indicators like that of Aruoba, Diebold, and Scotti (2009) and qualitative recession dummies as those published by the NBER for the United States. Romer and Romer (2017a) acknowledge that financial stress occurs along a continuum, and that “(t)reating a continuous variable as discrete introduces measurement error, both because the variation across crises is omitted and because a small inaccuracy in evaluating an observation can cause a large change in the value assigned to it.” Financial stress indexes like the CISS avoid such problems. Like Reinhart and Rogoff (2009) and Romer and Romer (2017a), the CISS considers stress in several systemically important segments of the financial system, but aggregates stress in these market segments based on weights derived from sound economic and statistical principles operationalising the concept of systemic risk. By construction, the CISS also facilitates a finer delimitation of the start and end dates of crisis episodes.

Figure 1 plots the three qualitative crisis indicators for the US together with our CISS index, which is bounded between zero and one by construction. Panel A shows the Reinhart/Rogoff and the Laeven/Valencia indicators, and Panel B the Romer/Romer indicator at a monthly frequency, assigning the value of the annual or semiannual data for all months of a year or half-year, respectively. There is clearly some positive association between the qualitative indicators and the CISS during the crisis years indicated by the former. This is particularly true for the Romer/Romer indicator which is explicitly built to better reflect crisis severity. This notwithstanding, the US financial system seem to have experienced several major stress events which are not captured by any of the qualitative indicators. One may argue that FSIs can only measure financial stress net of the impact of policy interventions, while most qualitative indicators take the size of policy reactions explicitly into account when assessing a potential crisis event.¹ As a result, FSIs categorise crisis severity according to the measured stress level which, however, may reflect very different intensities of policy intervention. This notwithstanding, it may still be possible to construct appropriate counterfactuals which allow cleaning the FSI from the effects of policy interventions, or, which is the same, quantifying the impact of policy interventions on financial distress which would be interesting in its own right.

Second, our paper links the literature on FSIs with that on empirical measures of systemic risk. For overviews of the FSI literature see Illing and Liu (2006) and Kliesen, Owyang,

¹An exception is Carlson, Lewis, and Nelson (2012) which uses policy interventions in identifying crisis events and constructing the FSI.

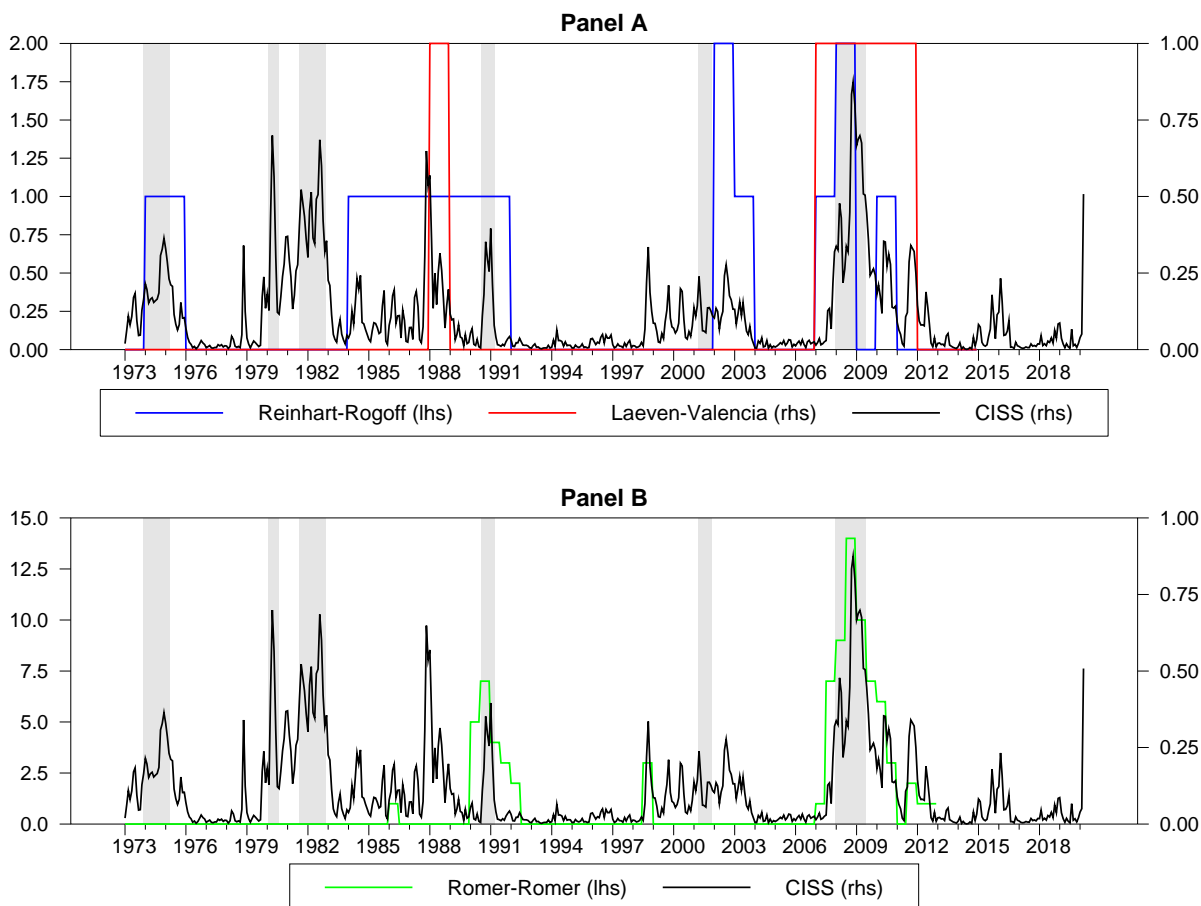


Fig. 1. The figure plots the Composite Indicator of Systemic Stress (CISS, black line) along with three qualitative financial crisis indicators for the United States. Panel A shows the crisis indicators of Reinhard and Rogoff (2009, blue line) and Laeven and Valencia (2018, red line); the indicator of Romer and Romer (2017, green line) is displayed in Panel B. Data is shown monthly from January 1973 to March 2020 for the CISS, annual until 2017 for Laeven and Valencia, annual until 2012 for Reinhard and Rogoff, and semiannual until 2012:2 for Romer and Romer. National Bureau of Economic Research recessions are represented by the shaded areas.

and Vermann (2012), and systemic risk measures are surveyed in Bisias et al. (2012) and Freixas et al. (2015), Chapter 7. The seminal paper on FSIs is Illing and Liu (2006). They discuss several approaches to aggregate a given set of individual stress indicators into a composite stress index. The preferred FSI is chosen according to which variant performs best in capturing crisis events in the Canadian financial system identified on the basis of a survey among Bank of Canada policy-makers and staff. The preferred FSI comprises 11 daily financial market variables aggregated on the basis of weights determined by the relative size of the market to which each of the indicators pertains. Cardarelli, Elekdag, and Lall

(2011) present a monthly financial stress index for 17 advanced economies computed as the arithmetic mean of twelve standardised financial stress indicators. Nelson and Perli (2007) and Carlson, King, and Lewis (2011) present a weekly financial fragility indicator for the US computed in two steps from twelve market-based financial stress measures. The standardised input series are first reduced to three summary indicators, namely a level factor, a rate-of-change factor and a correlation factor. In the second step, the financial fragility indicator is computed as the fitted probability from a logistic regression model with the three summary indicators as explanatory variables and a binary pre-defined crisis indicator as the dependent variable. Refining the last step of that approach, Grimaldi (2010) computes a weekly FSI for the euro area, where the binary crisis indicator is systematically derived from crisis events identified on the basis of a keyword-search algorithm applied to relevant parts of the ECB Monthly Bulletin. The Fed Cleveland FSI developed by Oet, Eiben, Bianco, Gramlich, and Ong (2011) integrates 11 daily financial market indicators grouped into four sectors. The raw indicators are normalised by applying the probability integral transform, and are then aggregated into the composite indicator by computing a weighted average with time-varying credit weights proportional to the quarterly financing flows in the four markets. Hakkio and Keeton (2009) construct a monthly FSI for the US applying principal components analysis based on the idea that financial stress is the latent factor driving the observed correlation between the input series. The weekly index developed by Brave and Butters (2011, 2012) also builds on factor analysis but is more complex and sophisticated than its competitors in terms of the number and the heterogeneity of the input data and the statistical indicator design. The computation of what is published as the Federal Reserve Bank of Chicago's National Financial Conditions Index (NFCI) is cast into a dynamic factor model in state-space form which includes a maximum of 100 indicators, where Kalman filtering takes account of the missing data problem resulting from the different sample lengths and frequencies of the input data.

These two factor model approaches can also be regarded as empirical systemic risk measures. Both models include stress indicators from a broad range of financial market segments, and both determine the relative weights of the input series in the composite indicator based on their sample cross-correlations which can be interpreted as applying systemic risk weighting (Brave and Butters (2011)). Principal component analysis of a set of asset returns is also behind systemic risk measures proposed by Kritzman, Li, Page, and Rigobon (2011) and Billio, Getmansky, Lo, and Pelizzon (2012). The "absorption ratio" by Kritzman et al. (2011) takes daily returns of 51 US stock market sectors and computes, over a moving 500-day window, the fraction of the total variance of these returns explained by the first ten eigenvectors of the covariance matrix. The ratio is higher if there is stronger commonality

among sector returns. The idea is that high values of the absorption ratio indicate a state of enhanced system fragility since a given shock to financial stability would tend to propagate more quickly and broadly when markets are more closely linked. In a similar vein, Billio et al. (2012) takes the monthly returns of 100 individual hedge funds, banks, broker/dealers, and insurance companies and estimates, over 36-month rolling windows, the “cumulative risk fraction” that corresponds to the fraction of total return variance explained by a certain number of statistically significant principal components. The turbulence index of Kritzman and Li (2010) aggregates asset returns based on the Mahalanobis distance. The index shall capture the statistical “unusualness” of a set of returns given their historical pattern of behaviour.

Our paper presents a general statistical framework to measure systemic financial stress, a notion that blends the ideas behind financial stress indexes and systemic risk indicators. In Appendix C we illustrate that several of the above-mentioned indicator designs from both fields can be represented as special cases of our general statistical framework. According to the taxonomy of systemic risk measures by the analytical time horizon (pre-, contemporaneous, and post-event) proposed in Bisias et al. (2012), our general systemic stress index counts as a contemporaneous systemic risk measure. Since our general stress index—and thus also the CISS—combines measures of extremeness and co-dependence (connectedness), it fits in both the “fragility” and the “crisis monitoring” subcategories of contemporaneous systemic risk measures similar to Kritzman et al. (2011), Billio et al. (2012) and the aggregate CoVaR (Adrian and Brunnermeier (2016)) and SRISK (Brownlees and Engle (2017)) measures. While the main purpose of the latter two papers is to measure the contributions of individual financial institutions to the aggregate systemic risk in the sector comprising these institutions, the CISS measures the degree of connectedness of various aggregate market segments and how it contributes to the level of systemic stress in the financial system as a whole. Both perspectives can thus be seen as complementary. In contrast to the factor models of Hakkio and Keeton (2009) and Brave and Butters (2012)—which rely on sample cross-correlations—, we estimate state-dependent connectedness as autoregressive cross-correlations borrowing from the GARCH literature. The PCA-approaches of Kritzman et al. (2011) and Billio et al. (2012) also measure time-variation in cross-correlations, but do so by way of fixed moving-window estimation (0-1-weighting of historical information). In contrast, our exponentially-weighted-moving-average approach puts decaying weights on more distant information, which proves flexible enough to also capture abrupt changes in correlation patterns as they typically occur during periods of stress. Last but not least, our paper amends Hóllo et al. (2012) by providing further support to the basic CISS design.²

²The statistical concept of the CISS has become the blueprint for the financial stress indexes of several

For instance, we substantiate the use of relative ranks (the probability integral transform) as a more robust and homogeneous transformation of the raw input indicators. The paper also introduces an enhanced daily variant of the CISS for the US and the euro area which has facilitated the real-time monitoring of financial stability conditions since the outbreak of the Covid-19 crisis.³

Third, the paper contributes to the rich literature on the real effects of financial distress. In linear growth prediction frameworks, several papers capture financial stress by using FSIs (including the CISS) or financial conditions indexes (FCIs)—which are often similarly constructed—and uniformly find economically and statistically significant effects (Hakkio and Keeton (2009), Cardarelli et al. (2011), Carlson et al. (2011), Mallick and Sousa (2013), Dovern and van Roye (2014), Kremer (2016), Hatzius, Hooper, Mishkin, Schoenholtz, and Watson (2010)). In line with theoretical predictions, studies applying nonlinear regression frameworks find stronger real effects in bad states of the world. The set of nonlinear approaches includes, *inter alia*, threshold VARs (Hóllo et al. (2012)), Markov-switching VARs (Davig and Hakkio (2010), Hubrich and Tetlow (2015), Hartmann, Hubrich, Kremer, and Tetlow (2015)), quantile VARs (Chavleishvili and Manganelli (2019), Chavleishvili, Engle, Fahr, Kremer, Manganelli, and Schwaab (2021)), and single-equation quantile regressions as common in the recent growth-at-risk literature (Adrian, Boyarchenko, and Giannone (2019), Figueres and Jarociński (2020)). Similar studies using alternative measures of financial distress (qualitative crisis variables, systemic risk indicators or individual asset price indicators such as credit spreads or stock market volatilities) also find strong and, where applicable, enhanced state-dependent effects of financial distress on economic activity (Romer and Romer (2017a), Giglio, Kelly, and Pruitt (2016), Adrian et al. (2019), Bloom (2009), Gilchrist and Zakrajšek (2012)). Giglio et al. (2016), Adrian et al. (2019) and Figueres and Jarociński (2020) demonstrate that composite indicators perform superior in terms of predictive power than a broad range of single indicators capturing financial stress and/or systemic risk.

Our paper complements this literature by running linear and quantile VARs on euro area

financial stability authorities, such as the Swedish Riksbank, Norges Bank, Bank of England, Banco de España, the Spanish Comisión Nacional del Mercado de Valores, Banco de Portugal, Bank of Greece, Czech National Bank, European Securities and Markets Authority (ESMA), Peoples Bank of China, Bank Negara Malaysia, Banco de la República Colombia, and Bank of Jamaica.

³Compared to the older version of the CISS, the enhanced version first estimates asset volatilities as integrated GARCH processes which allows moving to a daily computation of the index. Second, the new version aggregates the index components in one step, without prior aggregation into market-specific subindexes; it therefore uses the full 15×15 matrix of cross-correlations instead of the 5×5 correlation matrix of the previous version. Third, the new variant uses equal weights per indicator instead of segment-specific real-impact weights, the latter being often perceived as arbitrary. Fourth, the set of input series changes somewhat and becomes more harmonised across the US and the euro area indexes. Fifth, the euro area CISS carries a longer data history, starting in 1980.

data, and it corroborates previous findings of strong conditional mean effects of financial stress on future economic growth, effects which become amplified in the lower tails of the growth distribution. While Adams, Adrian, Boyarchenko, Giannone, Liang, and Qian (2020) and Figueres and Jarociński (2020) also use our new CISS for the US and the euro area, respectively, in quantile growth-at-risk regressions, our QVAR framework allows us to estimate feedback effects and to disentangle direct from indirect effects between the CISS and our two measures of economic activity; including the PMI likely achieves a better identification of systemic stress and economic activity shocks. Our paper furthermore quantifies the macroeconomic relevance of systemic stress by performing conditional projections based on historical shocks, taken from the linear VAR, to demonstrate the different nature of the GFC and the Covid-19 crisis. The projection for the Covid-19 event serves as a practical example of a real-time out-of-sample forecasting exercise at the outset of the crisis, representing the information available to policymakers at the time.

3. A general statistical framework for measuring systemic financial stress

Based on the theoretical definition of systemic financial stress discussed in the introduction, this section proposes a general statistical framework to measure systemic stress by means of a composite indicator, i.e. a systemic financial stress index (FSI). A systemic FSI \mathcal{S}_t can be thought of a statistical function that combines three types of ingredients: i) an N -dimensional vector of original stress indicators x_t for $t = 1, \dots, T$; ii) a conformable vector of “portfolio weights” w_t ; and iii) some matrix or vector of systemic risk weights \mathcal{C}_t . All stress indicators are constructed from the raw data such that they satisfy the following assumption:

Assumption 3.1. *Original stress indicators $x_{i,t}$ increase in the level of stress, i.e. stress at time t_1 is strictly higher (lower) than stress at time t_2 if $x_{i,t_1} > x_{i,t_2}$ ($x_{i,t_1} < x_{i,t_2}$).*

Since the original stress indicators are usually measured on different scales or on a common scale with widely differing ranges, a meaningful aggregation typically requires some form of normalisation for which we formulate a minimum requirement:

Assumption 3.2. *Original stress indicators $x_{i,t}$ are transformed into normalised stress factors $z_{i,t}$ by applying a monotone transformation $x_{i,t} \rightarrow g(x_{i,t}) \equiv z_{i,t}$ which is increasing, i.e. $u > v \Rightarrow g(v) > g(u)$.*

For instance, converting the raw scores of the original stress indicators into z-scores—by subtracting the sample mean from an individual raw score and then dividing the difference by the sample standard deviation—is the most common monotone transformation applied in the literature. However, as we will argue later, standardisation (z-score conversion) of variables has certain comparative disadvantages over our preferred alternative, the probability integral transform, in the context of FSIs.

For aggregation, stress factors $z_{i,t}$ are assigned portfolio weights $w_{i,t}$ with $0 < w_{i,t} < 1$, $i = 1, \dots, N$ and satisfying the adding-up constraint $\sum_{i=1}^N w_{i,t} = 1$. These assumptions simply rule out negative and zero portfolio weights, the latter assuming that none of the selected stress indicators can be regarded redundant. We borrow the term “portfolio weights” from standard finance portfolio theory where the weights $w_{i,t}$ would capture the share of asset i in the overall asset portfolio. Hence, in the present context a natural analogy of portfolio shares would be the share of a market segment covered by a certain stress indicator in the overall size of the financial system as measured, e.g., by the outstanding amounts of financial instruments issued in that market segment. Since the structure of a financial system typically changes over time, such size-based weights can be made time-varying which is why we add a time index to w_t .⁴ However, the easiest and most common way to calibrate w_t is to assume constant equal weights per indicator, $w_i = 1/N$.

A systemic FSI differs from standard FSIs by introducing some weighting of indicators that operationalises the concept of systemic risk. The main idea behind systemic weighting is that in order to assess the broader implications of stress in one particular market segment, one has to take into account how this stress relates to the stress in the other market segments, reflecting potentially state-dependent degrees of market interconnectedness, stress spillovers or contagion risks. If stress spreads widely, it may cause a systemic crisis with large macroeconomic costs, and the inherent non-linearity may be paraphrased with the famous Aristotle quote, “The whole is greater than the sum of the parts”. Allowing for bilateral differences in the somehow measured connectivity between stress indicators, the systemic risk weights may be conceived of some $N \times N$ matrix function \mathcal{C}_t .

We now approach statistical measurement of systemic financial stress. If the individual stress indicators satisfy Assumptions 3.1 and 3.2, then any increasing function of the vector z_t could in principle serve as a financial stress index. However, not every function will capture what we consider as characteristic for financial crises, namely widespread financial stress. To be more concrete, we define systemic financial stress as follows:

Definition 1 (Systemic Financial Stress). *Systemic financial stress is a state of the financial*

⁴As an alternative, Hólló et al. (2012) estimate weights based on the relative predictive power of market-specific subindexes of stress for industrial production growth.

system in which a representative set of stress factors $z_{i,t}$ are extremely high and strongly co-dependent.

According to Definition 1, systemic financial stress rules out cases in which high stress remains locally contained and does not spill over to other major parts of the financial system so that the general level of cross-market co-dependence remains weak. We can therefore define systemic stress also as the joint hypothesis (\mathcal{H}_t) that individual stress factors $z_{i,t}$ are extremely high (“extremeness” hypothesis $\mathcal{H}_{1,t}$) and that they are jointly so as a result of strong co-dependence (“co-dependence” hypothesis $\mathcal{H}_{2,t}$):

$$\mathcal{H}_t \equiv \{\mathcal{H}_{1,t} \wedge \mathcal{H}_{2,t}\}. \quad (1)$$

These two hypotheses are distinguished since extremely high stress is a necessary but not a sufficient condition for strong co-dependence: $\mathcal{H}_{1,t} \Rightarrow \mathcal{H}_{2,t}, \mathcal{H}_{1,t} \not\Leftarrow \mathcal{H}_{2,t}$. Put differently, while system-wide extreme stress implies strong co-dependence, the converse is not true since stress factors can also be highly correlated because they are jointly low, and not high. Hence, we can characterise systemic stress or financial instability as a state of “co-extremeness”, i.e., extremeness cum co-dependence, rendering the notion of widespread stress more concrete.⁵

The next definition introduces the last building block of our proposed general statistical framework for a systemic FSI.⁶

Definition 2 (Systemic Financial Stress Index). *Let the extremeness hypothesis $\mathcal{H}_{1,t}$ and the co-dependence hypothesis $\mathcal{H}_{2,t}$ be formalised by the $N \times N$ bounded real-valued matrices \mathcal{E}_t and \mathcal{C}_t , respectively. A systemic financial stress index \mathcal{S}_t can be defined as a matrix association index quantifying the joint hypothesis expressed in Equation (1):*

$$\mathcal{S}_t \equiv \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N (\mathcal{E}_t)_{i,j} (\mathcal{C}_t)_{i,j}. \quad (2)$$

The scaling factor $1/N^2$ just reflects the assumption of equal weighting, i.e., $w_i = 1/N$. The extremeness matrix \mathcal{E}_t is often a simple function of time- t realisations of the stress factors z_t , and co-dependence matrix \mathcal{C}_t represents the matrix of systemic risk weights either calibrated or estimated. The index in model (2) can be seen as a scaled matrix association index proposed by Mantel (1967). Such indexes are extensively used in different fields of science to model and test for the similarity of multidimensional data observed in matrix

⁵In this paper, we use the terms *extreme* and *extremeness* only in the meaning of *extremely high*.

⁶The following notation is used throughout the paper. The scalar $(A)_{i,j}$ represents the i, j th element of an $N \times N$ matrix A . I_N denotes the $N \times N$ identity matrix, J_N the $N \times N$ all-ones matrix and O_N the $N \times N$ all-zeros matrix; ι_N and 0_N are $N \times 1$ vectors of ones and zeros, respectively.

form.⁷ In our case, the index shows to which extent the realisations z_t are jointly high and co-dependent by associating two matrix functions \mathcal{E}_t and \mathcal{C}_t that quantify extremeness and co-dependence. In Section 5 we offer a bootstrap procedure to statistically test the null hypothesis of low or normal stress, that is the hypothesis that stress factors are not jointly high and strongly co-dependent at any point in time. Hence, our co-extremeness hypothesis represents the alternative hypothesis in our proposed testing scheme.

4. Operationalising systemic stress - the CISS

As one possible operationalisation of the general statistical framework summarised in Equation 2, we propose a modified variant of the ECB’s Composite Indicator of Systemic Stress (CISS) originally developed by one of the authors and collaborators in the working paper Hólló et al. (2012). The present paper introduces new daily CISS series for the United States and the euro area.⁸ As we will argue, the CISS possesses certain specific advantages which, in our view, distinguish it from alternative FSI designs. Appendix C demonstrates how other popular FSI designs can also be represented as special cases of our general framework.

4.1. Original stress indicators

The operationalisation starts with systematically selecting $N = 15$ stress indicators $x_{i,t}$, with $i = 1, \dots, N$ and $t = 1, \dots, T$, from the most important segments of the US and the euro area financial system. Taken together, these market segments should cover the main flows of financial funds channeled either indirectly through financial intermediaries or directly via short-term and long-term securities markets from ultimate lenders to final borrowers. These market segments include: i) equity markets for nonfinancial corporations; ii) equity markets for financial institutions (listed banks and other traded financial entities); iii) money markets (interbank, commercial paper and T-bill markets); iv) sovereign and corporate bond markets; and v) foreign exchange markets. Our choice of indicators is limited by the following considerations. First, to ensure representativeness, the original stress indicators should either be based on broader market indices or on assets with benchmark status for the pricing of a wider range of substitutes (e.g., on-the-run government bonds). Second, we require daily data (with a publication lag of one business day at most) to support the CISS as a quasi real-

⁷The scaled version of this measure can be seen as Moran’s I or Geary’s C statistics that are used extensively while modeling economic networks based on the geographical distance.

⁸Daily updates of the new daily CISS are available for a broader set of countries via the ECB’s Statistical Data Warehouse.

time financial stability monitoring tool. And third, stress indicators should carry long data histories to allow for meaningful historical benchmarking of stress events and to render the CISS potentially useful in macro-financial econometric time-series analysis. These limitations imply that the CISS is mainly based on fairly standard financial market data.

Stress is measured in several ways, with each indicator capturing certain observable stress symptoms and increasing in the level of stress in line with Assumption 3.1.⁹ All market segments include (at least) one measure of historical volatility, computed as an exponentially-weighted moving average (EWMA, or exponential smoother) of squared daily log returns or squared daily interest rate changes with smoothing parameter $\lambda = 0.85$; initial values are set to the return variance over the first two years of the respective data sample (Engle (2009), p. 30f.).¹⁰ Further stress measures include several interest rate differentials as risk spreads (e.g., Ted spread, corporate bond spreads) as well as book-price ratios and cumulative valuation losses (the CMAX by Patel and Sarkar (1998)) to capture strains in equity markets. A detailed description of all original stress indicators included in the US and the euro area can be found in Tables 2 and 3 in Appendix A).

4.2. Stress factors

Next, all original stress indicators $x_{i,t}$ are transformed into stress factors $z_{i,t}$ by applying the probability integral transformation (PIT), which involves estimation of the empirical cumulative distribution function (cdf):¹¹

$$z_{i,t} = \hat{F}(x_{i,t}) := \begin{cases} \frac{1}{T_0-1} \sum_{s=1}^{T_0-1} \mathcal{I}(x_{i,s} \leq x_{i,t}), & \text{for } t = 1, \dots, T_0 - 1 \\ \frac{1}{t} \sum_{s=1}^t \mathcal{I}(x_{i,s} \leq x_{i,t}), & \text{for } t = T_0, \dots, T, \end{cases} \quad (3)$$

with $\hat{F}(x_{i,t})$ defined as the empirical cdf of indicator $x_{i,t}$. Hence, each realisation of stress factor $z_{i,t}$ simply results from replacing the original observation $x_{i,t}$ with its estimated cdf value $\hat{F}(x_{i,t})$. Since $\hat{F}(x_{i,t})$ is non-decreasing, the PIT meets Assumption 3.2. One important feature of how we implement the PIT is its recursive computation over expanding data samples. The integer T_0 sets the start date of the recursion. In our application to US and euro area data, the recursion starts in January 1, 2002.¹² Recursive transformation immunises the composite indicator against look-ahead bias (Brownlees, Chabot, Ghysels,

⁹For a general discussion of individual stress indicators and their information content see Hólló et al. (2012) and the literature cited therein.

¹⁰The 2012 version of the CISS measures volatility as the average absolute daily return or interest rate change over the five business days of a week.

¹¹See, e.g., Casella and Berger (2002).

¹²This date is chosen to guarantee at least several years of data for the initialisation of a few euro area time series which only start in 1999 or shortly before.

and Kurz (2020)) and, related, against event reclassification, i.e. the risk of recalling a stress event indicated in the past (based on past data only) that fails to be identified as a stress event when also taking into account future information (Hóllo et al. (2012))).

There are several reasons why we prefer the PIT to z-score transformation for our purposes.¹³ 1. *Homogeneity*. – The PIT leads to indicators which are unit-free and—whatever their original distribution—(approximately) unconditionally standard uniform distributed: $z_{i,t} \sim U(0, 1)$. The stress factors resulting from the PIT are thus homogenised not only in terms of scale (i.e., $z_{i,t} \in (0, 1]$) but also in terms of distribution. The latter fact does not hold true for z-score standardisation. 2. *Robustness*. – The use of ranks in the PIT robustifies the stress factors and thus the composite indicator against outliers since it limits an outlier to the value of its relative rank (Stuart and Ord (1994)). This property is important in our context for two main related reasons. It firstly ensures that even though the PIT is computed recursively, the information content of the composite indicator is still robust over time, i.e., it does not depend on whether the composite indicator is computed over the full data sample or any meaningful subset of it (see Figure 13 in Appendix B). Consequentially, the CISS is also robust to selecting different starting dates T_0 for the recursive PIT (Hóllo et al. (2012)). And, secondly, this robustness property matters most when the composite indicator is most useful, namely during episodes of financial instability which may add many outliers to the expanding data samples of the original stress indicators. Figure 2 below demonstrates—in a constructed example based on the US Ted spread—that z-score standardisation can lead to transformed indicators whose information content becomes unstable over time when many outlying observations are added to the data sample (panel b). While recursive standardisation indicates several stress peaks prior to the GFC, most of these indicated events get almost completely “ironed out” when using the full-sample mean and standard deviation for the transformation. The picture appears much less dramatic in the case of the probability integral transformation (panel a). A third main advantage of the PIT—namely robustness of our measure of co-dependence—will be discussed in Subsection 4.4.

4.3. Extremeness

The vector of stress factors z_t is a natural measure of extremeness. Stress in any particular market segment is most extreme if the respective indicator $x_{i,t}$ reaches its historical maximum, in which case $z_{i,t} = 1$. However, we want to capture extremeness for the system as a whole. For this purpose, we quantify extremeness as the cross-product between all

¹³We are not the first using the PIT in FSI designs. Illing and Liu (2006) conduct a study comparing different indicator transformations (PIT and z-score), and they find that FSIs based on z-score transformed components perform better in regressions of crisis-event dummies on FSIs.

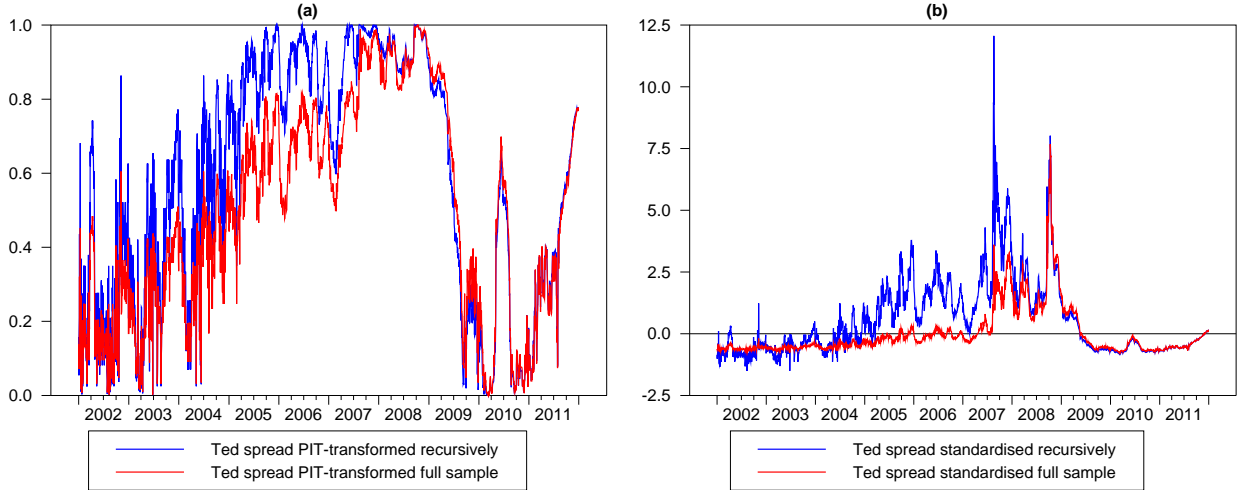


Fig. 2. The figure assumes that the Ted spread were available only from January 2002 until December 2011, i.e. for ten years of daily data. The subprime mortgage turmoil and the Global Financial Crisis occur in the second half of the sample period. In panel (b), the Ted spread is standardised (z-score) in two ways. First, standardisation is done recursively as from January 2006 (blue line). Second, it is standardised using full sample information at all points in time (red line). The figure demonstrates that in particular in small samples it can happen that the information content of z-score transformed indicators becomes unstable. Panel (a) suggests that this problem is less severe for probability integral transformed indicators.

non-centered stress factors, $z_{i,t} \cdot z_{j,t}$, and collect them in the symmetric matrix \mathcal{E}_t of order N :

$$\mathcal{E}_t = z_t z_t'. \quad (4)$$

Since for each matrix element it holds that $(\mathcal{E}_t)_{i,j} = (z_t z_t')_{i,j} \in (0, 1]$, the extremeness matrix becomes an all-ones matrix $J_N = \iota_N \iota_N'$ in case of maximum system stress and approaches a zero matrix $O_N = 0_N 0_N'$ in case of minimum stress.

4.4. Co-dependence

We compute the co-dependence nonparametrically as the conditional rank correlation (Spearman's ρ) between each of the $N(N - 1)/2$ pairs of stress factors. The rank correlation shall tell us whether and to which extent two stress factors are, by historical standards, similarly high or low at a certain point in time. For instance, if two stress factors jointly move towards the upper ranges of their empirical cdf, there is a higher risk that financial stress becomes more widespread and thus systemic, *ceteris paribus*. If they are moving asynchronously, that risk tends to be lower.

Since our stress factors inherit some of the autocorrelation and heteroskedasticity of the original data¹⁴, we estimate the rank correlation between two stress factors $z_{i,t}$ and $z_{j,t}$ as the conditional time series expectation $\rho_{i,j,t} := E_t[\rho_{i,j,t+1}|z_{i,t-k}, z_{j,t-k}, \rho_{i,j,0}]$ for $k = 0, 1, \dots, t-1$ and some initial value $\rho_{i,j,0}$. The conditional expectation, in turn, is modelled in a nonparametric way as an autoregressive exponentially-weighted moving average (EWMA) process drawing on the multivariate GARCH literature (Engle (2002)). The rank correlation coefficient is defined vis-à-vis a centrality measure. We take the population median 0.5 as the centrality measure and accordingly define the vector of centered stress factors as $\tilde{z}_t = (z_t - \iota_N \cdot 0.5)$. The EWMA-filter is implemented for the variance-covariance matrix H_t of the centered stress factors \tilde{z}_t :

$$H_t = \lambda H_{t-1} + (1 - \lambda) \tilde{z}_t \tilde{z}_t' \quad (5)$$

with a calibrated smoothing parameter (or decay factor) $\lambda = 0.85$. On the one hand, this value of 0.85 provides relatively smooth estimates of the rank correlation in normal times; on the other hand, it still accommodates abrupt and large shifts in correlation patterns as they typically occur during crisis periods. The elements $\rho_{i,j,t}$ of correlation matrix R_t are computed from the elements $h_{i,j,t}$ of H_t as $\rho_{i,j,t} = h_{i,j,t} / \sqrt{h_{i,i,t} h_{j,j,t}}$. The Spearman's ρ matrix R_t is our desired matrix of systemic risk weights:

$$\mathcal{C}_t = R_t. \quad (6)$$

This definition of Spearman's rank correlation is equivalent to the usual Pearson correlation coefficient between PIT-transformed data. As such, it is a nonparametric measure of co-dependence and consequently more robust to distributional assumptions and outliers (Engle (2009), p. 22). For instance, Spearman's rank correlation only relies on the existence of a monotonic relationship between the original stress indicators. In contrast, raw stress indicators with very heterogeneous statistical distributions may co-vary in nonlinear ways despite monotonicity. Linear correlation between the raw stress indicators or their z-scores may therefore fail to properly capture such nonlinear dependencies. Since we are working with the PIT and with rank correlations, our approach to measure co-dependence loosely borrows from the copula literature (Engle (2009)).

Figure 3 visualises realisations of matrices \mathcal{E}_t and \mathcal{C}_t for four days characterising stress conditions before and during the GFC and the Covid-19 crisis, respectively. Panel (a) represents the period ahead of the GFC in which volatilities and risk premia were persistently

¹⁴We do not prewhiten nor de-GARCH the raw stress indicators before applying the PIT to the "cleaned" residuals as it is the case, for instance, in the standard copula literature.

low by historical standards in basically all market segments and across the globe. This apparent “pricing for perfection” (Gieve (2006)) is reflected in the dominant dark blue colours on and below the main diagonal, indicating that stress factors were at or below the 30th or 10th percentile of their historical distribution on 31 January 2007. The dominant dark red colour above the main diagonal illustrates the commonality of very low stress conditions across market segments, indicating rank correlations in the range of 0.8 to 1.0.¹⁵ That benign picture at the beginning of 2007 soon morphed into the subprime mortgage crisis and ultimately into the GFC. At the height of the GFC, stress levels were at or slightly below their historical maxima across the board, indicated by panel (b) showing dark red colour almost everywhere for 10 October 2008. Panels (c) and (d) emphasise how quickly financial stability deteriorated when the US was struck by the Covid-19 shock. On 28 February 2020, financial stress conditions were mixed across markets and indicators and overall very benign. Accordingly, panel (c) covers more or less the full colour spectrum. Within just one month, the situation changed completely, with financial stress rising sharply in all market segments. The systemic dimension of the increased financial stress caused by the coronavirus manifests itself in the similarity of panels (b) and (d). Like at the height of the GFC, the contour plot in panel (d) is dominated by the dark red colour; the blue and yellow “cross” reflects the fact that only the price-book ratio of nonfinancial corporations did not jump towards historically high levels by 31 March 2020.

4.5. Composite Indicator

Equipped with all ingredients, the CISS can now be easily computed as the matrix association index \mathcal{S}_t as defined in Equation (2) based on the quantified inputs from Equations (4) and (6):

$$\text{CISS}_t = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N (z_t z'_t)_{i,j} (R_t)_{i,j}, \quad (7)$$

with $0 < \text{CISS}_t \leq 1$. As mentioned, the scaling parameter $1/N^2$ simply represents equal portfolio weighting $w_i = 1/N$ of stress factors $z_{i,t}$. Since R_t has only ones on its main diagonal, each (squared) stress factor enters this scaled sum of element-by-element products with a unit systemic risk weight, while the risk weights of the cross-products (the upper and lower diagonal elements of R_t) are typically less than 1 and approach 1 only if stress is persistently high or low.

This formulation of the composite indicator problem is not only statistically appealing but also economically intuitive. The formula is well-known in finance, describing how to

¹⁵See the notes section of Figure 3 on how to infer correlation levels from the colour bars.

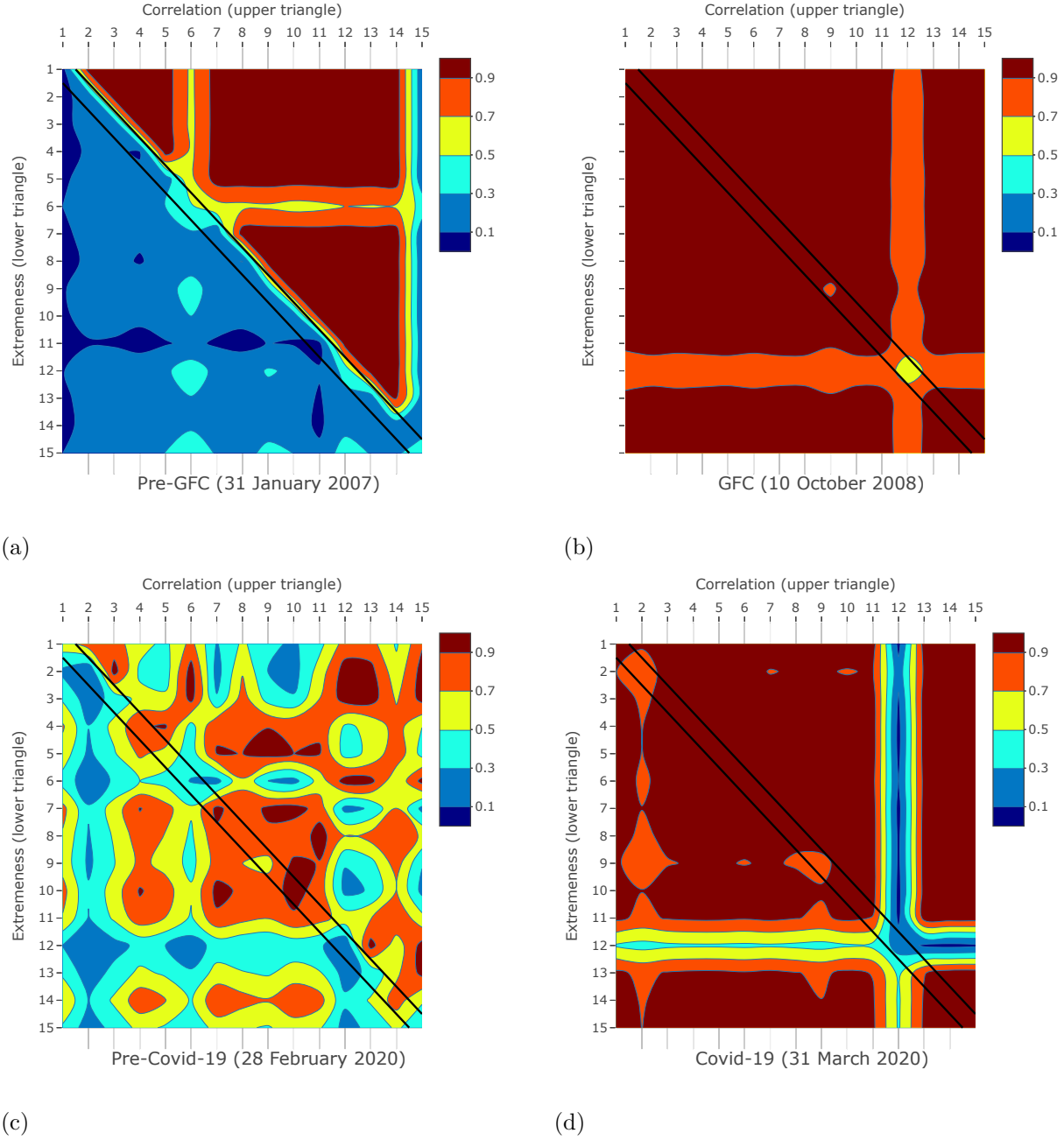


Fig. 3. The figure shows contour plots visualising realisations of the extremeness and the correlation matrices on four days representing financial stability conditions before and during the Global Financial Crisis (GFC) and the Covid-19 crisis. Since both matrices are symmetric, extremeness is shown in the triangle below and correlation in the triangle above the main diagonal in each of the four panels. The main diagonal is highlighted by the two parallel black solid lines and shows the levels of the 15 stress factors $z_{i,t}$ for US data. The order of the stress factors follows the numbering in Table 2. Extremeness is plotted as the square root of the cross-products $\sqrt{z_{i,t}z_{j,t}}$ (the geometric average) conformable to the level scale of the stress factors. For graphical purposes, the plotted correlation coefficients $\tilde{\rho}_{i,j,t}$ are rescaled to fit into the unit interval according to the transformation $\tilde{\rho}_{i,j,t} = 0.5(\rho_{i,j,t} + 1)$.

compute the return variance (risk) of an asset portfolio based on the return variances and covariances (or standard deviations and correlations) of a set of assets with equal portfolio shares (see, e.g., the seminal paper on portfolio selection by Markowitz (1952)). Appendix C illustrates this point by computing the variance of the financial system as a whole based on a subset of the CISS components.

The CISS formula (7) can be equivalently expressed as a quadratic form:

$$\text{CISS}_t = (w \circ z_t)' R_t (w \circ z_t), \quad (8)$$

with \circ denoting element-by-element multiplication. Figure 4 shows the CISS for the US and the euro area since January 1973 and January 1980, respectively. Most pronounced spikes in the two indicators occur simultaneously and can be associated with well-known global stress events such as the stock market crash in 1987, the LTCM collapse in 1998, the 9/11 terrorist attacks in 2001, the subprime mortgage turmoil in 2007, the Lehman Brothers default in 2008, and the recent Covid-19 crisis. However, there are also a few episodes reflecting more localised shocks like the Fed monetary policy tightening regime under Chairman Paul Volcker from 1979 to 1982 and the associated Savings and Loans crisis in the early 1980s, the European Exchange Rate Mechanism (ERM) crisis in 1992 and the euro area sovereign debt crisis in 2011 and 2012.

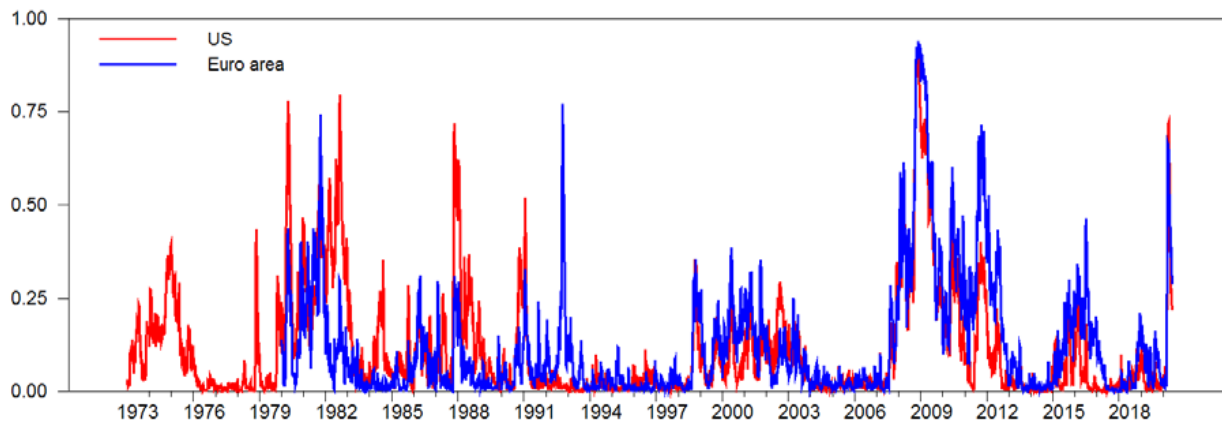


Fig. 4. This figure plots daily data of the US and the euro area CISS until 31 May 2020. The US data start on 3 January 1973 and the euro area data on 1 January 1980.

4.6. A Decomposition

The following simple decomposition of the CISS helps better understand its rationale. The decomposition starts from the assumption that all stress factors were perfectly correlated at all times. In that case, the correlation matrix R_t becomes an all-ones matrix J_N and thus redundant, simplifying the CISS formula to the following expression:

$$\begin{aligned} \text{CISS}_t &= (w \circ z_t)' J_N (w \circ z_t) = (w \circ z_t)' (w \circ z_t) \\ &= \left(\sum_{i=1}^N w_i \cdot z_{i,t} \right)^2 = \bar{z}_t^2. \end{aligned} \quad (9)$$

Hence, in case of perfect correlation, the square of the equally-weighted average of stress factors \bar{z}_t^2 emerges as the upper limit of the CISS. From this follows that in a systemic crisis, when stress is high across the board, the CISS converges towards \bar{z}_t^2 . On the contrary, when cross-correlations are generally weaker, the CISS deviates more strongly from the simple-average FSI. We accordingly call the difference between the CISS and \bar{z}_t^2 the *correlation discount* and use it for the following decomposition of the CISS:

$$\text{CISS}_t = \sum_{i=1}^N \frac{\bar{z}_t}{N} z_{i,t} - \left(\frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N z_{i,t} z_{j,t} (1 - \rho_{i,j,t}) \right). \quad (10)$$

The first term of Equation (10) is simply \bar{z}_t^2 split into N stress factor contributions, and the second term is the correlation discount. Figure 5 shows the decomposition of \bar{z}_t^2 where the stress factors are aggregated into five market segments.¹⁶ The figure confirms that the CISS and the stacked contributions almost coincide during the GFC such that the correlation discount approaches zero. More importantly, the pattern of the correlation discount—the grey shaded area in Figure 5—clearly demonstrates the main advantage of the CISS: it helps better identify episodes of truly widespread financial stress (systemic crises) by giving lower weight to situations in which elevated stress remains a more local market-specific event. The dot-com boom-bust period in the late 1990s and early 2000s serves as a prime example of a largely non-systemic stress episode.

5. Statistical analysis

In Definition 2 we motivated a systemic financial stress index \mathcal{S}_t as a quantification of the hypothesis that stress factors are jointly high (co-extremeness). This section proposes

¹⁶Such a decomposition chart for the euro area weekly CISS of Hólló et al. (2012) is regularly shown in the ESRB Risk Dashboard (<https://www.esrb.europa.eu/pub/rd/html/index.en.html>).

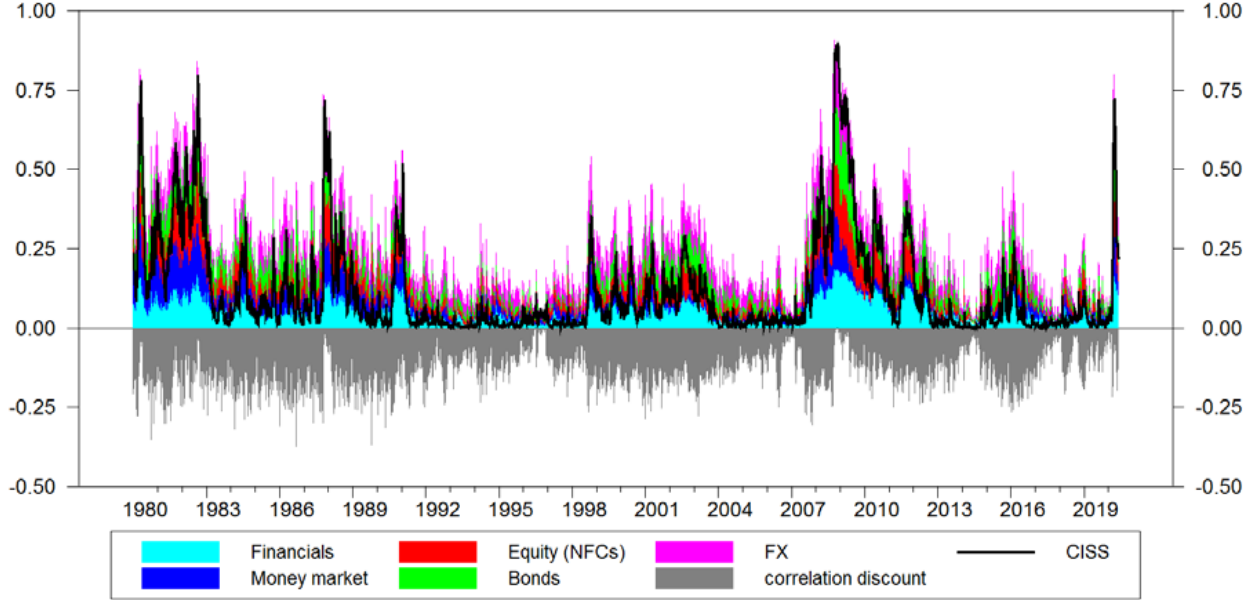


Fig. 5. The figure plots the decomposition of the US CISS according to Equation (10). The contributions from the individual stress factors are aggregated into money, bond, equity (financial and nonfinancial corporations) and foreign exchange market contributions in line with Table 2. The data sample is daily from 2 January 1980 to 31 May 2020.

a bootstrap algorithm to derive test statistics to assess this hypothesis. Let realisations of stress index $\{S_t : t = 1, \dots, T\}$ be computed as the CISS according to Equation 7. Our aim is to make statistical statements as to whether any particular realisation S_t can be regarded as unusually high or low (“normal”) instead. For this purpose we formulate the following null hypothesis for any realisation S_t :

$$H_0 : S_t \in \mathcal{Y}^{low} \quad (11)$$

with $\mathcal{Y}^{low} = \{\tilde{S}_t : t = \tilde{T}_0, \dots, \tilde{T}_1\}$ denoting a subsample of the data representing a period with low stress, and \tilde{T}_0 and \tilde{T}_1 being the start and end date of this period, respectively. If $\mathcal{Y}^{full} = \{S_t : t = 1, \dots, T\}$ denotes the full data sample, it holds that $\mathcal{Y}^{low} \subset \mathcal{Y}^{full}$ and $\tilde{T}_1 - \tilde{T}_0 < T - 1$. In order to test the null hypothesis, we derive a critical stress value $\gamma_{1-\alpha}$ at some statistical significance level α , implicitly defined in

$$\Pr\left(\tilde{S}_t > \gamma_{1-\alpha} | \mathcal{Y}^{low}\right) \leq \alpha. \quad (12)$$

If the data-generating process of \tilde{S}_t were known, Equation (12) could be solved analytically for $\gamma_{1-\alpha}$. Since we do not assume any particular data-generating process, we need to estimate

the critical value nonparametrically by simulation. We propose a bootstrap-based sampling procedure that starts with approximating the empirical distribution of S_t under no-stress conditions, i.e. under the null hypothesis, and then calculates a critical value for the null hypothesis as the $1 - \alpha$ percentile of this approximate distribution.

The idea of our bootstrap algorithm is similar to the random permutation approach of Mantel (1967) which provides a test for the similarity of two (distance) matrices. Since the CISS summarises the information contained in the matrix of cross-products $(z_t z_t')_{i,j} (R_t)_{i,j}$, any statistical comparison of a realised CISS with a benchmark value (e.g., a critical value) for a low-stress CISS may be interpreted as a test for the similarity of the two underlying cross-product matrices. However, our procedure benefits from the assumption that some observations of S_t stem from a low-stress environment and are thus generated under the null hypothesis. We judge the subperiod from 1 January 1992 (\tilde{T}_0) to 29 December 2006 (\tilde{T}_1) to be one representing normal stress conditions in the US financial system.

We start from a simple linear relationship between each cross-product and the CISS that defines the set of $N(N - 1)/2$ different “residuals” $\psi_{i,j,t}$ to draw from:

$$(z_t z_t')_{i,j} (R_t)_{i,j} = \tilde{S}_t + \psi_{i,j,t}, \quad i, j = 1, \dots, N, \quad t = \tilde{T}_0, \dots, \tilde{T}_1, \quad (13)$$

where $\psi_{i,j,t}$ has zero mean by construction. We now assume that in a low-stress environment—with extremeness and correlations being rather low on average—the $\psi_{i,j,t}$ are expected to be evenly allocated. We formalise this expectation in the following assumption:

Assumption 5.1. *Suppose there exists an array of random variables $\{\psi_{i,j,t}, 1 \leq i, j \leq N\}$ which are i.i.d. with zero mean, finite variance and satisfy $|\psi_{i,j,t}|^3 < \infty$.*

If this assumption holds, we can randomly draw from this array of random variables, and the following bootstrap algorithm can be used to construct the empirical distribution of \tilde{S}_t (Wasserman (2003), Chapter 8) and to compute any critical value $\gamma_{1-\alpha}$:¹⁷

Algorithm 1. *Let $t = \tilde{T}_0, \dots, \tilde{T}_1$ and perform the following steps:*

- (i) *Let $\tilde{S}_t = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N (z_t z_t')_{i,j} (R_t)_{i,j}$, and define the auxiliary residuals $\psi_{i,j,t} = (z_t z_t')_{i,j} (R_t)_{i,j} - \tilde{S}_t$;*
- (ii) *From the array of random variables $\{\psi_{i,j,t}, 1 \leq i, j \leq N\}$, draw randomly i and j , $n = 1, \dots, N_b$ times, and construct N_b simulated values of array $\{\psi_{i,j,t}^{(n)}, 1 \leq i, j \leq N\}$;*
- (iii) *Compute N_b simulated values of each cross-product term as $\left((z_t z_t')_{i,j} (R_t)_{i,j} \right)^{(n)} = \tilde{S}_t + \psi_{i,j,t}^{(n)}$;*

¹⁷By a loose interpretation, this would mean there is only a negligible difference between realizations $\psi_{i,j,t}$ across $i, j = 1, \dots, N$.

(iv) Compute N_b simulated values of the stress index $\tilde{S}_t^{(n)} = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \left((z_t z_t')_{i,j} (R_t)_{i,j} \right)^{(n)}$ and define

$$\gamma_{1-\alpha} = \frac{1}{\tilde{T}_1 - \tilde{T}_0 + 1} \sum_{t=\tilde{T}_0}^{\tilde{T}_1} Q_{1-\alpha} \left(\tilde{S}_t^{(1)}, \dots, \tilde{S}_t^{(N_b)} \right),$$

where $Q_{1-\alpha}(\cdot)$ is an empirical quantile function that delivers the required critical value.

The algorithm is applied to daily US data based on the entire low-stress subsample, and the critical value is computed for a significance level of $\alpha = 0.01$ (1%). Figure 6 plots the US CISS together with a straight line representing the estimated critical value of 0.123.¹⁸

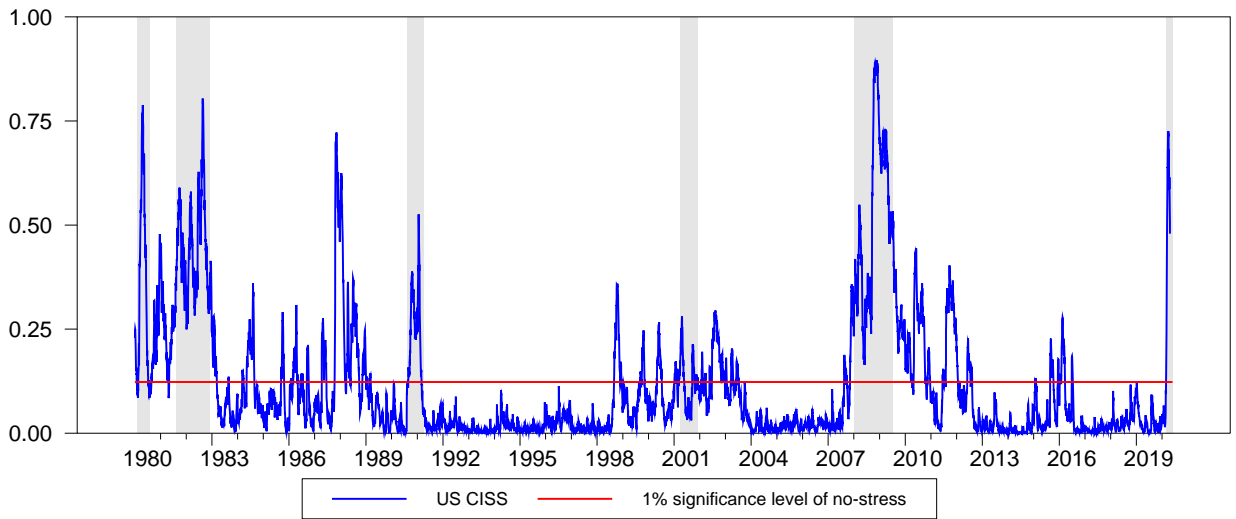


Fig. 6. US CISS and high-stress threshold. The threshold value (0.123) represents the critical value of the null hypothesis of low or “non-systemic” stress at the 1% significance level. The threshold is generated using sampling algorithm 1. Shaded area represents NBER based recession indicator for the United States from the period following the peak through the trough. Data is daily from 2 January 1980 to 31 May 2020. Sources: Federal Reserve Bank of St. Louis for the NBER recession indicator.

A simple illustration may provide some tentative evidence as to the potential economic relevance of a thus measured threshold. We use the 1%-critical value to split the sample into low-stress and high-stress CISS observations and plot, for both regimes separately, the US CISS against annual real GDP growth led by two months (Figure 7). Red dots indicate joint observations from the low-stress regime, and blue dots from the high-stress state. The figure clearly suggests a negative relationship between the CISS and growth in the high-stress regime, while the red dots appear like a random cloud. This potential nonlinear pattern is in

¹⁸The critical values for the 5% and 10% significance levels are 0.115 and 0.111, respectively.

line with Hólo et al. (2012) who estimate a CISS threshold based on a bivariate threshold-VAR for the euro area CISS and annual growth in industrial production. That paper indeed finds statistical support for a single threshold, and it turns out that shocks in the CISS have strong adverse effects on economic activity only in the high-stress regime. The next section explores the real effects of systemic stress in more detail, including potential asymmetries in states of elevated financial stress.

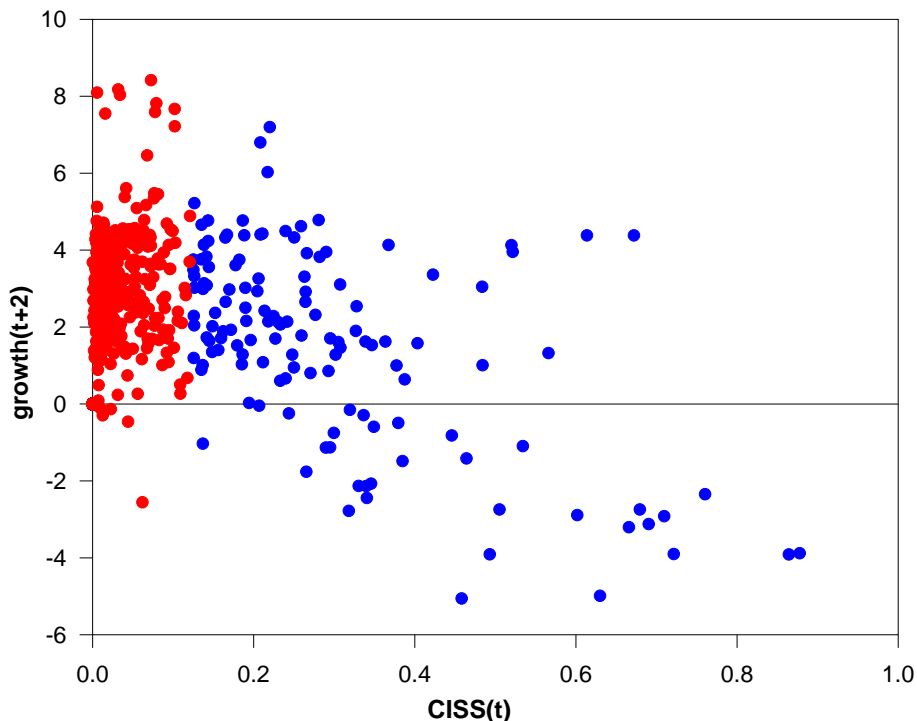


Fig. 7. Scatter plot of CISS and annual real GDP growth (in %, led by two months) in the US. CISS is monthly average of daily data, and real GDP is interpolated to monthly from quarterly data as described in Appendix D. Blue dots represent periods when the CISS is larger than the 1% critical value from sample algorithm 1 (high stress), and red dots symbolise periods in which the CISS is lower than that threshold (low stress). Data is monthly from January 1980 to March 2020.

6. Systemic financial stress and the macroeconomy

This section investigates the real effects of systemic financial stress measured by the CISS. For this purpose, we run linear and quantile VARs based on euro area data.¹⁹ We use these models to quantify the relative importance of financial stress as a driver of economic activity

¹⁹VARs with US data deliver results rather similar to those reported in this section.

on average and over two crisis episodes, the GFC in 2008/09 and the ongoing (at the time of writing) Covid-19 crisis. The quantile VARs helps identifying potential asymmetries in the underlying macro-financial linkages.

6.1. Linear VAR

We make our points based on parsimoniously specified VAR models. The models include three variables at monthly data frequency: annual differences in log real GDP, the level of the PMI Composite Output Index and the CISS (data are described in more detail in Appendix D). Annual real GDP growth is a measure of the cyclical component of real GDP (see Hamilton (2018)). Its time series properties square well with those of the PMI and the CISS which seem to follow stationary though rather persistent stochastic processes. The PMI tracks current business trends by collecting information on sales, new orders, employment, inventories and prices from over 5,000 companies from the manufacturing and service sectors. The PMI is timely available, with data for a given month published in that very month as a flash estimate and at the beginning of the subsequent month as final. Against this background, we take the PMI as an almost real-time business activity indicator that should help us identifying macroeconomic activity shocks and disentangle them from financial shocks as captured by the CISS. Figure 16 in Appendix D visualises the close association between the PMI and annual real GDP growth, an association which is strongly confirmed by the VAR estimates.²⁰ The estimation runs from October 1998 to February 2020. The lag order of the VAR is set to two according to standard information criteria.

The VAR estimation reveals rich interdependencies between the model variables. Standard exclusion F-tests—i.e., tests whether all the lags of one variable can be restricted to zero in the equation of another variable—corroborate the CISS as a highly significant *direct* predictor of economic activity in both the PMI and the GDP growth equations (with p -values of 0.0009 and 0.0091, respectively. Likewise, the PMI also predicts real GDP growth with very high statistical significance. Real GDP, by contrast, mainly receives predictive power from the other two variables but does not predict any of them at conventional confidence levels. This general pattern carries over to the impulse response functions (IRFs) which take into account also *indirect* effects²¹ and contemporaneous relationships between the model variables resulting from the structural shock identification. We identify structural shocks by applying the standard triangular Cholesky decomposition to the residuals variance-covariance matrix,

²⁰The macroeconomic predictive power of the PMI is found to be strong and robust for the cases of the US and the euro area in Jarocinski and Mackowiak (2017)

²¹ Jarocinski and Mackowiak (2017) provide formal tests of Granger causal priority in a Bayesian VAR setting. Granger causal priority tests whether one variable helps predict another variable by also taking into account indirect effects via third variables.

allowing for contemporaneous impacts of the CISS and the PMI on real GDP growth, and of the CISS on the PMI, respectively. The corresponding variable ordering CISS-PMI-GDP can be motivated as follows. First, the standard macro VAR literature would order the CISS after real GDP growth, assuming that slow-moving variables like output can only react with a lag to fast-moving asset prices.²² However, from an information perspective, such ordering would imply financial market participants to have perfect foresight since quarterly real GDP is published with a delay of at least one month. In addition, it may appear plausible to assume that financial stress shocks originate mainly from within the financial sector particularly during crisis times, and that producers or consumers respond to large uncertainty shocks by quickly adopting a wait-and-see attitude which, in turn, may give rise to an immediate reduction in economic activity (Bloom (2009)). And finally, this ordering is consistent with the pattern of exclusion F-tests found. Second, putting the CISS in front of the PMI is somewhat arbitrary. Both variables can be assumed to quickly adjust to incoming relevant information. But none of the reported results changes qualitatively with a different variable ordering in the Cholesky decomposition. We refer to some results from a reverse ordering of the CISS and the PMI in the main text and in Appendix E.

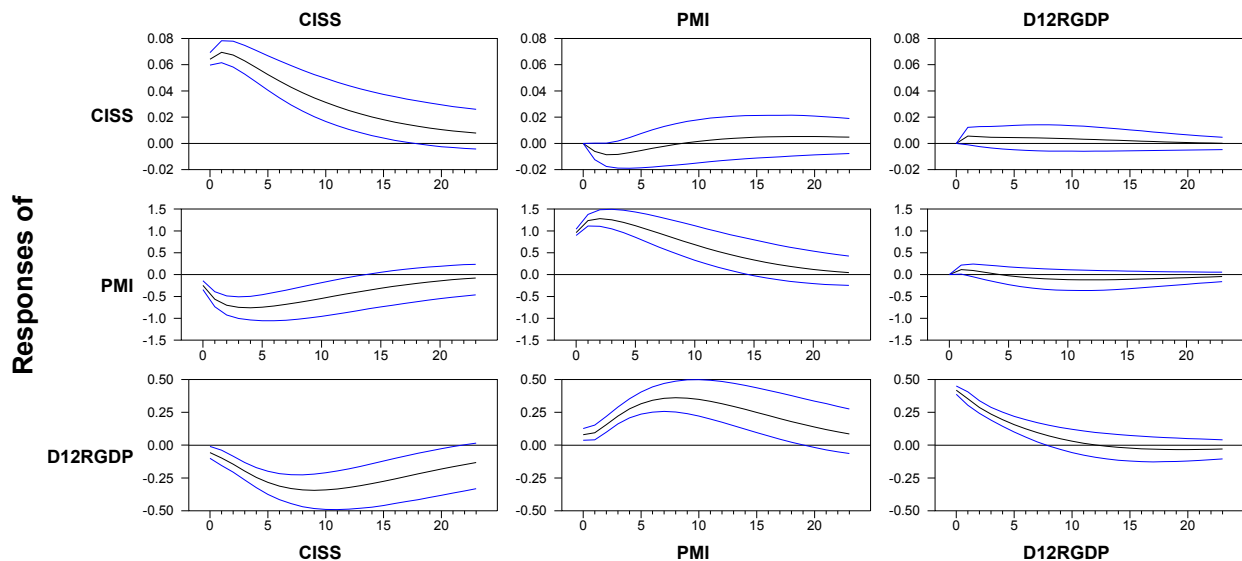


Fig. 8. The figure plots mean impulse response functions (black line) from the linear VAR for the case of structural identification by Cholesky decomposition with variables ordered CISS-PMI-GDP. The blue lines represent 90% error bands estimated via Monte Carlo integration. Forecast steps are monthly.

²²See Christiano, Eichenbaum, and Evans (1999) for an overview of different identification approaches and their economic rationale.

The impulse response functions (IRFs) appear consistent with our assumption of a stationary VAR, although the dynamic responses of the CISS and the PMI to own shocks suggest a rather sluggish mean reversion in both cases (see all diagonal elements of the IRFs shown in Figure 8). The first column of the IRFs plotted in Figure 8 shows the mean responses to a one-standard-deviation (0.063) shock in the CISS together with 90% error bands. Both the PMI and real GDP growth display swift, strong and statistically significant negative reactions to financial stress shocks. It takes a little longer for GDP growth to reach its peak effect than it takes for the PMI. Almost identical reaction patterns result from the alternative shock identification scheme, strengthening the case of the CISS as a decisive force behind the dynamics of our three-dimensional system (see Figure 17 in Appendix E). PMI shocks (one-standard-deviation shocks are 0.98 for plain residuals and 0.95 for structural shocks) exert strong and statistically significant influences on real GDP growth, confirming that this sentiment indicator contains important advance information for actual and subsequent real GDP growth. The reactions of the CISS to PMI shocks is the only case where the IRFs from the alternative Cholesky orderings look materially different. If the CISS precedes the PMI, there is no statistically significant reaction of the former to orthogonalised shocks in the latter; if the ordering is reversed, we obtain significantly negative reactions of the CISS to PMI shocks (compare the respective elements in the IRFs shown in Figures 8 and 17). That this IRF depends so strongly on the assumed direction of contemporaneous causality is consistent with the exclusion F-test failing to reject the null of no direct lagged impacts of the PMI on the CISS at conventional confidence levels. Finally, structural shocks in real GDP growth (0.41%, compared with 0.42% for non-orthogonalised shocks) do not generate significant adjustments neither in the CISS nor in the PMI.

Information about the relative strength of the in-sample explanatory power can be read off forecast error variance decompositions (FEVDs). Figure 9 summarises the most relevant information from this standard VAR tool for our purposes (all FEVDs for both identification schemes are shown in Table 5 in Appendix E). The blue line indicates that structural CISS shocks from our benchmark model explain on average about 22% of the forecast error variance of real GDP growth over a 6-month horizon, and this contribution gradually increases to 36% over a one-year horizon and 42% over a two-year horizon. The respective values for the reverse ordering between the CISS and the PMI are 10%, 18% and 22% (these are points on the red line of Figure 9). While the “true” contribution by CISS shocks may lie somewhere between the blue and the red lines, any number within that range is large by common standards. Furthermore, the joint contribution from CISS and PMI shocks—which is independent from the structural ordering of these two variables—explains 78% and 85% of the forecast error variance of GDP growth shocks over one-year and two-year horizons, respectively. This

leaves a mere 22% and 15% contribution of own GDP shocks over the two forecast horizons.

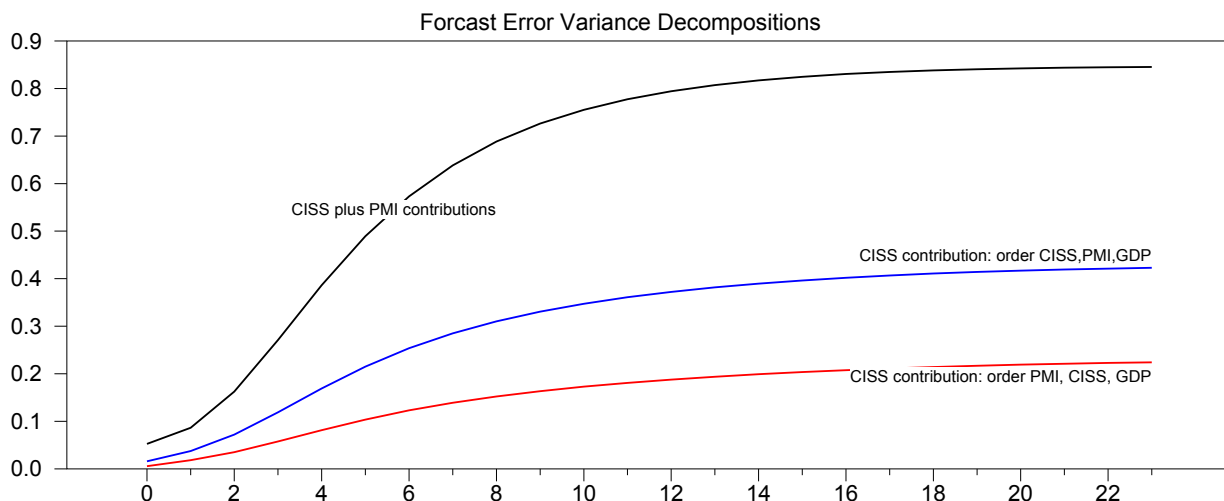


Fig. 9. The figure plots the individual and joint contributions from CISS and PMI shocks to the forecast error variance decomposition of real GDP growth from the linear VAR over monthly forecast horizons for two different variable orderings.

Simulation exercises further illustrate the strong macro-financial linkages captured by the model. We run three types of simulations over the GFC and the initial months of the Covid-19 crisis. Based on the full-sample coefficient estimates of our benchmark VAR, we compute conditional projections one to 24 months ahead starting in September 2008 (for the GFC) and in March 2020 (the outbreak of the Covid-19 crisis in the western world), respectively. We feed in three different sets of structural shocks: first, only CISS shocks (i.e., all other structural shocks set to zero); second, only PMI shocks; and third, CISS and PMI shocks jointly.

The GFC simulations assume that the system gets exposed to the paths of realised shocks from September 2008 to August 2010. This simulation exercise is therefore equivalent to a usual historical decomposition which decomposes the historical values of a VAR variable into an unconditional base projection and the accumulated effects of current and past shocks.²³ In Figure 10, the left vertical line highlights August 2008 as the forecast origin of the GFC projections, and the blue solid line shows the base projection assuming no disturbances during the forecast horizon. The red solid line depicts the scenario in which the system is hit by the sequence of structural CISS shocks as they actually occurred according to the model. In September and October, the crisis started off with two very large shocks of size +0.18

²³In contrast to a FEVD, the historical decomposition does not partition the deviation of a time series from its base projection since the historical decomposition can also produce negative shock contributions from some shocks in certain periods of time (the contributions only have to add up to the total deviation).

and +0.33 (three and five times the standard deviation of structural CISS shocks). These shocks lifted the CISS to its all-time highs given that systemic stress had already been very high when the Lehman Brothers shock arrived. Cumulated shocks further increased to 0.70 by March 2009, and at the end of the forecast horizon they stood at 0.81 reflecting the onset of the euro area sovereign debt crisis. Real GDP started falling very rapidly on impact, and by April 2009 it recorded with almost -6% its strongest rate of decline ever. CISS shocks explained 3.2 percentage points of this drop, that is more than half of it. Economic sentiment as captured by PMI shocks added another 1.2 percentage points to the maximum fall in GDP (the solid purple line in Figure 10). CISS and PMI shocks jointly thus explained three quarters of the record GDP slump (the solid turquoise line). Throughout the period with negative economic growth, the joint explanatory power of both shocks was similarly high or even higher.²⁴

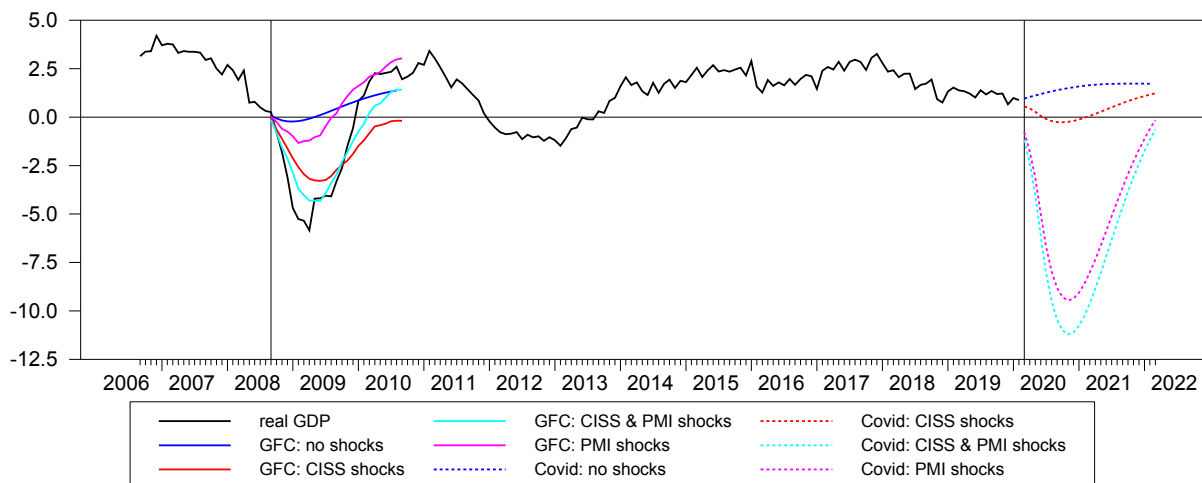


Fig. 10. The figure plots different simulated paths of real GDP growth over 2-year horizons. Solid coloured lines refer to the global financial crisis (GFC) with forecast origin Aug. 2008, and dashed coloured lines to the Covid-19 crisis with forecast origin Feb. 2020. The blue lines refer to the base projection with no shocks. The remaining coloured lines represent scenarios triggered by shocks in the variables indicated in the series key. For the GFC, the shocks are identical to the structural residuals observed over the simulation period. For the Covid-19 crisis, only the structural CISS and PMI shocks for Mar. and Apr. 2020 are fed into the simulation.

For the Covid-19 crisis, we perform a quasi real-time out-of-sample conditional forecasting exercise. The VAR estimation runs until February 2020 to which we also set the forecast origin. While monthly data of the CISS and the PMI are available up to and including

²⁴When ordering the PMI before the CISS, the picture still looks rather similar, with CISS and PMI shocks contributing more equally to the “meltdown” of the euro area economy.

Table 1: Covid-19 related CISS and PMI shocks

Month	model residuals		ordering CPG		ordering PCG	
	CISS	PMI	CISS	PMI	CISS	PMI
March 2020	+0.480	-22.044	+0.480	-20.228	+0.133	-22.044
April 2020	-0.152	-8.686	-0.152	-9.263	-0.289	-8.686

Notes: The table shows estimated VAR residuals and structural shocks for the CISS and the PMI from two different variable orderings in the Cholesky decomposition: CPG stands for the order CISS-PMI-GDP, and PCG for PMI-CISS-GDP.

April 2020, the last original final GDP data is from the fourth quarter of 2019; however, since the quarterly real GDP series is interpolated into a monthly series with the help of industrial production, we also obtain data points for January and February 2020. By means of iterative VAR projections for March and April, we back out model residuals for the CISS and the PMI and compute the respective structural shocks based on the Cholesky factor from the benchmark VAR estimate. For each of the three simulation cases, we then compute conditional VAR projections for real GDP growth over the period March 2020 to February 2022 (i.e., one- to 24-month ahead) feeding in the structural shocks in the CISS and/or the PMI for March and April 2020.

In our preferred identification setup, the outbreak of the Covid-19 crisis triggered a structural shock in the CISS of 0.48, the largest in history. However, already in April this huge shock was partly compensated by a mitigating shock of -0.15 (on the size of Covid-19 related shocks, see Table 1). This reversal may go back to the financial system—in particular banks—having become more resilient thanks to the regulatory and institutional overhaul since the GFC, and forceful policy interventions—by fiscal, monetary and supervisory authorities at national and union-wide levels—geared towards reassuring the financial system (Giese and Haldane (2020)). As a result, these two CISS shocks related to the outbreak of the pandemic would forecast, on its own, a strong but still limited and short-lived contraction in economic activity over a two-year horizon (red dotted line in Figure 10), implying about 2 percentage points lower growth than the base projection (the blue dotted line).

When adding the Covid-19 related PMI shocks, the picture changes dramatically. The structural PMI shock in March 2020 is truly unprecedented and represents an extreme outlier. If we were to assume normally distributed shocks, its ex ante probability of occurrence is basically zero since its value of -20.2 is, in absolute terms, more than 20 times larger than the standard deviation of orthogonalised PMI shocks. According to that yardstick, the subsequent PMI shock of -9.1 in April also counts as extreme. Based on only these two shocks, the VAR predicts a sharp and quick drop in real GDP that even eclipses the

dimensions of the GFC (see the dotted purple line in Figure 10). By November 2020, real GDP is forecast to drop by 9.5% year-on-year. Joining PMI with CISS shocks makes things worse, with real GDP now predicted to decline by about 11% annually in the last quarter of 2020 (the turquoise dotted line). According to the simulation, without expansionary shocks (originating, e.g., from abroad or from policy stimulus), year-on-year real GDP growth is bound to stay negative throughout the entire two-year forecast horizon.²⁵

All in all, the simulations support the narrative that the GFC represents a true systemic financial crisis that drove the euro area (and the world) economy into a deep recession, and that such devastating financial strains can be well captured by the CISS. On the other hand, the model also supports the claim that the Covid-19 crisis is mainly a real phenomenon, representing a huge shock to economic activity which affected the stability of the financial system only temporarily. While it is likely that the initial financial strains somewhat propagated the huge macro shocks, it seems as if policymakers successfully stepped in with resolute large-scale support measures that helped the financial system to maintain its crucial intermediation function in highly critical and uncertain times, thereby avoiding a vicious macro-financial cycle to (re-)appear (Altavilla, Barbiero, Boucinha, and Burlon (2020)).

6.2. *Quantile VAR*

We now investigate whether asymmetries exist in the macro-financial linkages present in our VAR setup. The recent more closely related literature regularly finds strong impacts of financial stress or financial conditions indicators on the left tail of the distribution of output growth (i.e., in recessions), while the impact of such financial variables is found to be much weaker in the central and upper parts of the growth distribution (i.e., in economic expansions). Such asymmetric responses of economic activity to financial distress may reflect several mechanisms put forward in the theoretical macro-finance literature, such as occasionally binding credit constraints in combination with fire sales amplifying the original shock (e.g, Bianchi (2011), He and Krishnamurthy (2012), Lorenzoni (2008) and Mendoza (2010)).

For this purpose, we reestimate our VAR model with quantile regression techniques. Specifically, we estimate a fully-fledged recursively identified Quantile VAR (QVAR) model as introduced in Chavleishvili and Manganelli (2019). In contrast to the more common

²⁵Reversing the variable ordering in the structural identification implies that two thirds of the large non-orthogonalised CISS shock in March could be attributed to the massive PMI shock in that months (Table 1), resulting in a mere +0.15 structural CISS shock. Given the second huge PMI shock in April, the corresponding orthogonalised CISS innovation for that month would amount to -0.29 . As a consequence, idiosyncratic CISS shocks would provide a modest economic stimulus over the projection horizon. However, we entertain some doubt that the OLS estimate of the Cholesky factor offers a realistic description of the contemporaneous causality running from PMI to CISS innovations in the far tails of the joint distribution.

single-equation predictive quantile regression frameworks applied in Giglio et al. (2016) and Adrian et al. (2019), the QVAR model allows for exploring all potential asymmetries in the direct and indirect relationships between the model variables, including feedback effects from the real to the financial sector. The basis structure of our recursive QVAR with three variables and two lags can be sketched as follows:²⁶

$$y_t = \phi^\theta + \sum_{l=1}^2 \Phi_l^\theta y_{t-l} + \Gamma^\theta y_t + \varepsilon_t^\theta, \quad (14)$$

with $y_t \in \mathbb{R}^3$ a vector collecting the endogenous variables in the order CISS, PMI and annual growth in real GDP; ϕ^θ denoting the vector of intercepts for a particular quantile $\theta \in (0, 1)$; Φ_l^θ the two quantile-specific 3×3 matrices containing the lag coefficients; and Γ^θ a lower triangular matrix with zeros on the main diagonal and the identified contemporaneous coefficients as the non-zero elements, again conditional on a particular quantile. Let $\mathcal{F}_{1,t} = (y'_{t-1}, y'_{t-2})'$ and $\mathcal{F}_{i,t} = (\mathcal{F}'_{i-1,t}, y_{i-1,t})'$ for $i = 2, 3$ be the recursive information set, then the conditional quantile functions can be used to explain the relationships among the variables across the entire conditional distribution and can be obtained by solving the following conditional quantile restrictions:

$$\Pr(y_{i,t} < Q_{y_{i,t}}(\theta | \mathcal{F}_{i,t}) | \mathcal{F}_{i,t}) = \theta, \quad i = 1, 2, 3, \quad (15)$$

with

$$Q_{y_t} = \phi^\theta + \sum_{l=1}^2 \Phi_l^\theta y_{t-l} + \Gamma^\theta y_t,$$

denoting the conditional quantile functions identified through restrictions (15).

The regression is equivalent to solving the numerical problem in Koenker and Bassett (1978):

$$\hat{\beta}_i(\theta) = \arg \min_{\beta \in \mathbb{R}^d} \sum_{t=3}^T \rho_\theta(y_{i,t} - Q_{y_{i,t}}(\theta | \mathcal{F}_{i,t})),$$

where $\beta_i(\theta) = ((\phi^\theta)_i, (\Phi_1^\theta)_{i,\cdot}, \dots, (\Phi_p^\theta)_{i,\cdot}, (\Gamma^\theta)_{i,\cdot})'$ a vector containing the d regression parameters in the equation for variable i , and $\rho_\theta(u) = u(\theta - I(u < 0))$ is an asymmetric loss function.

All direct and indirect, lagged and contemporaneous interactions between the model variables can be summarised in terms of quantile impulse response functions (QIRFs) proposed in Chavleishvili and Manganelli (2019). The idea of a QIRF is to explore how the entire con-

²⁶For a detailed exposition see Chavleishvili and Manganelli (2019).

ditional distribution of a variable of interest reacts to unexpected changes in one of the model variables, as opposed to standard impulse response functions that look only into conditional mean effects. For a horizon $h \geq 1$, the QIRFs are defined as

$$Q_{y_{t+h}}(\theta|\mathcal{F}_{t+1}, \delta_i) - Q_{y_{t+h}}(\theta|\mathcal{F}_{t+1}), \quad \theta \in (0, 1)^3, \quad (16)$$

with δ_i denoting the time $t + 1$ value of variable $i \in (1, 2, 3)$, $y_{i,t+1}$, after being hit a certain structural shock. It should be noted that θ is now defined as a three-dimensional vector with elements $\theta_i \in (0, 1)$, implying that the QIRFs can be computed for cases in which the model variables follow paths along different quantiles of their distributions.²⁷

Figure 11 shows QIRFs from simulations in which each variable, in turn, experiences an adverse shock amounting to one standard deviation of its unconditional empirical distribution. The respective post-shock $t + 1$ values δ_i for the CISS, the PMI and real GDP growth are accordingly set to $\delta = (0.2196, 47.3143, -0.9631)$. The QIRFs are computed for horizons $h = 1, \dots, 24$ months and for three different quantiles $\theta \in (0.1, 0.5, 0.9)$, i.e., 10%, 50% and 90%.²⁸ The responses of a variable to the three different shocks are displayed row by row, such that the responses of each of the three variables to a particular shock are shown column by column.

Overall, and as expected, the shape of the mean and median responses are similar to those from the linear VAR when taking into account the opposite signs of the PMI and GDP shocks, as well as the different shock sizes. However, the QIRFs in the first column and in the first row—all associated with shocks to or responses of the CISS—do all feature some asymmetries when comparing the 10% and the 90% quantile responses. Having said that, only the differential responses to CISS shocks are also economically material, while the reactions of the CISS to shocks in the PMI and in real GDP growth are rather small and most likely statistically insignificant for all quantiles. Most interesting in our context are the much stronger impacts of CISS shocks to the PMI and real GDP growth at the 90% quantile compared to the 10% quantile and the median responses. For instance, a CISS shock of +0.18 lowers the annual growth rate of real GDP by maximum 1.1% over a roughly one year horizon at the 90% quantile of the growth distribution, compared with -0.7% and -0.4% for the median and the 10% quantile, respectively. The differences are even starker for PMI responses to the same CISS shock, with the PMI declining by 2.5

²⁷This property of conditional quantile forecasting renders QVAR particularly interesting for macro stress testing exercises as demonstrated in Chavleishvili and Manganelli (2019)

²⁸The QIRFs are constructed using sampling algorithms proposed in Chavleishvili and Manganelli (2019). In this paper we use a fine grid for $\theta_i \in [0.05, 0.1, \dots, 0.9, 0.95]$, $i \in (1, 2, 3)$, and 10^5 sampling repetitions. When the data generating process is a Gaussian VAR, the QIRFs defined at median levels would equal the standard IRFs from a linear VAR calculated as $E(y_{t+h}|\mathcal{F}_{t+1}, \delta_i) - E(y_{t+h}|\mathcal{F}_{t+1})$.

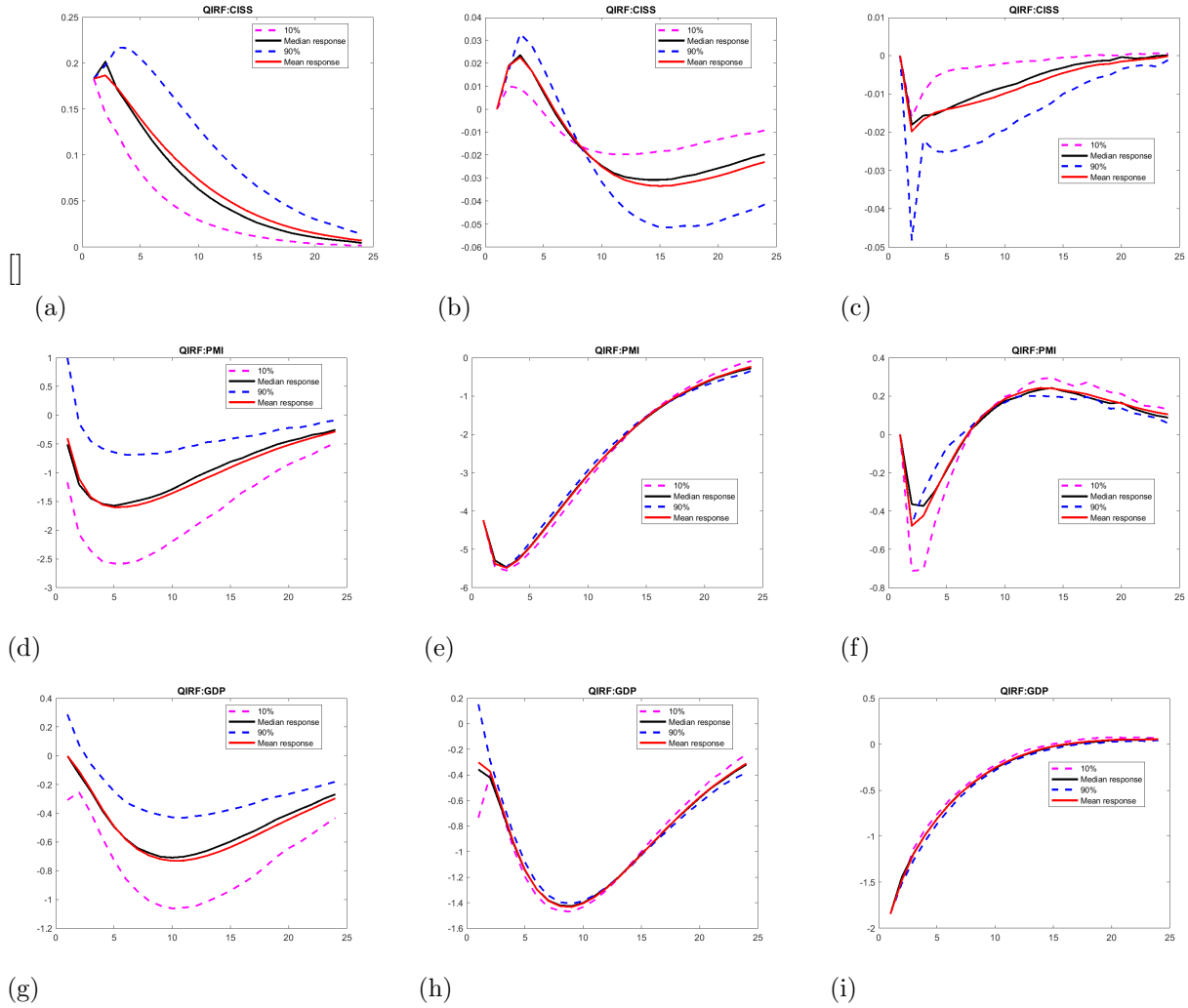


Fig. 11. The figure shows structural quantile impulse response functions (QIRFs) for one standard deviation adverse shocks in each variable. An adverse shock in the CISS is an increase, while an adverse shock for the PMI and annual real GDP growth is a decline. Responses of the CISS, the PMI and real GDP growth are displayed in the first, second and third row, respectively. Correspondingly, responses to shocks in the CISS, in the PMI and in real GDP growth are shown in the first, second and third column, respectively. Each panel shows QIRFs for three quantiles $\theta \in (0.1, 0.5, 0.9)$ of the distribution of the response variable, and mean responses are displayed for comparison.

points at the 90% quantile, 1.5 points at the median and 0.5 points at the 10% quantile. The response of the CISS to own shocks reveals some temporary overshooting at the 90% quantile. This reflects a local nonstationarity in the CISS equation at the upper quantiles of the conditional CISS distribution in line with previous findings (Chavleishvili et al. (2021)). In contrast to the asymmetries found in the macro-financial linkages, the QIRFs between the two economic activity variables appear overall symmetric and similar to those from the linear VAR. Hence, shocks in the PMI bring about similar changes to real GDP growth in all parts of the conditional growth distribution.

We now use the QVAR to forecast real GDP growth out-of-sample conditional on the initial Covid-19 shocks equivalent to the exercise conducted with the linear VAR. Based on the parameters estimated with data until February 2020, we compute the projected values of the conditional density forecasts for real GDP growth up to two years ahead feeding in the realised values (and thus the shocks) of first the CISS and the PMI in March and April 2020. The second exercise only uses these two CISS observations as the forecasting condition. The counterfactuals are computed using the sampling approach suggested in Chavleishvili and Manganelli (2019). The outcomes of the two exercises are shown in Figure 12.

The left panel plots the simulations conditional on shocks in both the CISS and the PMI. These quantile-specific forecast paths are to be compared with the dashed turquoise line in the right part of Figure 10 tracking mean projections from the linear VAR. Not surprisingly, the mean and median responses from the QVAR and the linear VAR are very similar, with peak effects on annual real GDP growth somewhat below -10% by the end of 2020. Looking at the lower tails, growth is projected to decline by around 14% and 13% at the 1% and 5% quantiles, respectively. In typical growth-at-risk analyses, such numbers would be reported. The estimated asymmetries imply that the differences between the median and the upper tail forecasts are smaller with about -8.5% and -7.5% at the 95% and 99% quantiles. The right hand panel displays the forecast paths resulting from the partly offsetting March and April 2020 shocks in the CISS. Consistent with the red dashed line in Figure 10, the maximum mean effect predicts a mild economic contraction over the short term. At the 5% quantile, however, growth appears at risk to experience a sharp recession, with a drop in growth by -3% over a one year horizon, only conditional on the two CISS shocks.

7. Conclusions

Financial stress indexes like the CISS have become an integral part of the macroprudential analytical surveillance toolkit of most authorities. The CISS serves macroprudential policy in particular in its crisis management role by enabling the monitoring of materialised systemic

risk in real time (Freixas et al. (2015)). In addition, the use of the CISS in typical growth-at-risk frameworks (Adrian et al. (2019), Adams et al. (2020)) shows that it predicts well future downside tail risks in real GDP. In this context, Chavleishvili et al. (2021) demonstrate how the CISS, when combined with a measure of the financial cycle and real GDP growth, can also be used for macro stress testing and thus for ex ante macroprudential policy purposes. It is not hard to imagine that such predictive powers can and will also benefit central banks in analysing the short- and long-run costs and benefits from potential monetary policy actions, potentially including those geared towards leaning against financial winds. The CISS may also help predicting macroeconomic variables under conditions of extreme uncertainty caused by extraordinary financial shocks, a point which we illustrated in this paper on the basis of the Covid-19 crisis.

The paper also raises issues which require further research. For instance, why do composite indicators usually possess superior predictive power for standard macroeconomic variables than individual asset prices, and which data compression method works best in this regard? And does this superiority vary with how the composite indicator is designed, or does it simply reflect the fact that combined forecasts tend to improve forecast accuracy over individual forecasts quite generally?

Finally, our proposed method can also be applied to aggregate certain variables into a composite indicator in different economic contexts. For instance, Schüler, Hiebert, and Peltonen (2020) applies time-varying correlation-weighting to combine four market-specific financial cycle measures into an overall financial cycle indicator.

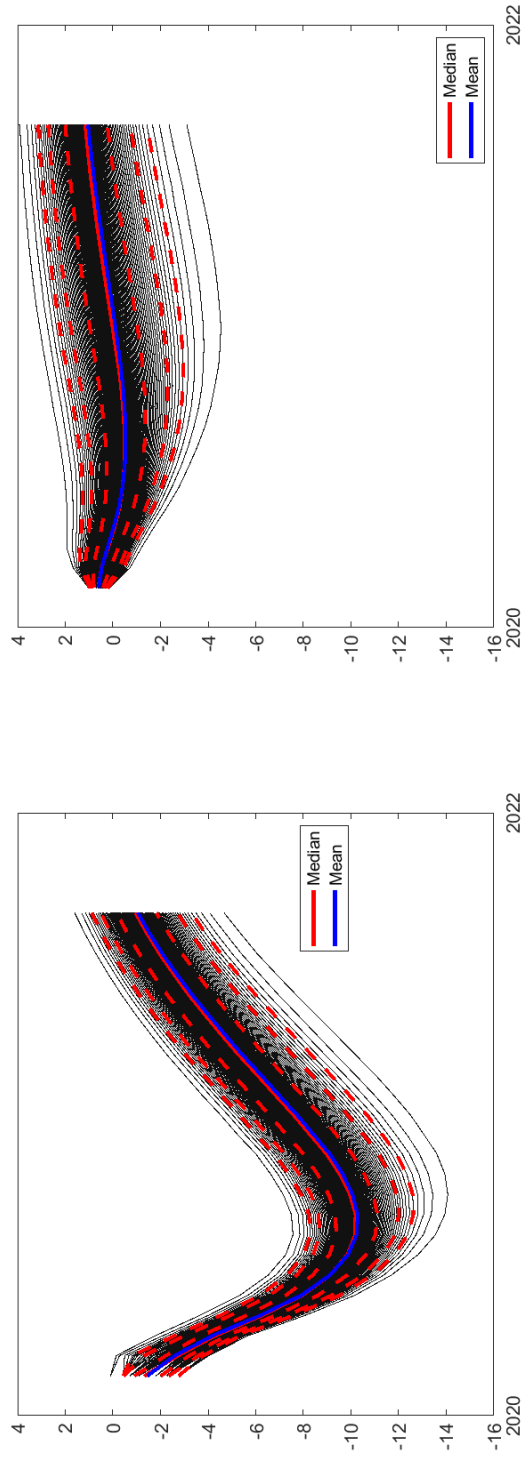


Fig. 12. This figure plots QVAR counterfactual density forecasts of real GDP growth over a two-year horizon with February 2020 the forecast origin. The density forecasts are conditional on realisations of the CISS and the PMI in March and April 2020. Solid black lines correspond to empirical percentiles from 1% to 99% with a step size of 1%. Dashed red lines highlight empirical percentiles from 5% to 95% with a step size of 5%. The left panel shows density forecasts conditional on March and April realisations of both the CISS and the PMI, while the forecast densities in the right panel are conditional only on the CISS.

Appendix A. Composition of the daily CISS

Table 2: Components of the US CISS

Money market

1. Volatility of 3-month nonfinancial AA-rated Commercial Paper (CP)
2. Rate spread 3-month LIBOR against Treasury bills (“Ted spread”)
3. Rate spread 3-month nonfinancial AA-rated CP against Treasury bills

Bond market

4. Volatility of 10-year benchmark government bond price index
5. Yield spread 10-year Moody’s seasoned AAA-rated nonfinancial corporate bonds against Treasury bonds
6. Yield spread 10-year Moody’s seasoned BAA-rated against AAA-rated nonfinancial corporate bonds

Equity market

7. Volatility of nonfinancial stock price index
8. Maximum cumulated loss (CMAX) of nonfinancial stock price index over moving 2-year window: $CMAX_t = 1 - x_t / \max[x_i \in (x_{t-j} | j = 0, 1, \dots, 520)]$
9. Book-price ratio of nonfinancial stock price index

Financial intermediaries

10. Volatility of financial stock price index
11. CMAX of financial stock price index
12. Book-price ratio of financial stock price index

Foreign exchange market

13. Volatility of US dollar exchange rate vis-à-vis euro
14. Volatility of US dollar exchange rate vis-à-vis Japanese Yen
15. Volatility of US dollar exchange rate vis-à-vis Canadian dollar

Notes: Volatilities are computed as exponentially-weighted moving-averages of squared daily log returns with smoothing parameter $\lambda = 0.85$. Data start in January 1973 or when becoming available.

Sources: All raw data is from Refinitiv Datastream; own calculations. Daily US CISS updates are available from the ECB’s Statistical Data Warehouse at <https://sdw.ecb.europa.eu> with series key CISS.D.US.Z0Z.4F.EC.SS.CIN.IDX.

Table 3: Components of the euro area CISS

Money market

1. Volatility of 3-month Euribor
2. Rate spread 3-month Euribor against French Treasury bill

Bond market

3. Volatility of German 10-year benchmark government bond price index
4. Yield spread 10-year interest rate swap against German government bonds
5. Yield spread 7-year A-rated nonfinancial corporate bonds against AAA-rated government bonds
6. Yield spread 7-year A-rated financial corporate bonds against AAA-rated government bonds

Equity market

7. Volatility of nonfinancial stock price index
8. Maximum cumulated loss (CMAX) of nonfinancial stock price index over moving 2-year window: $CMAX_t = 1 - x_t / \max[x_i \in (x_{t-j} | j = 0, 1, \dots, 520)]$
9. Book-price ratio of nonfinancial stock price index

Financial intermediaries

10. Volatility of financial stock price index
11. CMAX of financial stock price index
12. Book-price ratio of financial stock price index

Foreign exchange market

13. Volatility of euro exchange rate vis-à-vis US dollar
14. Volatility of euro exchange rate vis-à-vis Japanese Yen
15. Volatility of euro exchange rate vis-à-vis British pound

Notes: Volatilities are computed as exponentially-weighted moving-averages of squared daily log returns with smoothing parameter $\lambda = 0.85$. Data start in January 1980 or when becoming available.

Sources: All raw data is from Refinitiv Datastream; own calculations. Daily CISS updates are available from the ECB's Statistical Data Warehouse at <https://sdw.ecb.europa.eu> with series key CISS.D.U2.Z0Z.4F.EC.SS.CIN.IDX.

Appendix B. Robustness

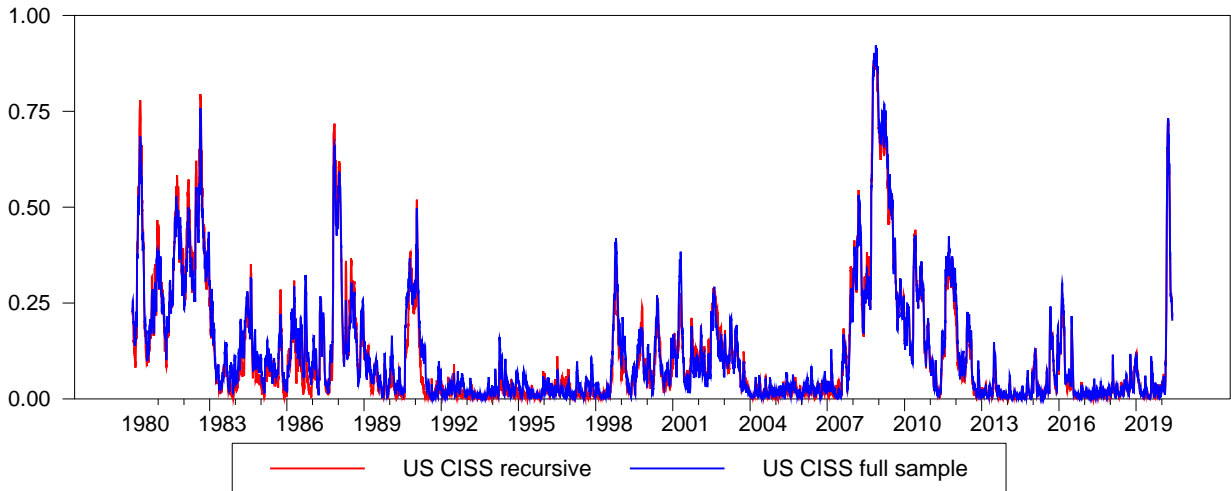


Fig. 13. The figure shows daily values of the standard US CISS computed from stress factors recursively transformed as from 1 January 2002 (red line) and a variant of the CISS using full-sample information at all times (blue line). Data is from 2 January 1980 to 29 May 2020.

Appendix C. Special cases

This Appendix presents different designs of systemic FSIs as special cases of the general statistical framework introduced in Section 3.

1. Variance of the financial system as a whole

This example shall illustrate the economic intuition behind the CISS approach. Assume we observe an N -dimensional vector of daily asset returns, r_t , covering a broad range of financial market segments. The aim is to measure the variance of the financial system as a whole, $\sigma_{FS,t}^2$, based on an estimate of the return variances, $\sigma_{i,t}^2$, and covariances, $\sigma_{i,j,t}$ for $i \neq j$, collected in the variance-covariance matrix Σ_t . Assuming equal weights $1/N$ for all market segments, the total system variance can be computed in the same way as portfolio risk is computed from the risk characteristics of individual assets, i.e. from their return variances and covariances (see, e.g., Brealey and Myers (2000)):

$$\sigma_{FS,t}^2 = \sum_{i=1}^N \sum_{j=1}^N \frac{1}{N} \frac{1}{N} \sigma_{i,j,t}. \quad (17)$$

With $w = (1/N, 1/N, \dots, 1/N)'$ being the vector of equal weights, Equation (17) can also be written in matrix form:

$$\sigma_{FS,t}^2 = w' \Sigma_t w. \quad (18)$$

Using the decomposition $\Sigma_t = D_t R_t D_t$ with $D_t^2 = \text{diag}\{\Sigma_t\}$ the diagonal variance matrix and R_t the correlation matrix, the following equivalent expressions result:

$$\begin{aligned} \sigma_{FS,t}^2 &= w' D_t R_t D_t w \\ &= (w \circ \sigma_t)' R_t (w \circ \sigma_t) \end{aligned} \quad (19)$$

$$\begin{aligned} &= w' \cdot [(\sigma_t \cdot \sigma_t') \circ R_t] \cdot w \\ &= \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N (\sigma_t \cdot \sigma_t')_{i,j} (R_t)_{i,j}. \end{aligned} \quad (20)$$

Interpreting the vector of standard deviations σ_t as a vector of stress indicators, Equation (19) is equivalent to our CISS formula (8). Accordingly, Equation (20) represents the system variance as a scaled matrix association index as proposed in Equations (2) and (7), with extremeness matrix $\mathcal{E}_t = (\sigma_t \cdot \sigma_t')$ and co-dependence matrix $\mathcal{C}_t = R_t$.

In the empirical implementation, we compute 9 daily log price changes (nonfinancial corporation stock prices, financial corporation stock prices, and the three US-dollar exchange rates) and interest rate changes (commercial paper rate, T-bill rate, 10-year government bond yield and the BAA corporate bond yield) from the raw input data of the US CISS. These daily return series are then standardised (z-score transformation) to control for different measurement units. The time-varying variance-covariance matrix Σ_t of standardised returns is estimated on the basis of a DCC-IGARCH model with asymmetry effects and Student's t -distributed errors (Engle (2002)). Figure 14 plots the such measured variance of the US financial system along with the US CISS and the VIX. During well-known episodes of extreme stress, all three series display similar pronounced spikes. However, while both the total system variance and the VIX tend to recede from those spikes relatively quickly, the higher levels of the CISS tend to be more persistent in general. This behaviour of the CISS is due to the inclusion of stress measures like risk spreads which also show a stronger persistence than asset volatilities.

2. Special cases from the literature

This part illustrates how other popular index designs from the literature can also be represented as special cases of our general statistical framework for systemic FSIs. The way how the raw input series x_t are transformed is one important design feature of any

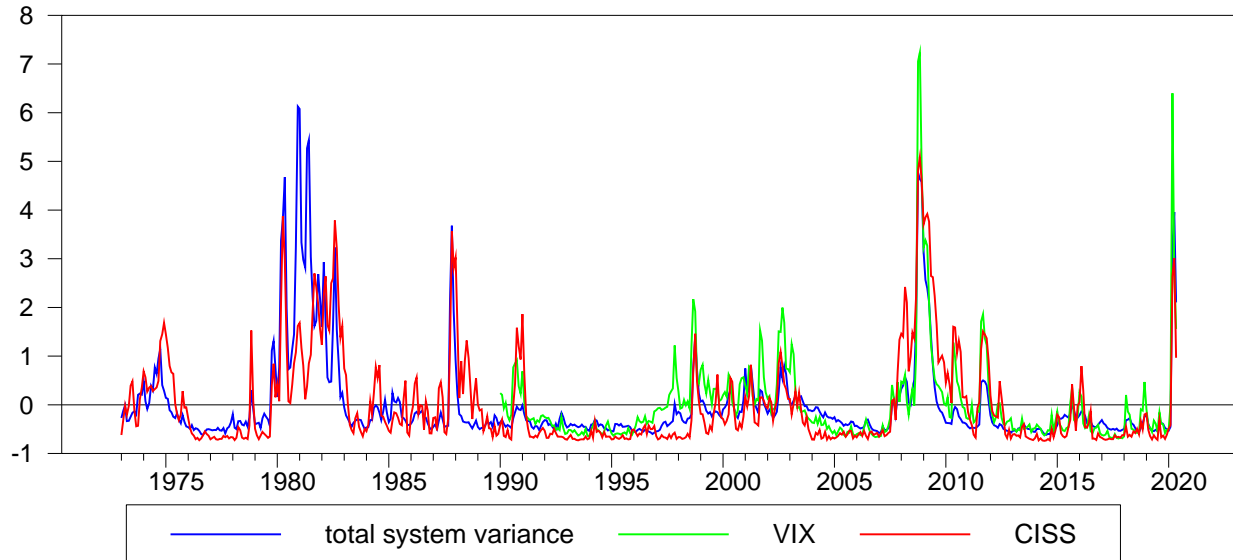


Fig. 14. The figure plots the estimated variance of the US financial system as a whole against the VIX and the US Daily CISS. Total system variance is computed according to Equation (20) based on 9 standardised daily return series for the US market, and the variance-covariance matrix of standardised returns estimated by a DCC-IGARCH model. The VIX series shown is the square of the CBOE Volatility Index which measures the risk-neutral expectation of the S&P 500 index variance based on options on that index. For the purpose of this chart, the three series are standardised and original daily data converted into monthly averages. Data run from January 1973 to May 2020. Sources: Refinitiv Datastream and own calculations.

FSI. We consider two different transformations: the probability integral transform based on the empirical cdf (giving rise to stress factors z_t), and the z-score standardisation ($\tilde{x} = (x_t - \bar{x})/\sigma_x$). In terms of aggregation scheme, we differentiate between simple weighted averages (as in Oet et al. (2011) and Cardarelli et al. (2011)), principal components analysis (like in Hakkio and Keeton (2009)), a dynamic factor model approach (in the spirit of Brave and Butters (2012)), and the turbulence index (by Kritzman and Li (2010)).

Table 4 describes the design of seven FSIs, including the CISS, in terms of how they operationalise the extremeness and co-dependence dimensions (columns 2 and 3), and how these are combined in the index formula (column 4). The simple-average FSIs for the two transformations (FSI-avg-cdf and FSI-avg-std) are trivial cases since they do not perform any systemic risk weighting, which is why co-dependence is represented by an all-ones vector.²⁹ The two PCA-based indexes (FSI-PCA-cdf and FSI-PCA-std) and the dynamic factor model

²⁹Averaging over standardised indicators has also been labelled *variance-equal weighting*. If an FSI averages using weights calibrated to reflect the relative size of market segment, this could also be considered as a form of systemic risk weighting.

(FSI-DF-std) follow the idea that the systemic dimension of financial stress can be captured by a latent factor driving the observed correlation between the input indicators. The latent factor accordingly serves as the FSI. In the PCA approaches, the latent factor is identified as the first principal component of the sample correlation matrix of the transformed indicators. The systemic risk weights of the input series in the composite indicator are accordingly computed from their loadings to the first principal component, i.e. from the first eigenvector. The dynamic factor model follows the same logic but estimates the latent common factor f_t based on state space methods. The estimated loadings γ to the common factor can be interpreted as the systemic risk weights since they implicitly reflect to which extent each input series contributes to the dynamics of the estimated common factor.³⁰ The turbulence index measures the Mahalanobis distance between the standardised indicators over time. The Mahalanobis distance is a multi-dimensional generalisation of the idea of how many standard deviations a set of variables are away from their means. The turbulence index thus captures how “unusual” the input series jointly behave at each point in time, and it turns out that periods of thus measured unusualness broadly coincide with well-known episodes of heightened financial stress. The formula behind the turbulence index looks similar to that of the CISS, but weights the cross-products by the inverse of the full-sample correlation matrix.

In the empirical application we apply the different indicator designs to the same set of input series, namely to the 15 components of the US CISS. This ensures that any differences between the FSIs exclusively reflect differences in how the components are transformed and aggregated. The resulting seven FSIs are plotted in the two panels of Figure C. Both panels reveal, first, that the simple-average FSIs and the PCA-based FSIs are almost identical for both transformations, reflecting that the first eigenvectors of the spectral decomposition does not differ much from the vector of equal weights. This in turn suggests that all components contribute on average to a similar extent to the total variability of the data, implying that no indicator appears to be significantly more systemic than any other. Second, while the FSI estimated from a simplified dynamic factor model appears somewhat smoother than the PCA index, they still do not deviate much from each other in general. Third, in terms of cdf-transformed indicators, the CISS differentiates more strongly between systemic and nonsystemic stress episodes due to its time-varying correlation weighting. As shown in Section 4.6, the CISS and the simple-average FSI only coincide when all input series are

³⁰The dynamic factor model is set up as a simplified version of that in Stock and Watson (1991). The state space model is estimated by Maximum Likelihood methods applying the Kalman filter. The measurement equation is $\hat{x}_t = \gamma f_t + v_t$; the state equation is $f_t = \phi f_{t-1} + w_t$, with γ the vector of loadings to the single latent factor f_t , ϕ the autoregressive coefficient of the dynamic factor in the state equation, and v_t and w_t are i.i.d. error terms, where the variance of w_t is normalised to 1 to achieve identification of the dynamic factor.

Table 4: Special cases of systemic financial stress indexes

FSI concept	Extremeness (\mathcal{E}_t)	Co-dependence (\mathcal{C}_t)	Index formula
CISS	$z_t z_t'$	R_t	$(w \circ z_t)' R_t (w \circ z_t)$
FSI-average-cdf	z_t	ι_N	$w' z_t$
FSI-PCA-cdf	z_t	e_1	$e_1' z_t$
FSI-average-std	\tilde{x}_t	ι_N	$w' \tilde{x}_t$
FSI-PCA-std	\tilde{x}_t	e_1	$e_1' \tilde{x}_t$
FSI-DF-std	\tilde{x}_t	γ	$\tilde{x}_t - \hat{v}_t = \gamma f_t$
Turbulence	$\tilde{x}_t \tilde{x}_t'$	R^{-1}	$(w \circ \tilde{x}_t)' R^{-1} (w \circ \tilde{x}_t)$

Notes: The table describes different FSI designs in terms of different measurements of the extremeness and co-dependence dimensions according to Definition 2 and Equation 2. For the sake of brevity, the index formulas are written in matrix notation. x_t denotes the N -dimensional vector of raw stress indicators, \tilde{x}_t their z-score transforms and z_t their probability integral transforms; R denotes the correlation matrix of stress factors; ι_N represents the trivial case of unitary co-dependence (systemic risk) weighting; e_1 denotes the first eigenvector of the spectral decomposition of the variance-covariance matrix of stress factors; γ is the vector of estimated loadings of the stress factors on a single latent dynamic factor f_t (the FSI-DF-std), with \hat{v}_t denoting the residuals from the measurement equation of the dynamic factor model; w is the vector of equal weights with $w_i = 1/N$.

close to being perfectly correlated. Fourth, the turbulence index indeed mainly identifies extreme events. It displays a few large jumps around well-known stress episodes while hovering around low levels most of the time.

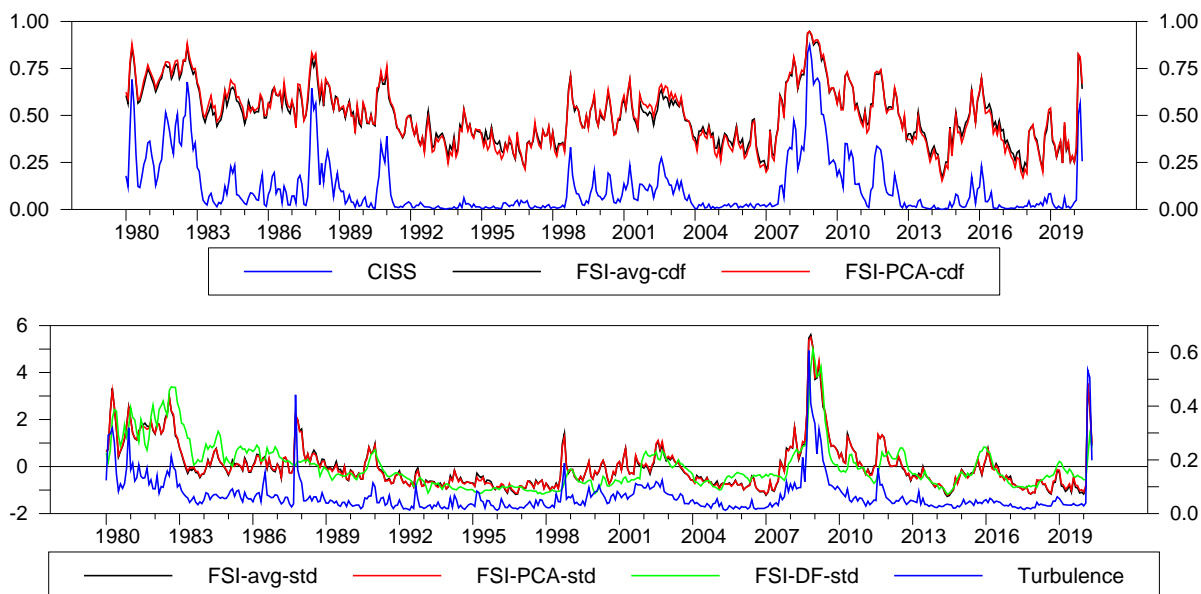


Fig. 15. The figure plots seven different FSIs designed according to Table 4 and applied to the 15 component series of the US CISS as described in Table 2. The first three series in the lower panel are standardised for graphical purposes. Data is monthly averages of daily data from January 1980 to May 2020.

Appendix D. VAR data

The CISS for the euro area and the US is computed as the monthly average of daily data. Data is from own calculations and published in the ECB’s SDW.

The original quarterly real GDP data is interpolated to the monthly frequency applying state space methods, using monthly industrial production as an interpolator variable and assuming that the interpolation error can be described as a log-linear ARIMA(1,1,0) process as in Litterman (1983). Estimation is implemented using the procedure DISAGGREGATE.SRC in WinRATS. Euro area data is taken from the ECB’s SDW, US data from the FRED database.

The Markit Eurozone PMI Composite Output Index tracks business trends across both the manufacturing and service sectors, based on data collected from a representative panel of over 5,000 companies (60 percent from the manufacturing sector and 40 percent from

the services sector). The index tracks variables such as sales, new orders, employment, inventories and prices. National data are included for Germany, France, Italy, Spain, Austria, the Netherlands, Greece and the Republic of Ireland. Both the euro area and the US PMI are from Markit.

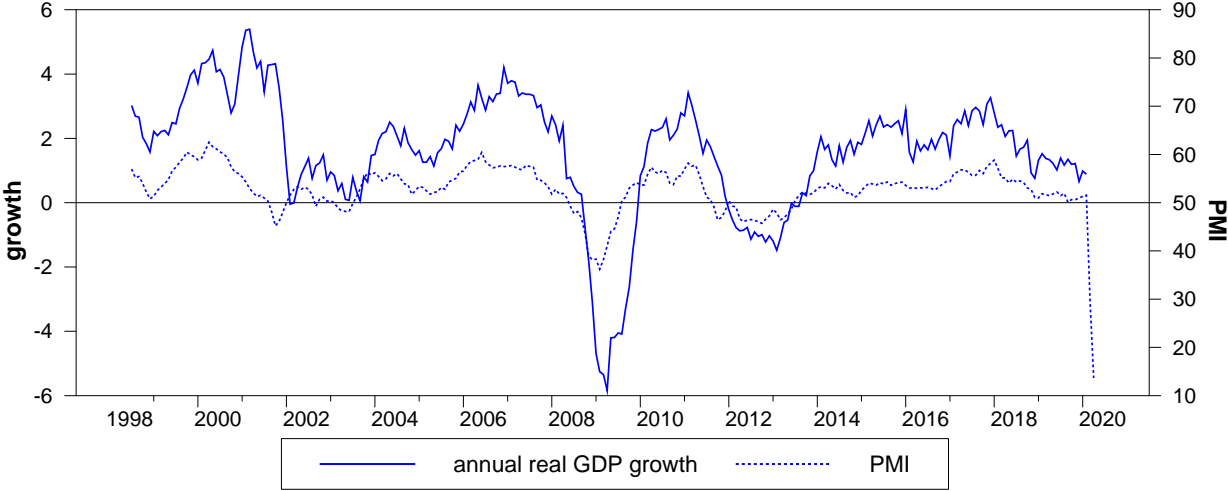


Fig. 16. The figure shows annual growth in log real GDP and the PMI Composite Output Index for the euro area. A PMI reading above 50 indicates expansion in business activity and below 50 indicates contraction. Log real GDP is interpolated from quarterly to monthly frequency informed by industrial production. Data is monthly from July 1998 to April 2020 (GDP: to 2019:Q4, industrial production: to February 2020).

Appendix E. Supplementary VAR results

Table 5: Forecast Error Variance Decomposition

A. Variable ordering: CISS, PMI, GDP					
Decomposition of Variance for Series CISS					
Step	Std Error	CISS	PMI	GDP	
1	0.063	100.00	0.00	0.00	
3	0.114	98.71	0.97	0.31	
6	0.151	98.38	1.26	0.36	
12	0.175	98.51	1.03	0.45	
24	0.183	97.73	1.79	0.48	
Decomposition of Variance for Series PMI					
Step	Std Error	CISS	PMI	GDP	
1	0.976	5.90	94.10	0.00	
3	2.192	16.79	82.73	0.48	
6	3.233	22.65	77.09	0.26	
12	4.049	27.02	72.56	0.43	
24	4.234	28.46	70.68	0.86	
Decomposition of Variance for Series real GDP					
Step	Std Error	CISS	PMI	GDP	
1	0.422	1.41	3.74	94.85	
3	0.663	6.92	9.16	83.92	
6	0.972	20.95	27.71	51.34	
12	1.518	35.37	42.12	22.51	
24	1.833	41.57	42.79	15.65	
B. Variable ordering: PMI, CISS, GDP					
Decomposition of Variance for Series PMI					
Step	Std Error	PMI	CISS	GDP	
1	0.976	100.00	0.00	0.00	
3	2.192	95.89	3.63	0.48	
6	3.233	93.09	6.66	0.26	
12	4.049	90.19	9.39	0.43	
24	4.234	88.66	10.48	0.86	
Decomposition of Variance for Series CISS					
Step	Std Error	PMI	CISS	GDP	
1	0.063	5.90	94.10	0.00	
3	0.114	10.53	89.16	0.31	
6	0.151	11.71	87.93	0.36	
12	0.175	10.51	89.04	0.45	
24	0.183	9.89	89.64	0.48	
Decomposition of Variance for Series real GDP					
Step	Std Error	PMI	CISS	GDP	
1	0.422	4.69	0.46	94.85	
3	0.663	12.75	3.33	83.92	
6	0.972	38.65	10.01	51.34	
12	1.518	59.89	17.60	22.51	
24	1.833	62.45	21.91	15.65	

Notes: The table shows forecast error variance decompositions from the linear VAR for two different variable orderings reported in Panels A and B, respectively. Forecast steps are monthly, and the contributions from the structural shocks in the three variables add up to 100 in each row.

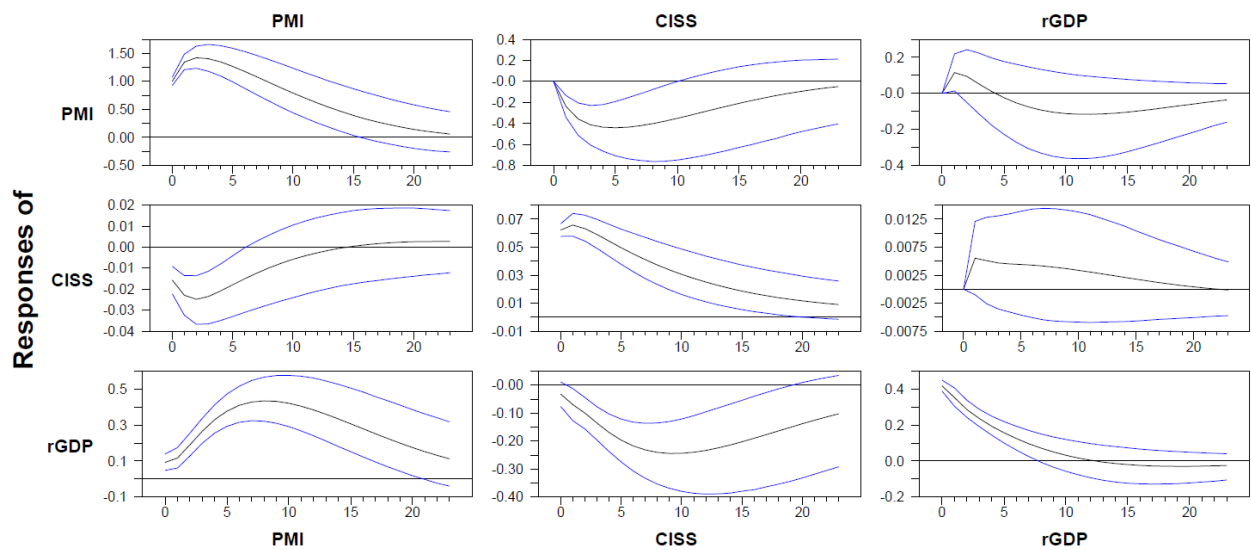


Fig. 17. The figure presents mean impulse response functions (black line) from the linear VAR for the case of structural identification by Cholesky decomposition with variables ordered PMI-CISS-GDP. The blue lines represent 90% error bands estimated via Monte Carlo integration. Forecast steps are monthly.

References

- Adams, P., Adrian, T., Boyarchenko, N., Giannone, D., Liang, N., Qian, E., 2020. What do financial conditions tell us about risks to GDP growth?, Federal Reserve Bank of New York Liberty Street Economics, May 21, <https://libertystreeteconomics.newyorkfed.org/2020/05/what-do-financial-conditions-tell-us-about-risks-to-gdp-growth.html>.
- Adrian, T., Boyarchenko, N., Giannone, D., 2019. Vulnerable growth. *American Economic Review* 109, 1263–1289.
- Adrian, T., Brunnermeier, M. K., 2016. CoVaR. *American Economic Review* 106, 1705–1741.
- Altavilla, C., Barbiero, F., Boucinha, M., Burlon, L., 2020. The great lockdown: pandemic response policies and bank lending conditions, ECB Working Paper No. 2465, September.
- Aruoba, S. B., Diebold, F. X., Scotti, C., 2009. Real-time measurement of business conditions. *Journal of Business and Economic Statistics* 27, 417–427.
- Beck, T., 2014. Finance, Growth, and Stability: Lessons from the Crisis. *Journal of Financial Stability* 10, 1–6.
- Bekaert, G., Hoerova, M., 2014. The VIX, the Variance Premium and Stock Market Volatility. *Journal of Econometrics* 183, 181–192.
- Bianchi, J., 2011. Overborrowing and systemic externalities in the business cycle. *American Economic Review* 79, 3400–3426.
- Billio, M., Getmansky, M., Lo, A. W., Pelizzon, L., 2012. Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of Financial Economics* 104, 535–559.
- Bisias, D., Flood, M., Lo, A. W., Valavanis, S., 2012. A survey of systemic risk analytics. *Annual Review of Financial Economics* 4, 255–296.
- Bloom, N., 2009. The impact of uncertainty shocks. *Econometrica* 77, 623–685.
- Brave, S., Butters, A., 2011. Monitoring financial stability: a financial conditions index approach, Federal Reserve Bank of Chicago Economic Perspectives, First Quarter.
- Brave, S., Butters, A., 2012. Diagnosing the financial system: Financial conditions and financial stress. *International Journal of Central Banking* 8, 191–239.

- Brealey, R. A., Myers, S. C., 2000. Principles of corporate finance. McGraw-Hill, Boston, 6th ed.
- Brownlees, C., Chabot, B., Ghysels, E., Kurz, C., 2020. Back to the future: Backtesting systemic risk measures during historical bank runs and the Great Depression. *Journal of Banking and Finance* 113.
- Brownlees, C., Engle, R. F., 2017. SRISK: A conditional capital shortfall measure of systemic risk. *Review of Financial Studies* 30, 48–79.
- Cardarelli, R., Elekdag, S., Lall, S., 2011. Financial stress and economic contractions. *Journal of Financial Stability* 7, 78–97.
- Carlson, M. A., King, T., Lewis, K., 2011. Distress in the financial sector and economic activity. *B.E. Journal of Economic Analysis and Policy* 11, Art. 35.
- Carlson, M. A., Lewis, K. F., Nelson, W. R., 2012. Using policy intervention to identify financial stress, Finance and Economics Discussion Series 2012-02. Washington: Board of Governors of the Federal Reserve System.
- Carr, P., 2017. Why is VIX a fear gauge? *Risk and Decision Analysis* 6, 179–185.
- Casella, G., Berger, R., 2002. Statistical inference. Brooks/Cole, second ed.
- Chavleishvili, S., Engle, R., Fahr, S., Kremer, M., Manganelli, S., Schwaab, B., 2021. The risk management approach to macro-prudential policy, ECB Working Paper, forthcoming.
- Chavleishvili, S., Manganelli, S., 2019. Forecasting and stress testing with quantile vector autoregression, ECB Working Paper No. 2330, November.
- Christiano, L., Eichenbaum, M., Evans, C., 1999. Monetary policy shocks: what have we learned and to what end? In: Taylor, J., Woodford, M. (eds.), *Handbook of Macroeconomics, Vol. 1, Chap. 2*, Elsevier, pp. 65–148.
- Davig, T., Hakkio, C., 2010. What is the effect of financial stress on economic activity?, Federal Reserve Bank of Kansas City Economic Review, Second Quarter.
- de Bandt, O., Hartmann, P., 2000. Systemic risk: a survey, ECB Working Paper No. 35.
- Dovern, J., van Roye, B., 2014. International transmission and business-cycle effects of financial stress. *Journal of Financial Stability* 13, 1–17.
- Engle, R., 2009. Anticipating correlations. Princeton University Press, Princeton and Oxford.

- Engle, R. F., 2002. Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business and Economic Statistics* 20, 339–350.
- Figueres, J. M., Jarociński, M., 2020. Vulnerable growth in the euro area: Measuring the financial conditions. *Economics Letters* 191, 109126.
- Freixas, X., Laeven, L., Peydró, J.-L., 2015. Systemic risk, crises, and macroprudential regulation. MIT Press.
- Giese, J., Haldane, A., 2020. Covid-19 and the financial system: a tale of two crises. *Oxford Review of Economic Policy* 36, S200–S214.
- Gieve, J., 2006. Pricing for perfection. Bank of England, Speech at the Bank of England, 14 December 2006 .
- Giglio, S., Kelly, B., Pruitt, S., 2016. Systemic risk and the macroeconomy: An empirical evaluation. *Journal of Financial Economics* 119, 457–471.
- Gilchrist, S., Zakrajšek, E., 2012. Credit spreads and business cycle fluctuations. *American Economic Review* 102, 1692–1720.
- Grimaldi, M. B., 2010. Detecting and interpreting financial stress in the euro area, ECB Working Paper No. 1214, June.
- Hakkio, S. C., Keeton, W. R., 2009. Financial stress: what is it, how can it be measured, and why does it matter?, *Federal Reserve Bank of Kansas City Economic Review*, Second Quarter.
- Hamilton, J. D., 2018. Why you should never use the Hodrick-Prescott filter. *Review of Economics and Statistics* 100, 831–843.
- Hartmann, P., Hubrich, K., Kremer, M., Tetlow, R. J., 2015. Melting down: Systemic financial instability and the macroeconomy, mimeo.
- Hatzius, J., Hooper, P., Mishkin, F. S., Schoenholtz, K. L., Watson, M. W., 2010. Financial conditions indexes: A fresh look after the financial crisis, NBER Working Paper No. 16150, July.
- He, Z., Krishnamurthy, A., 2012. A model of capital and crises. *Review of Economic Studies* 79, 735–777.

- Hóllo, D., Kremer, M., Lo Duca, M., 2012. CISS - A composite indicator of systemic stress in the financial system, ECB Working Paper No. 1426, March.
- Hubrich, K., Tetlow, R. J., 2015. Financial stress and economic dynamics: The transmission of crises. *Journal of Monetary Economics* 70, 100–115.
- Illing, M., Liu, Y., 2006. Measuring financial stress in a developed country: an application to Canada. *Journal of Financial Stability* 2, 243–265.
- Jarocinski, M., Mackowiak, B., 2017. Granger causal priority and choice of variable in vector autoregressions. *Review of Economics and Statistics* 99, 319–329.
- Kliesen, K. L., Owyang, M. T., Vermann, E. K., 2012. Disentangling diverse measures: a survey of financial stress indexes. *Federal Reserve Bank of St. Louis Review* 94, 369–398.
- Koenker, R., Bassett, G., 1978. Regression Quantiles. *Econometrica* 46, 33–49.
- Kremer, M., 2016. Macroeconomic effects of financial stress and the role of monetary policy: A var analysis for the euro area. *International Economics and Economic Policy* 13, 105–138.
- Kritzman, M., Li, Y., 2010. Skulls, financial turbulence, and risk management. *Financial Analysts Journal* 66, 30–41.
- Kritzman, M., Li, Y., Page, S., Rigobon, R., 2011. Principal components as a measure of systemic risk. *Journal of Portfolio Management* 37, 112–126.
- Laeven, L., Valencia, F., 2008. Systemic banking crises: a new database, IMF Working Paper No. 08/224.
- Laeven, L., Valencia, F., 2013. Systemic banking crises database. *IMF Economic Review* 61, 225–270.
- Laeven, L., Valencia, F., 2018. Systemic banking crises revisited, IMF Working Paper No. 18/206.
- Levine, R., 2005. Finance and growth: Theory and evidence. In: Aghion, P., Durlauf, S. N. (eds.), *Handbook of Economic Growth, Vol. 1A*, Amsterdam: Elsevier, pp. 865–934.
- Litterman, D., 1983. A random walk, Markov model for the distribution of time series. *Journal of Business and Economic Statistics* 1, 169–173.
- Lorenzoni, G., 2008. Inefficient credit booms. *Review of Economic Studies* 75, 809–833.

- Mallick, S. K., Sousa, R. M., 2013. The real effects of financial stress in the eurozone. *International Review of Financial Analysis* 30, 1–17.
- Mantel, N., 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research* 27, 209–220.
- Markowitz, H., 1952. Portfolio selection. *Journal of Finance* 7, 77–91.
- Mendoza, E., 2010. Sudden stops, financial crisis, and leverage. *American Economic Review* 100, 1941–1966.
- Nelson, W. R., Perli, R., 2007. Selected indicators of financial stability. *Irving Fisher Committee’s Bulletin on Central Bank Statistics* 23, 92–105.
- Oet, M., Eiben, R., Bianco, T., Gramlich, D., Ong, S., 2011. The financial stress index: identification of systemic risk conditions, Federal Reserve Bank of Cleveland Working Paper No. 11-30, November.
- Patel, S. A., Sarkar, A., 1998. Crises in developed and emerging stock markets. *Financial Analysts Journal* 54, 50–61.
- Reinhart, C. M., Rogoff, K. S., 2009. *This time is different: eight centuries of financial folly*. Princeton University Press, Princeton, NJ.
- Romer, C. D., Romer, D. H., 2017a. New evidence on the aftermath of financial crises in advanced countries. *American Economic Review* 107, 3072–3118.
- Romer, C. D., Romer, D. H., 2017b. New evidence on the aftermath of financial crises in advanced countries: Dataset. *American Economic Review* .
- Schüler, Y., Hiebert, P., Peltonen, T. A., 2020. Financial cycles: characterisation and real-time measurement. *Journal of International Money and Finance* 100, 82–102.
- Stock, J., Watson, M., 1991. A probability model of the coincident economic indicators. In: Lahiri, K., Moore, G. (eds.), *Leading Economic Indicators: New Approaches and Forecasting Records, Chapter 4*, Cambridge University Press, Cambridge.
- Stuart, A., Ord, J., 1994. Distribution theory. In: *Kendall’s Advanced Theory of Statistics, Vol. 1*, Arnold, London.
- Wasserman, L., 2003. *All of Statistics: A Concise Course in Statistical Inference*. Springer, Berlin.