EUROPEAN CENTRAL BANK
EUROSYSTEM

Working Paper Series

Sulkhan Chavleishvili, Manfred Kremer

Measuring systemic financial stress and its risks for growth

No 2842

Abstract

This paper proposes a general statistical framework for systemic financial stress indices which measure the severity of financial crises on a continuous scale. Several index designs from the financial stress and systemic risk literature can be represented as special cases. We introduce an enhanced daily variant of the CISS (composite indicator of systemic stress) for the euro area and the US. The CISS aggregates a representative set of stress indicators using their time-varying cross-correlations as systemic risk weights, computationally similar to how portfolio risk is computed from the risk characteristics of individual assets. A bootstrap algorithm provides test statistics. Single-equation and system quantile growth-at-risk regressions show that the CISS has stronger effects in the lower tails of the growth distribution. Simulations based on a quantile VAR suggest that systemic stress is a major driver of the Great Recession, while its contribution to the COVID-19 crisis appears to be small.

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

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

Non-technical Summary

This paper proposes a general statistical framework for systemic financial stress indices that is 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. Financial stress indices quantify the aggregate level of stress in the financial system by compressing a certain number of stress indicators from individual financial market segments into a single statistic. We consider systemic stress as realised systemic risk, and thus as a measure of the severity of financial crises. Our statistical framework defines systemic stress as a state of the financial system in which a representative set of individual stress measures is considered to be 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 demonstrate that several indicators from the literature on financial stress indices and systemic risk indicators can be represented as special cases of our general framework.

The paper also introduces an enhanced daily variant of the ECB's Composite Indicator of Systemic Stress (CISS, pronounced /kɪs/) as a non-parametric operationalisation of the general statistical framework. Since its first release in 2011, the CISS concept has been adopted by many financial stability authorities around the world as a blueprint for their own financial stress index. Moreover, the CISS has become a popular tool in academic research as a measure of crisis severity and general financial conditions, not least in the context of the recent macroeconomic "outcome-at-risk" literature. In this paper, we provide a rigorous statistical foundation for the design of the CISS. In the empirical application, we compute the CISS for the euro area and the US. However, daily updates of the new CISS are available for a broader set of countries via the ECB's Statistical Data Warehouse (https://sdw.ecb.europa.eu/browse.do?node=9689686). Both the euro area and the US CISS aggregate 15 components that capture stress symptoms in money, bond, equity and foreign exchange markets. All raw input variables are first transformed into relative ranks using 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 their 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. Accordingly, the CISS gives more weight to situations in which stress becomes widespread and thus systemic. Correlations can also capture externalities such as contagion or spillovers from one part of the financial system to the financial system as a whole, a feature that any measure of systemic risk should take into account. From a statistical point of view, the various steps in the design of the CISS aim to provide a composite indicator that does not suffer from look-ahead bias, is sufficiently robust to outliers, is largely unaffected by different distributional properties of the underlying raw data, and is easy to compute and update. Finally, we propose a bootstrap algorithm to test whether the CISS exceeds a level that can be considered as normal and harmless.

The final part of the paper investigates the empirical linkages between systemic stress and economic growth. It is a stylised fact that systemic financial crises lead to large losses in output and employment. Any meaningful measure of systemic financial stress should therefore be able to replicate this fact. We apply single-equation and VAR-based quantile regressions to assess the short-run forecasting properties of the CISS for real GDP growth. Indeed, we find a stronger predictive power of the CISS for short-term economic activity in the lower tails of the growth distribution, i.e. in bad states of the economy. In an in-sample quantile predictive regression horserace, the CISS is found to have superior short-term forecasting power compared to different index designs and other financial indicators from the literature. Simulations with a quantile VAR suggest a dominant role for the CISS in explaining the deep recession observed during the

GFC. This differs from the COVID-19 crisis, where financial stress shocks play a minor role relative to aggregate output shocks, despite the large initial jump in the CISS.

1. Introduction

> Money is a veil, but when the veil flutters, real output sputters.
>
> John G. Gurley (1961)

Finance and growth are twin sisters. Finance helps overcome frictions in the real sector arising from transaction and information costs, thereby supporting the savings and investment decisions of economic agents 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 interdependence between finance and growth holds true in good times as well as bad. The Global Financial Crisis (GFC) reminded us that financial development can sometimes 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 caused, for example, by externalities, information asymmetries, incomplete markets or limited human cognitive abilities (Bisias et al. (2012)). When such financial frictions intensify and prevail, leading to widespread financial stress, they tend to have severe repercussions on the real economy. A better understanding of these important macro-financial linkages requires - in addition to appropriate theoretical models - meaningful empirical measures of financial stress.

There are many indicators that measure stress in different market segments. Each of these indicators captures certain market- and instrument-specific stress symptoms, such as increased market volatility or wider default and liquidity risk premia, which in turn reflect stress reactions such as increased uncertainty, higher risk aversion or run-like phenomena such as flight-to-safety and flight-to-liquidity. For example, volatility measures derived from option prices provide information on the degree of risk aversion and uncertainty of market participants (Bekaert and Hoerova (2014)); the most well-known of such volatility measures, the VIX, is often used as a general "fear gauge" and thus as an indicator of financial stress (Carr (2017)). Such standard indicators form the backbone of any financial stability surveillance toolkit. However, the sheer number of individual stress measures complicates the task of inferring whether the stress observed in a particular market segment is either idiosyncratic or instead a more system-wide phenomenon. Sometimes you "can't see the wood for the trees."

Financial stress indices are one way of synthesising such scattered information. A financial stress index (FSI) quantifies the aggregate level of stress by compressing a number of individual stress indicators into a single statistic. This paper proposes a general statistical framework for *systemic* financial stress indices 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 et al. (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 at any point in time. A systemic FSI can thus be seen as a coincident indicator of financial (in)stability, measuring the severity of financial crises on a continuous scale. The statistical framework defines systemic stress as a state of the financial system in which a representative set of stress indicators is considered to

be extremely high and highly co-dependent. The composite indicator results from a matrix association index that combines two matrices quantifying the extremeness and the co-dependence stress dimensions.

We also propose an enhanced variant of the ECB's Composite Indicator of Systemic Stress (CISS, pronounced /kɪs/, originally developed by Hóllo et al. (2012)) as a non-parametric operationalisation of this framework. The present paper provides a rigorous statistical foundation for the CISS. In the empirical application, we introduce new daily CISS series for the US and the euro area. Both indicators aggregate 15 components that capture stress symptoms in money, bond, equity and foreign exchange markets. All raw input variables are first transformed into relative ranks using 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 their 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. Accordingly, the CISS gives more weight to situations in which stress becomes widespread and thus systemic. Correlations can also capture externalities such as contagion or spillovers from one part of the financial system to the financial system as a whole, a feature to which any measure of systemic risk should pay attention to (Freixas et al. (2015)). From a statistical point of view, the various steps in the design of the CISS aim to provide a composite indicator that does not suffer from look-ahead bias, is sufficiently robust to outliers, is largely unaffected by different distributional properties of the underlying raw data, and is easy to compute and update. Finally, we propose a bootstrap algorithm to test whether the CISS exceeds a level that can be considered as normal and harmless.

The final part of the paper investigates the empirical linkages between systemic stress and economic growth. It is a stylised fact that systemic financial crises lead to large losses in output and employment. Any meaningful measure of systemic financial stress should therefore be able to replicate this fact. We apply single-equation and VAR-based quantile regressions to assess the short-run forecasting properties of the CISS for real GDP growth. Indeed, we find a stronger predictive power of the CISS for short-term economic activity in the lower tails of the growth distribution, i.e. in bad states of the economy. In an in-sample quantile predictive regression horserace, the CISS is found to have superior short-term forecasting power compared to different index designs and other financial indicators from the literature. Simulations with a quantile VAR suggest a dominant role for the CISS in explaining the deep recession observed during the GFC. This differs from the COVID-19 crisis, where financial stress shocks play a minor role relative to aggregate output shocks, despite the large initial jump in the CISS.

The paper is organised as follows. The next section presents a selective literature review and outlines how the paper contributes to the various areas covered in the review. 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 proposes a statistical inference tool to address the joint extremeness-co-dependence hypothesis underlying the CISS. Section 6 examines the empirical linkages between the CISS and measures of economic activity. Section 7 concludes, and several appendices provide supplementary information.

2. Related Literature

The paper contributes to several strands of the literature. First, our paper speaks to the literature on financial crisis indicators. Crisis indicators aim to identify periods of extreme stress in either one or more 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 quantitative criteria, qualitative information and events. For example, Laeven and Valencia (2008, 2013, 2018) create annual crisis dummies that capture systemic distress in the banking system when 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 that identify 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 given country. Romer and Romer (2017a,b) go one step further in differentiating the severity of crises. Using the semi-annual OECD Economic Outlook as a real-time source of information, they identify and classify financial stress events into five groups, from pure credit disruptions, minor crisis, moderate crisis, major crisis up to extreme crisis. The severity of the stress within each group is further differentiated into three sub-categories (minus, regular, and plus events) to produce a "measure of financial distress" with values ranging from 0 for no distress periods to a maximum of 15 for an extreme crisis-plus. Unlike qualitative ordinal crisis indicators, FSIs use 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 (such as that of Aruoba et al. (2009)) and qualitative recession dummies, such as those published by the business cycle dating committees of the NBER for the United States and the CEPR-EABCN for the euro area. Romer and Romer (2017a) acknowledges 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 indices such as the CISS mitigate such problems. Like Reinhart and Rogoff (2009) and Romer and Romer (2017a), the CISS takes into account stress in several systemically important segments of the financial system, but aggregates stress in these market segments based on weights derived from economic and statistical principles that operationalise the concept of systemic risk. As a continuous measure, the CISS also allows 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 by construction is bounded between zero and one. Panel A shows the Reinhard/Rogoff and the Laeven/Valencia indicators, and panel B the Romer/Romer indicator at a monthly frequency, assigning the value of the annual or semi-annual data to 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 constructed to better reflect crisis severity. This notwithstanding, the US financial system appears to have experienced several major stress events that are not captured by any of the qualitative indicators. It can be argued that FSIs can

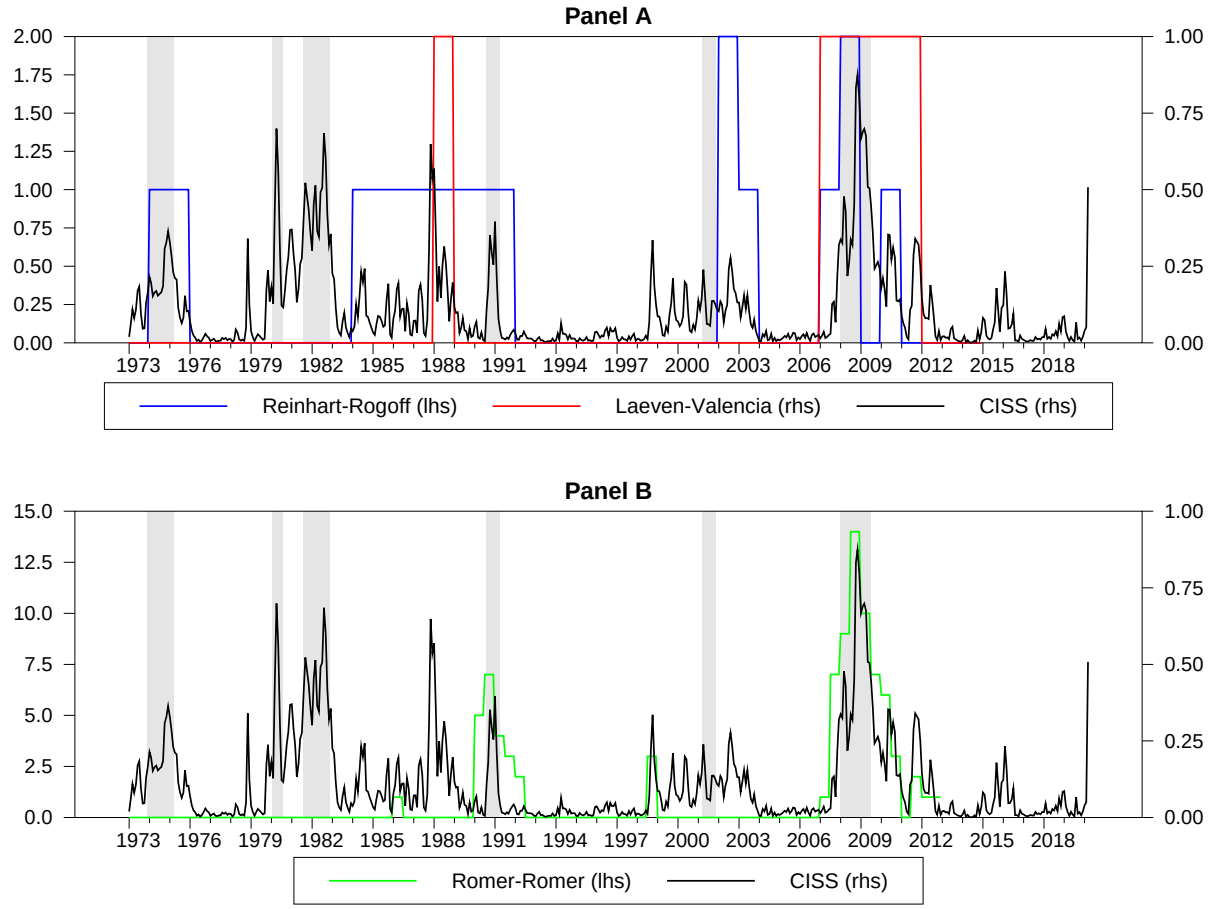

Fig. 1. The figure plots the Composite Indicator of Systemic Stress (CISS, black line) along with three qualitative financial crisis indicators for the United States. Panel A shows the crisis indicators of Reinhard and Rogoff (2009, blue line) and Laeven and Valencia (2018, red line); the indicator of 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.

only measure financial stress net of the impact of policy interventions, whereas most qualitative indicators explicitly consider the magnitude of policy reactions when assessing a potential crisis event.[1] As a result, the FSIs categorise crisis severity according to the measured level of stress, but this may reflect very different intensities of policy intervention. Notwithstanding this, it may still be possible to construct appropriate counterfactuals that allow the FSI to be cleansed of the effects of policy interventions, or, equally, to quantify 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 et al. (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

[1] An exception is Carlson et al. (2012), who use policy interventions to identify crisis events and construct the FSI.

aggregating a given set of individual stress indicators into a composite stress index. Their favoured design is determined by which variant performs best in signalling crisis events in the Canadian financial system by evaluating Type I and Type II errors; the crisis events are identified through a survey of Bank of Canada policymakers and staff. The preferred FSI consists of 11 daily financial market variables aggregated on the basis of weights determined by the relative size of the market to which each indicator relates. Cardarelli et al. (2011) present a monthly financial stress index for 17 advanced economies, calculated as the arithmetic mean of 12 standardised financial stress indicators. Nelson and Perli (2007) and Carlson et al. (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 predefined binary crisis indicator as the dependent variable. As a refinement of the last step of this approach, Blix 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 et al. (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. Kliesen and Smith (2010) and, more recently, Groen et al. (2022) follow the same approach, with the latter paper constructing monthly FSIs for 46 countries. The weekly index developed by Brave and Butters (2011, 2012) is also based 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 the National Financial Conditions Index (NFCI), which is published regularly by the Federal Reserve Bank of Chicago, is cast in a dynamic factor model in state-space form with a maximum of 100 indicators, where Kalman filtering deals with the problem of missing data resulting from the different sample lengths and frequencies of the input data.

Such factor models also provide non-trivial systemic risk weights. The models cited all include stress indictors from a wide range of financial market segments. In addition, factor models 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 weights (Brave and Butters (2011)). Principal component analysis of a set of asset returns is also behind the systemic risk measures proposed by Kritzman et al. (2011) and Billio et al. (2012). The "absorption ratio" of Kritzman et al. (2011) takes daily returns of 51 US stock market sectors and computes, over a 500-day moving window, the fraction of the total variance of these returns explained by the first ten eigenvectors of the covariance matrix. The ratio is higher when there is more commonality between sector returns. The idea is that high values of the absorption

ratio indicate a state of increased system fragility, since a given shock to financial stability would tend to propagate more quickly and more 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", which corresponds to the fraction of total return variance explained by a given number of statistically significant principal components. The turbulence index of Kritzman and Li (2010) aggregates asset returns based on the Mahalanobis distance. The index aims to capture the statistical "unusualness" of a set of returns given their historical pattern of behaviour.

Our paper presents a general statistical framework for measuring systemic financial stress, a concept that blends the ideas behind financial stress indices and systemic risk indicators. In Section 4.7, we illustrate that several of the aforementioned indicator designs from both fields can be represented as special cases of our general statistical framework. Following the taxonomy of systemic risk measures by the analytical time horizon (pre-event, contemporaneous, and post-event) proposed by Bisias et al. (2012), our general framework produces contemporaneous systemic risk measures that fit into both the "fragility" and the "crisis monitoring" subcategories of such 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)) indicators. While the main purpose of the latter two studies is to measure the contribution of individual financial institutions to the aggregate systemic risk of the sector comprising these institutions, the CISS measures the degree of interconnectedness of various aggregate market segments and how this contributes to the level of systemic stress in the financial system as a whole. These perspectives can therefore 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 interconnectedness as autoregressive cross-correlations, borrowing from the GARCH literature. The PCA approaches of Kritzman et al. (2011) and Billio et al. (2012) also measure the time variation in cross-correlations, but do so by way of a fixed moving-window estimation (0-1-weighting of historical information). In contrast, our exponentially-weighted-moving-average approach places decaying weights on more distant information, which is flexible enough to capture abrupt changes in correlation patterns as they typically occur during periods of stress. Finally, our paper complements Hóllo et al. (2012) by enhancing the statistical foundation of the CISS design. For instance, we substantiate the use of relative ranks (probability integral transform) as a more robust and homogeneous transformation of raw input indicators. This 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.[2] Since its introduction in Hóllo et al. (2012), the statistical concept of the CISS has been adopted by many financial stability authorities across the world as the blueprint for

[2]Compared 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; therefore, it 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 are often perceived as arbitrary. Fourth, the set of input series changes somewhat and becomes more harmonised across the US and euro area indices. Fifth, the euro area CISS carries a longer data history, starting in 1980.

their own financial stress index.[3] The CISS concept is also applied to aggregate different stress symptoms into composite indicators of stress prevailing in specific important market segments like the markets for European sovereign bonds (Garcia-de-Andoain and Kremer (2017)) and US corporate bonds (Boyarchenko et al. (2022)).[4]

Third, the paper contributes to the rich literature on the real effects of financial distress. In linear growth prediction frameworks, several studies capture financial stress using FSIs (including the CISS) or financial conditions indices (FCIs), and they 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 et al. (2010)). In line with theoretical predictions, studies applying nonlinear regression frameworks have found stronger real effects in bad states of the world. The set of nonlinear approaches include, inter alia, threshold VARs (Hóllo et al. (2012)), Markov-switching VARs (Davig and Hakkio (2010), Hubrich and Tetlow (2015), Hartmann et al. (2015)), quantile VARs (Chavleishvili and Manganelli (2019), Chavleishvili et al. (2021)), and single-equation quantile regressions as common in the recent growth-at-risk literature (Adrian et al. (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 et al. (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 single-equation and VAR-system quantile regressions on euro area and US data. The structural quantile VAR (QVAR) approach recently developed by Chavleishvili and Manganelli (2019) allows us to capture, apart from asymmetries in the dynamic relationships, potential feedback effects between financial stress and economic growth. The results from both approaches corroborate previous studies that find particularly strong effects of financial stress on future economic activity in the lower tails of the growth distribution. While Adams et al. (2020) and Figueres and Jarociński (2020) also use our new CISS for the US and the euro area, respectively, in single-equation quantile growth-at-risk regressions, this paper is the first to use it in the QVAR framework. Furthermore, we quantify the macroeconomic relevance of systemic stress by performing conditional growth projections based on historical shocks in the CISS. The simulation results demonstrate the different nature of the GFC and the COVID-19 crisis, with the first being mainly driven by shocks in systemic stress and the latter by macroeconomic activity shocks.

[3] 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.

[4] Such composite indicators are regularly published by the ECB and the Federal Reserve Bank of New York, respectively.

3. A general statistical framework for measuring systemic financial stress

This section proposes a general statistical framework for measuring systemic financial stress using a composite indicator, namely a systemic financial stress index (FSI). A systemic FSI $\mathcal{S}_t$ can be thought of as a statistical function combining three types of ingredients: i) an N-dimensional vector of raw stress indicators x_t, for $t = 1, \ldots, T$; ii) a conformable vector of "portfolio share 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 to satisfy the following assumption:

Assumption 3.1. *Raw 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 $x_{i,t_1} > x_{i,t_2}$ ($x_{i,t_1} < x_{i,t_2}$).*

Assumption 3.1 requires that the raw stress indicators are constructed in such a way that higher indicator levels are associated with higher levels of financial stress. In practice, this assumption also implies that a raw stress indicator should appear mean-reverting, possibly in a highly nonlinear manner, over the available data sample in order to allow a meaningful comparison of stress levels over time. Indicators that follow a deterministic upward or downward trend can therefore be excluded. Moreover, an observable raw stress indicator should cover a wider range of stress levels. This does not mean, however, that the support of an indicator's empirical distribution must cover all conceivable outcomes. Financial crises are rare events, and each new crisis can extend the scale and add significant mass to the upper tail of an indicator's empirical distribution function. We therefore expect the indicator design to face statistical problems arising from small data samples, such as imprecise or even biased estimates of an indicator's empirical distribution function. It is therefore all the more important, in our view, that the statistical properties of a systemic stress indicator are sufficiently robust to the addition of outlier observations. Abstracting from such small-sample issues, we could simply assume that the raw stress indicators follow weakly stationary and ergodic stochastic processes in the population, and amend Assumption 3.1 accordingly.

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

Assumption 3.2. *Raw stress indicators $x_{i,t}$ are transformed into normalised stress factors $z_{i,t}$ by applying a monotone transformation $x_{i,t} \to g(x_{i,t}) \equiv z_{i,t}$ 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 in the literature. However, as we argue later, z-score standardisation has certain comparative disadvantages over our preferred alternative, the probability integral transform, in the context of systemic FSIs.

For aggregation, the stress factors $z_{i,t}$ are assigned portfolio share weights $w_{i,t}$ with $0 < w_{i,t} < 1$ and $\sum_{i=1}^{N} w_{i,t} = 1$ for $i = 1, \ldots, N$. The positivity assumption ensures that no stress indicator is redundant. We borrow the term "portfolio share weights" from standard portfolio

theory in finance, where the weights $w_{i,t}$ would capture the share of asset i in the total asset portfolio. In the context of FSIs, $w_{i,t}$ could represent the relative size of the different market segments covered by the set of stress indicators. In turn, the relative size could be measured in terms of the stocks or flows of financial instruments traded in each 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.[5] However, the simplest 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 a standard FSI by introducing a weighting of indicators that operationalises the concept of systemic stress. The main idea behind the systemic weighting is that, in order to assess the broader implications of stress in a particular market segment, it is necessary to consider how that 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 trigger a systemic crisis with large macroeconomic costs. Allowing for bilateral differences in the measured connectivity between stress indicators, the systemic risk weights may be conceived of as some $N \times N$ matrix function $\mathcal{C}_t$.

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

Definition 1 (Systemic Financial Stress)**.** *Systemic financial stress is a state of the financial system in which stress factors $z_{i,t}$ are generally extremely high (extremeness) and strongly co-dependent (co-dependence).*

According to Definition 1, systemic financial stress rules out cases where high stress remains locally confined and does not spill over to other significant parts of the financial system due to generally weak cross-market co-dependencies. In fact, we assume that in a systemic crisis, a vast majority of stress factors $z_{i,t}$ jointly increase to extremely high levels due to strong co-dependencies among them. Hence, jointly high stress among the stress factors implies strong co-dependence. But the converse is not true, as stress factors can also be highly co-dependent when they are jointly extremely low, but not high. It is therefore useful to distinguish between the stress dimensions of extremeness and co-dependence. We can thus characterise systemic stress or financial instability as a state of "co-extremeness", i.e., extremeness cum co-dependence, rendering the notion of widespread financial stress more concrete.[6]

Definition 2 introduces the final building block of our proposed general statistical framework for a systemic FSI.[7]

Definition 2 (Systemic Financial Stress Index)**.** *Let extremeness and co-dependence be measured by some $N \times N$ bounded real-valued matrices $\mathcal{E}_t$ and $\mathcal{C}_t$, respectively. A systemic financial*

[5] As an alternative, Hóllo et al. (2012) estimate weights based on the relative predictive power of each subindex of stress for industrial production growth.

[6] In this paper, we use the terms *extreme* and *extremeness* only in the sense of *extremely high*.

[7] The following notation is used throughout the paper. The scalar $(A)_{i,j}$ represents the i,jth 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.

stress index $\mathcal{S}_t$ can be defined as a matrix association index that quantifies the degree of co-extremeness as expressed in Equation (1)*:*

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

The scaling factor $1/N^2$ reflects the assumption of equal weighting, i.e., $w_i = 1/N$. The extremeness matrix $\mathcal{E}_t$ is usually a simple function of the time-t realisations of the stress factors z_t, and the co-dependence matrix $\mathcal{C}_t$ represents the matrix of systemic risk weights that can be estimated or calibrated. The index in Equation (1) can be viewed as a scaled matrix association index as proposed by Mantel (1967). Such indices are widely used in various fields of science to model and test the similarity of multidimensional data observed in matrix form.[8] In our case, the index shows the extent to which realisations $z_{i,t}$ 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, i.e. the hypothesis that stress factors are not jointly high and strongly co-dependent at any given point in time. Hence, the co-extremeness hypothesis represents the alternative hypothesis in our proposed testing scheme.

4. Operationalising systemic stress - the CISS

As our favoured operationalisation of the general statistical framework summarised in Equation 1, we propose an improved version of the ECB's Composite Indicator of Systemic Stress (CISS), originally developed by one of the authors and collaborators in the working paper Hóllo et al. (2012). The present paper introduces new daily CISS series for the United States and the euro area.[9] As we will argue, the CISS has certain specific advantages which, in our view, distinguish it from alternative FSI designs. Section 4.7 demonstrates how other popular FSI designs can be represented as special cases of our general framework.

4.1. Raw stress indicators

We start with selecting $N = 15$ raw stress indicators, $x_{i,t}$, with $i = 1, \ldots, N$ and $t = 1, \ldots, T$, from the main segments of the US and the euro area financial systems. Taken together, these market segments should cover the main flows of financial funds that are channelled either indirectly through financial intermediaries or directly through 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.

[8] The scaled version of this measure can also be seen as Moran's I or Geary's C statistics, which are widely used in modelling economic networks based on geographical distances.

[9] Daily updates of the new daily CISS for Austria, Belgium, China, the euro area, Finland, France, Germany, Greece, Ireland, Italy, the Netherlands, Portugal, Spain, the UK and the US are available via the ECB's Statistical Data Warehouse: https://sdw.ecb.europa.eu/browse.do?node=9689686.

Our choice of indicators is constrained by the following considerations. First, to ensure representativeness, the raw stress indicators should be based either 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 maximum publication lag of one business day) to support the CISS as a quasi real-time financial stability monitoring tool. And third, the stress indicators should carry long data histories to allow meaningful historical benchmarking of stress events and to make the CISS potentially useful for 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.[10] All market segments include (at least) a measure of historical volatility, computed as an exponentially-weighted moving average (EWMA) of squared daily log returns or squared daily interest rate changes with a smoothing parameter of $\lambda = 0.85$; initial values are set to the return variance over the first two years of the respective data sample (Engle (2009), p. 30f.).[11] Other stress measures include various 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 of Patel and Sarkar (1998)) to capture equity market stress. All raw stress indicators included in the US and the euro area CISS are described in Tables 3 and 4 in Appendix A.

4.2. Stress factors

Next, all raw stress indicators $x_{i,t}$ are transformed into stress factors $z_{i,t}$ by applying the probability integral transform (PIT), which involves estimating the empirical cumulative distribution function (cdf):[12]

$$z_{i,t} = \hat{F}(x_{i,t}) := \begin{cases} \frac{1}{T_0-1}\sum_{s=1}^{T_0-1} \mathcal{I}(x_{i,s} \leq x_{i,t}), & \text{for} \quad t = 1, \ldots, T_0 - 1 \\ \frac{1}{t}\sum_{s=1}^{t} \mathcal{I}(x_{i,s} \leq x_{i,t}), & \text{for} \quad t = T_0, \ldots, T, \end{cases} \tag{2}$$

where $\hat{F}(x_{i,t})$ is defined as the empirical cdf of the indicator $x_{i,t}$ and $\mathcal{I}(x)$ is the indicator function. Hence, each realisation of the stress factor $z_{i,t}$ simply results from replacing the original score $x_{i,t}$ with its cdf value $\hat{F}(x_{i,t})$. Since $\hat{F}(x_{i,t})$ is non-decreasing, the PIT satisfies Assumption 3.2. To address some anticipated statistical problems with small data samples, we compute the PIT recursively 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 on 1 January 2002.[13] The recursive transformation immunises the composite indicator against look-ahead bias (Brownlees et al. (2020)) and, relatedly, against event reclassification, i.e. the risk of recalling a stress event reported in the past (based only on past data) that is not identified

[10] For a general discussion of individual stress indicators and their information content, see Hakkio and Keeton (2009), Hóllo et al. (2012) and the literature cited therein.

[11] The 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.

[12] See, e.g., Casella and Berger (2002).

[13] This date is chosen to ensure at least several years of data for the initialisation of a few euro area time series that start only in 1999 or shortly before.

as a stress event when future information is also taken into account (Hóllo et al. (2012)).

There are several reasons why we prefer the PIT to the z-score transformation for our purposes.[14] 1. *Homogeneity.* – The PIT leads to indicators that are unitless 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 is not true for z-score standardisation. 2. *Robustness.* – The use of ranks in the PIT makes the stress factors, and thus the composite indicator, more robust to outliers by limiting 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. First, it 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 a meaningful subset of it (see Figure 14 in Appendix B). Consequently, the CISS is also robust to the choice of different starting dates T_0 for the recursive PIT (Hóllo et al. (2012)). And second, this robustness is most important when the composite indicator is most useful, namely during episodes of financial instability, which can add many outliers to the expanding data samples of the raw stress indicators. Figure 2 below shows - 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 identifies several stress peaks prior to the GFC, most of these indicated events are almost completely "ironed out" when using the full-sample mean and standard deviation for the transformation. The picture is much less dramatic when the probability integral transform is used (panel a). A third major advantage of the PIT - namely the robustness of our measure of co-dependence - is discussed in Subsection 4.4.

4.3. Extremeness

The vector of stress factors z_t is a natural measure of extremeness. Stress in a particular market segment is most extreme when 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. To do this, we quantify extremeness as the cross-product of all non-centered stress factors, $z_{i,t} \cdot z_{j,t}$, and collect them in the symmetric matrix $\mathcal{E}_t$ of order N:

$$\mathcal{E}_t = z_t z_t'. \tag{3}$$

Since 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 $J_N = \iota_N \iota_N'$ at maximum system stress and approaches a zero matrix $O_N = 0_N 0_N'$ at minimum stress.

[14] We are not the first to use 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.

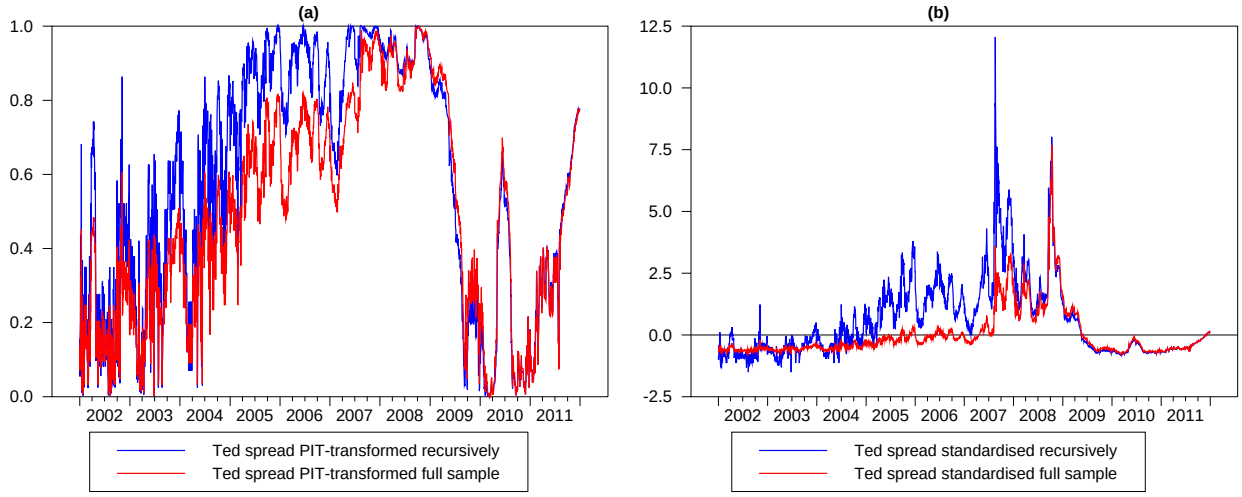

Fig. 2. The figure assumes that Ted spread data were only available from January 2002 to December 2011, i.e. ten years of daily data. The subprime mortgage turmoil and the Global Financial Crisis took place in the second half of the sample period. In panel (b), the Ted spread is z-score standardised recursively from January 2006 (blue line) and non-recursively using the full sample information at all points in time (red line). The figure shows that, particularly in the case of small samples, the information content of z-score transformed indicators can become unstable. Panel (a) suggests that this problem is less severe for probability integral transformed indicators.

4.4. Co-dependence

We compute the co-dependence non-parametrically as the conditional rank correlation (Spearman's ρ) between each of the $N(N-1)/2$ pairs of stress factors. The rank correlation tells us whether and to what extent two stress factors are similarly high or low at any given time. For instance, if two stress factors move together towards the upper ranges of their empirical cdfs, then, ceteris paribus, there is a higher risk that financial stress has become more widespread and thus systemic. If they move asynchronously, this risk tends to be lower.

Since our stress factors inherit some of the autocorrelation and heteroskedasticity of the original data[15], 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 $k = 0, 1, ..., t-1$ and some initial value $\rho_{i,j,0}$. The conditional expectation is non-parametrically modelled as an autoregressive exponentially-weighted moving average (EWMA) process, following the multivariate GARCH literature (Engle (2002)). The rank correlation coefficient is defined against a centrality measure. We take the population median 0.5 as the centrality measure and accordingly define the vector of centered stress factors as $\widetilde{z}_t = (z_t - \iota_N \cdot 0.5)$. The EWMA filter is implemented on the variance-covariance matrix H_t of the centered stress factors $\widetilde{z}_t$:

$$H_t = \lambda H_{t-1} + (1-\lambda)\widetilde{z}_t\widetilde{z}_t' \tag{4}$$

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

[15] We do not pre-whiten or de-GARCH the raw stress indicators before applying the PIT to the "cleaned" residuals, as is done, for example, in the standard copula literature.

hand, it still accommodates abrupt and large shifts in the correlation patterns that typically occur during crisis periods. The elements $\rho_{i,j,t}$ of the 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:

$$\mathcal{C}_t = R_t. \tag{5}$$

Estimating the rank correlation as the correlation coefficient between PIT-transformed data ensures that our non-parametric measure of co-dependence is fairly robust to distributional assumptions and outliers (Engle (2009), p. 22). For example, Spearman's rank correlation relies only on the existence of a monotonic relationship between the original stress indicators. In contrast, raw stress indicators with heterogeneous statistical distributions may co-vary in a nonlinear manner despite monotonicity. A linear correlation between the raw stress indicators or their z-scores may therefore fail to capture such nonlinear dependencies. Since we are working with the PIT and with rank correlations, our approach to measuring co-dependence is loosely based on the copula literature (Engle (2009)).

Figure 3 visualises realisations of the matrices $\mathcal{E}_t$ and $\mathcal{C}_t$ on four different days, characterising stress conditions before and during the GFC and the COVID-19 crisis, respectively. Panel (a) represents the pre-GFC period, when volatilities and risk premia were persistently low by historical standards in virtually 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.[16] This benign picture in early 2007 soon gave way to the subprime mortgage crisis and ultimately to the GFC. At the height of the GFC, stress levels were consistently at or slightly below their historical maxima across the board, as shown in panel (b), which is coloured dark red almost everywhere on 10 October 2008. Panels (c) and (d) highlight how quickly financial stability deteriorated when the US was struck by the COVID-19 shock. By 28 February 2020, financial stress conditions were mixed across markets and indicators and very benign overall. Accordingly, panel (c) covers more or less the full colour spectrum. Within just one month, the situation changed completely, with financial stress rising sharply across all market segments. The systemic dimension of the coronavirus-induced spike in financial stress is evident in the similarity of panels (b) and (d). As 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 to historically high levels by 31 March 2020.

4.5. Composite Indicator

Equipped with all the ingredients, the CISS can now be easily computed as the matrix association index $\mathcal{S}_t$ as defined in Equation (1) based on the quantified inputs from Equations

[16] See the notes section of Figure 3 for how to infer correlation levels from the colour bars.

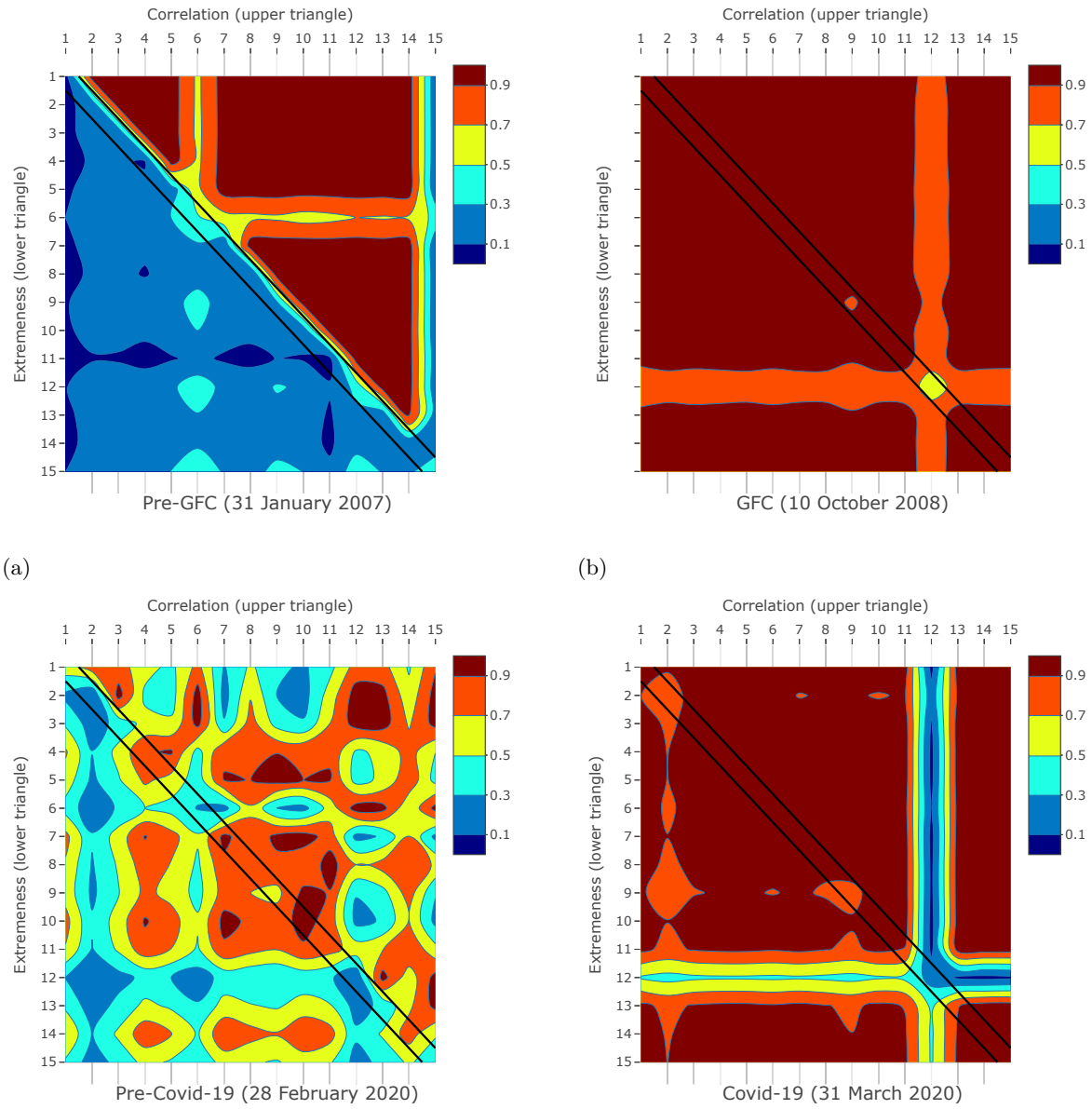

Fig. 3. The figure shows contour plots visualising realisations of the extremeness and the correlation matrices on four days representing financial stability conditions before and during the Global Financial Crisis (GFC) and the COVID-19 crisis. As both matrices are symmetric, extremeness is shown in 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 the US data. The order of the stress factors follows the numbering in Table 3. Extremeness is plotted as the square root of the cross-product $\sqrt{z_{i,t}z_{j,t}}$ (the geometric mean) consistent with 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 $\tilde{\rho}_{i,j,t} = 0.5(\rho_{i,j,t} + 1)$.

(3) and (5):

$$\text{CISS}_t = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \left(z_t z_t'\right)_{i,j} \left(R_t\right)_{i,j}, \tag{6}$$

with $0 < \text{CISS}_t \leq 1$. As mentioned above, the scaling parameter $1/N^2$ simply represents an equal portfolio weighting $w_i = 1/N$ of the 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 only approach 1 if stress is persistently high or low.

This formulation of the composite indicator problem is not only statistically attractive but also economically intuitive. The formula is well-known in finance and describes how to compute the return variance (risk) of a portfolio of assets from the return variances and covariances of a set of assets with equal portfolio shares (see, e.g., the seminal paper on portfolio selection by Markowitz (1952)).

The CISS formula (6) can be equivalently expressed as a quadratic form:

$$\text{CISS}_t = (w \circ z_t)' R_t (w \circ z_t), \tag{7}$$

where $\circ$ denotes element-by-element multiplication. Figure 4 shows the CISS for the US and the euro area since January 1973 and January 1980 respectively. Most of the pronounced peaks in the indicators occur simultaneously and can be associated with well-known global stress events such as the stock market crash of 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 some episodes that reflect more localised shocks, such as the Fed's monetary tightening regime under Chairman Paul Volcker from 1979 to 1982 and the associated savings and loan 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.

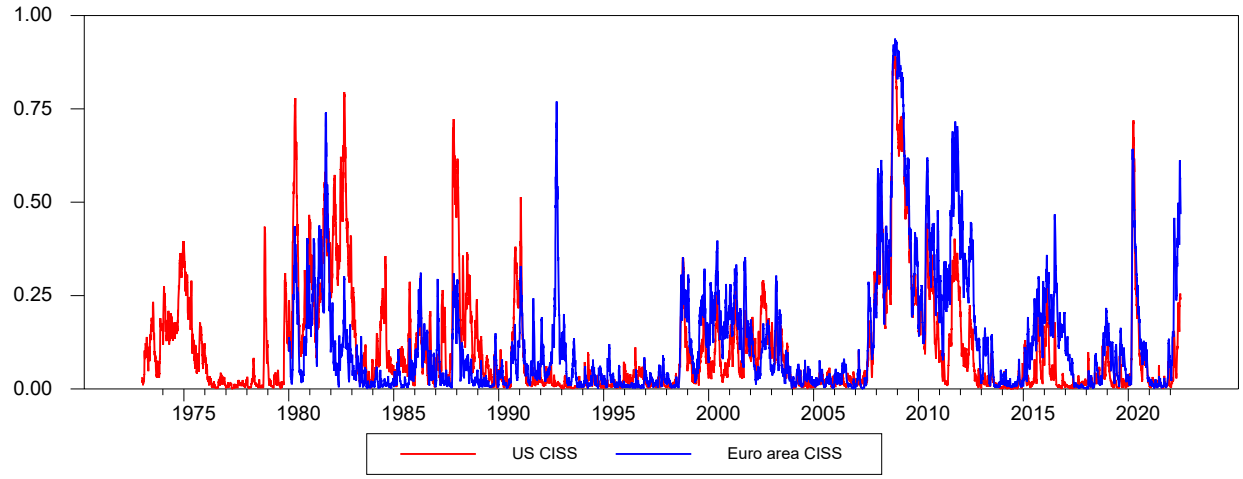

Fig. 4. This figure plots daily CISS data for the US and the euro area up to 1 July 2022. US and euro area data start on 3 January 1973 and 1 January 1980 respectively.

4.6. A Decomposition

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

$$\begin{aligned} \text{CISS}_t &= (w \circ z_t)' J_N (w \circ z_t) = (w \circ z_t)'(w \circ z_t) \\ &= (\textstyle\sum_{i=1}^{N} w_i \cdot z_{i,t})^2 = \bar{z}_t^2. \end{aligned} \tag{8}$$

Hence, in the case of perfect correlation, the square of the equally-weighted average of the stress factors $\bar{z}_t^2$ emerges as the upper bound of the CISS. This implies that in a systemic crisis, when stress is high across the board, the CISS converges to $\bar{z}_t^2$. Conversely, when cross-correlations are generally weaker, the CISS deviates more from the simple-average FSI. Accordingly, we call the difference between the CISS and $\bar{z}_t^2$ the *correlation discount* and use it for the following decomposition of the CISS:

$$\text{CISS}_t = \sum_{i=1}^{N} \frac{\bar{z}_t}{N} z_{i,t} - \left(\frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} z_{i,t} z_{j,t} (1 - \rho_{i,j,t}) \right). \tag{9}$$

The first term of Equation (9) is simply $\bar{z}_t^2$ decomposed 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.[17] The figure confirms that the CISS and the stacked contributions are close to coincide during the GFC, with the correlation discount approaching zero. More importantly, the pattern of the correlation discount - the grey shaded area in Figure 5 - clearly illustrates the main advantage of the CISS: it helps to better identify episodes of truly widespread financial stress (systemic crises) by giving less weight to situations where elevated stress remains a more local, market-specific event. The dot-com boom and bust of the late 1990s and early 2000s is a prime example of a largely non-systemic episode of stress. Thus, in the case of the CISS, the systemic risk weighting achieves that "the whole is [smaller] than the sum of the parts."

4.7. Special cases from the literature

This section illustrates how several designs of financial stress and systemic risk indicators from the literature can be represented as special cases of the general statistical framework introduced in Section 3. The way in which the raw input series x_t are transformed is an important design feature of any FSI. We consider two different transformations: the probability integral transform based on the empirical cdf (which gives rise to the stress factors z_t), and the z-score standardisation ($\tilde{x} = (x_t - \bar{x})/\sigma_x$). Regarding the aggregation scheme, we distinguish between simple weighted averages (as in Oet et al. (2011) and Cardarelli et al. (2011)), principal component analysis (as in Hakkio and Keeton (2009)), a dynamic factor model approach (in the

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

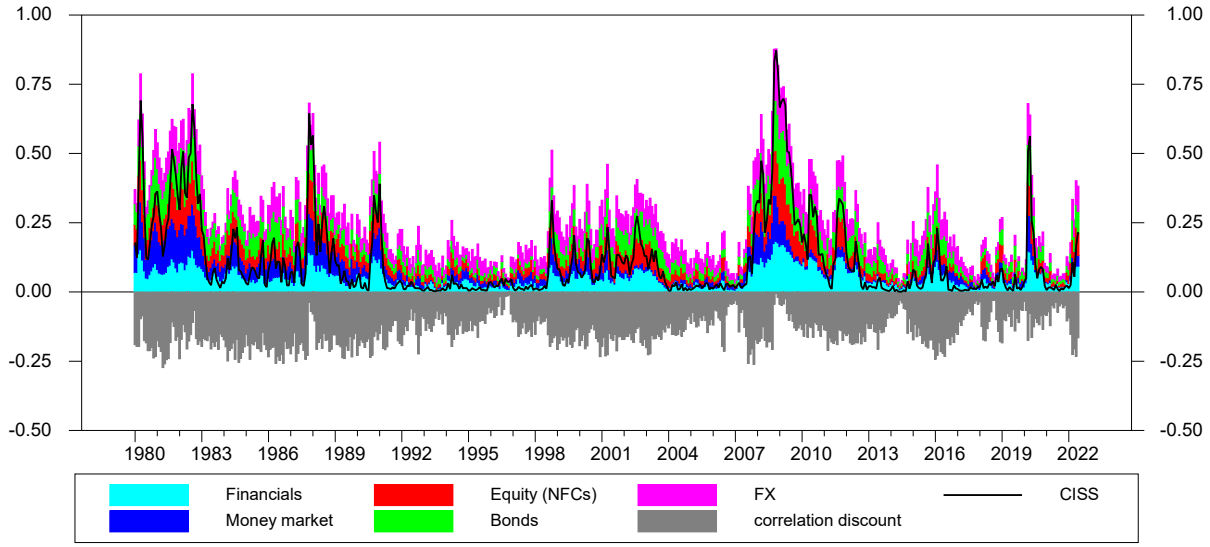

Fig. 5. The figure plots the decomposition of the US CISS according to Equation (9). The contributions from the individual stress factors are aggregated into money, bond, equity nonfinancial corporations, equity financials and foreign exchange market contributions in line with Table 3. The data are monthly averages of daily data from January 1980 to June 2022.

spirit of Brave and Butters (2012)), and the turbulence index (from Kritzman and Li (2010)).

Table 1 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). Financial authorities that regularly publish stress indices based on the respective design are listed in column 5, together with the main literature references. The simple-average FSIs for the two transformations (FSI-avg-cdf and FSI-avg-std) are trivial cases as they do not perform any systemic risk weighting, so the co-dependence is represented by an all-ones vector.[18] The two PCA-based indices (FSI-PCA-cdf and FSI-PCA-std) and the dynamic factor model (FSI-DF-std) follow the idea that the systemic dimension of financial stress can be captured by a latent factor that drives the observed correlation between the input indicators. The latent factor thus 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. Accordingly, the systemic risk weights of the input series in the composite indicator are accordingly computed from their loadings on the first principal component, i.e. 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 on the common factor can be interpreted as the systemic risk weights, as they implicitly reflect the extent to which each input series contributes to the dynamics of the estimated common factor.[19] The turbulence index measures the Mahalanobis distance between the standardised indicators

[18] Averaging over standardised indicators has also been labelled *variance-equal weighting*. If the FSI uses weights calibrated to reflect the relative size of the market segment, this could also be seen as a form of systemic risk weighting.

[19] The 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 using the Kalman filter. The measurement equation is $\tilde{x}_t = \gamma f_t + v_t$; the state equation is $f_t = \phi f_{t-1} + w_t$, where γ is the vector of loadings on 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 1: Different FSI designs as special cases of the general framework

FSI concept	$\mathcal{E}_t$	$\mathcal{C}_t$	Index formula	Published by; main references
CISS	$z_t z_t'$	R_t	$(w \circ z_t)' R_t (w \circ z_t)$	ECB, New York Fed; [1], [2]
FSI-average-cdf	z_t	ι_N	$w_t' z_t$	Cleveland Fed; [3]
FSI-PCA-cdf	z_t	e_1	${e_1}' z_t$	
FSI-average-std	$\tilde{x}_t$	ι_N	$w_t' \tilde{x}_t$	[4], [5], [6]
FSI-PCA-std	$\tilde{x}_t$	e_1	${e_1}' \tilde{x}_t$	Kansas City & St. Louis Fed; [7], [8]
FSI-DF-std	$\tilde{x}_t$	γ	$\tilde{x}_t = \gamma f_t + \hat{v}_t$	Chicago Fed, OFR; [9], [10], [11]
Turbulence	$\tilde{x}_t \tilde{x}_t'$	R^{-1}	$(w \circ \tilde{x}_t)' R^{-1} (w \circ \tilde{z}_t)$	[12]

Notes: The table describes different FSI designs in terms of different measures of the extremeness ($\mathcal{E}_t$) and co-dependence ($\mathcal{C}_t$) dimensions according to Definition 2 and Equation 1. For brevity, the index formulae are written in matrix notation: x_t denotes the N-dimensional vector of raw stress indicators, $\tilde{x}_t$ their z-score transformations and z_t their probability integral transforms; R denotes the correlation matrix of the 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 the stress factors; γ is the vector of estimated loadings of the stress factors on a single latent dynamic factor f_t, where $\hat{v}_t$ denotes the residuals from the measurement equation of the dynamic factor model; In this model, the dynamic factor f_t represents the FSI; w is the vector of equal weights with $w_i = 1/N$; w_t denotes a set of convex weights that deviate from equal weights and can also vary over time. Main references: [1] Hóllo et al. (2012); [2] Boyarchenko et al. (2022); [3] Oet et al. (2011); [4] Illing and Liu (2006); [5] Cardarelli et al. (2011); [6] Vermeulen et al. (2015); [7] Hakkio and Keeton (2009); [8] Kliesen and Smith (2010); [9] Brave and Butters (2012); [10] Monin (2017); [11] van Roye (2014); [12] Kritzman and Li (2010).

over time. The Mahalanobis distance is a multidimensional 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 behave together at each point in time, and it turns out that periods of 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 exercise, 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 only reflect differences in the way the components are transformed and aggregated. The resulting seven FSI series are plotted in the two panels of Figure 4.7. Both panels reveal, first, that the simple-average FSIs and the PCA-based FSIs are almost identical for both transformations, reflecting that the first eigenvector of the spectral decompositions is not very different from the vector of equal weights. This in turn suggests that, on average, all components contribute similarly 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 the dynamic factor model appears somewhat smoother than the PCA index, they still generally

overlap to a large extent. Third, in terms of cdf-transformed indicators, the CISS discriminates more strongly between systemic and non-systemic 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 close to being perfectly correlated. Fourth, the turbulence index does indeed mainly identify extreme events and thus acts similarly to a dummy variable. It shows some large jumps around known episodes of stress, while most of the time it hovers around low levels.

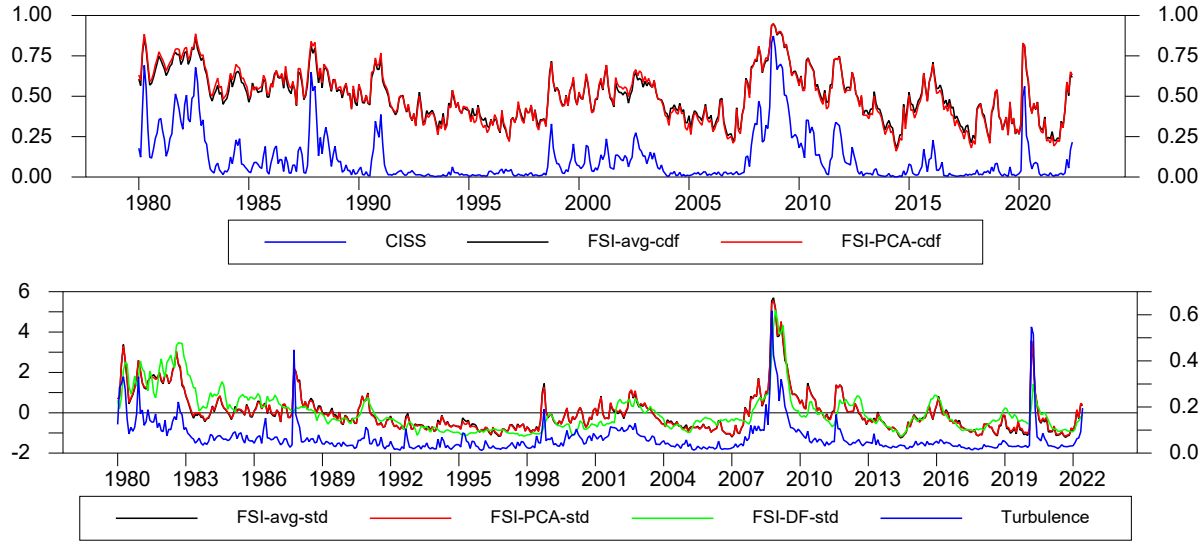

Fig. 6. The figure plots seven different FSIs designed according to Table 1 and applied to the 15 component series of the US CISS as described in Table 3. The first three series in the bottom panel are standardised for graphical purposes. The series are monthly averages of daily data from January 1980 to June 2022.

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 a bootstrap algorithm to derive test statistics to assess this hypothesis. Let the realisations of the stress index $\{S_t : t = 1, ..., T\}$ be computed as the CISS according to Equation 6. Our goal is to make statistical statements about whether a particular realisation S_t can be considered unusually high or low ("normal"). For this purpose we formulate the following null hypothesis for any realisation S_t:

$$H_0 : S_t \in \Upsilon^{low} \tag{10}$$

where $\Upsilon^{low} = \{\tilde{S}_t : t = \tilde{T}_0, ..., \tilde{T}_1\}$ denotes a subsample of the data representing a low stress period, and where $\tilde{T}_0$ and $\tilde{T}_1$ are the start and end dates of this period, respectively. If $\Upsilon^{full} = \{S_t : t = 1, ..., T\}$ denotes the full data sample, then $\Upsilon^{low} \subset \Upsilon^{full}$ and $\tilde{T}_1 - \tilde{T}_0 < T - 1$. To test the null hypothesis, we derive a critical stress value $\gamma_{1-\alpha}$ at some statistical significance level α, implicitly defined in

$$\Pr\left(\tilde{S}_t > \gamma_{1-\alpha} | \Upsilon^{low}\right) \leq \alpha. \tag{11}$$

If the data generating process of $\tilde{S}_t$ were known, Equation (11) 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 non-parametrically by simulation. We propose a bootstrap-based sampling procedure that first approximates the empirical distribution of S_t under no-stress conditions, i.e. under the null hypothesis, and then computes a critical value for the null hypothesis as the $1-\alpha$ percentile of this approximated 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 can be interpreted as a test of the similarity of the two underlying cross-product matrices. However, our procedure benefits from the assumption that some observations of S_t come from a low-stress environment and are thus generated under the null hypothesis. We consider the subperiod from 1 January 1992 ($\tilde{T}_0$) to 29 December 2006 ($\tilde{T}_1$) to represent normal stress conditions in the US financial system.

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

$$\left(z_t z_t'\right)_{i,j}(R_t)_{i,j} = \tilde{S}_t + \psi_{i,j,t}, \quad i,j = 1,\ldots,N, \quad t = \tilde{T}_0,\ldots,\tilde{T}_1, \tag{12}$$

where $\psi_{i,j,t}$ has a zero mean by construction. We now assume that in a low-stress environment - where extremeness and correlations are on average rather low - 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 $\gamma_{1-\alpha}$:[20]

Algorithm 1. *Let* $t = \tilde{T}_0, ..., \tilde{T}_1$ *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, $n = 1,\ldots,N_b$ *times, and construct* N_b *simulated values of array* $\{\psi^{(n)}_{i,j,t}, 1 \leq i,j \leq N\}$;

(iii) *Compute* N_b *simulated values of each cross-product term as* $\left((z_t z_t')_{i,j}(R_t)_{i,j}\right)^{(n)} = \tilde{S}_t + \psi^{(n)}_{i,j,t}$;

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

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

[20] By a loose interpretation, this would mean that there is only a negligible difference between realisations $\psi_{i,j,t}$ across $i,j = 1,\ldots,N$.

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

The algorithm is applied to US daily data based on the entire low-stress subsample, and the critical value is computed for a significance level of $\alpha = 0.01$ (1%). Figure 7 plots the US CISS together with a red horizontal line representing the estimated critical value of 0.123.[21]

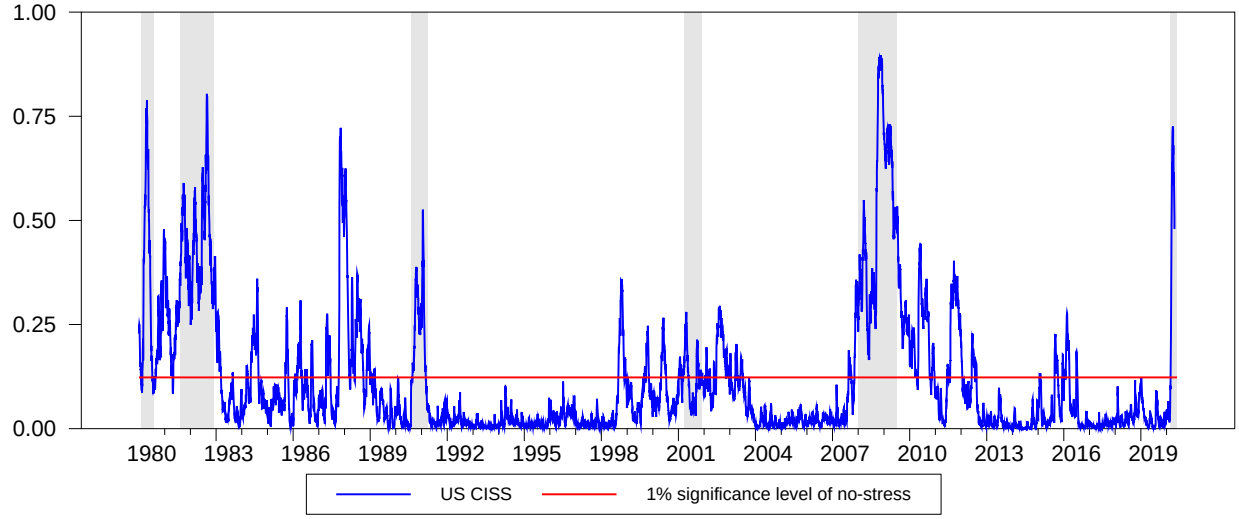

Fig. 7. US CISS and high-stress threshold. The threshold value (0.123) represents the critical value of the null hypothesis of low or "non-systemic" stress at the 1% significance level. The threshold is generated using the sampling algorithm 1. The shaded areas are NBER recessions for the United States (peak to trough). Data are daily from 2 January 1980 to 31 May 2020. Sources: Federal Reserve Bank of St. Louis for the NBER recession indicator.

A simple illustration can provide some preliminary evidence of the economic relevance of such a threshold estimate. We use the 1%-critical value to split the sample into low-stress and high-stress CISS observations (monthly averages of daily data) and plot, separately for each regime, the US CISS against annual real GDP growth led by two months (Figure 8). The red dots indicate joint observations from the low-stress regime and the blue dots from the high-stress regime. The figure suggests a negative relationship between the CISS and growth in the high-stress regime, at least when growth is negative. The red dots, on the other hand, appear to be a pure random cloud. Such a nonlinear relationship is consistent with Hóllo et al. (2012) who estimate a crisis threshold for the CISS within a bivariate threshold-VAR model with the euro area CISS and the annual growth rate of industrial production as endogenous variables. That paper indeed finds statistical support for a single threshold, and it turns out that shocks to the CISS have strong negative effects on economic activity only in the high-stress regime. The next section examines the real effects of systemic stress in more detail.

6. Downside risks to growth from financial stress

This section investigates the real effects of systemic financial stress, with a particular focus on potential asymmetries in these effects. Recent empirical studies typically find strong short-term predictive power of financial stress or financial conditions indicators in the lower tails of

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

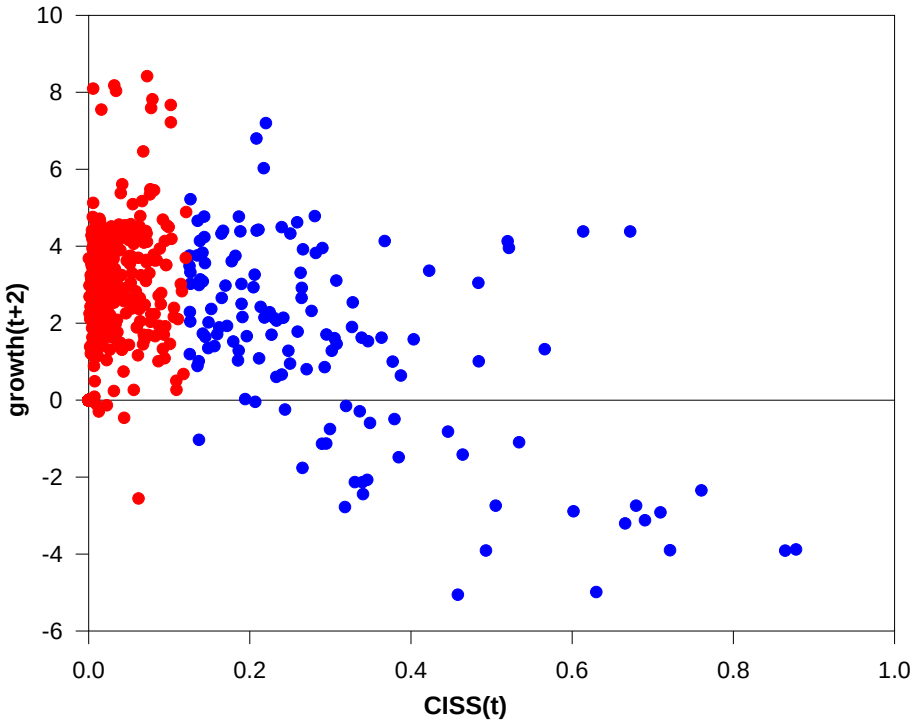

Fig. 8. Scatter plot of CISS and annual real GDP growth (in %, led by two months) in the US. CISS is a monthly average of daily data, and real GDP is interpolated to monthly from quarterly data as described in Appendix C. Blue dots represent periods when the CISS is above the 1% critical value from the sample algorithm 1 (high stress), and red dots symbolise periods when the CISS is below this threshold (low stress). Data are monthly from January 1980 to March 2020.

the conditional output growth distribution, while the effects are much weaker in the central and upper parts of the growth distribution. Such asymmetric responses of economic activity to financial distress may reflect several mechanisms put forward in the recent theoretical macro-finance literature. For example, the effects of sharply rising financial frictions and increasing uncertainty may be amplified by asset fire sales and credit constraints becoming binding (e.g., Bianchi (2011), He and Krishnamurthy (2012), Lorenzoni (2008) and Mendoza (2010)). Indeed, the popular growth-at-risk literature relies on this macro-financial asymmetry as a stylised fact.

We first run and compare single-equation predictive quantile regressions of real GDP growth on the CISS, the alternative FSI designs presented in Section 4.7, and some standard financial business cycle predictors from the literature. The purpose of such a horse race is to assess the potential value added of the CISS and its particular design features in forecasting, in-sample, the short-term downside risks to the economy. We then assess the dynamic effects of the CISS on measures of economic activity based on a parsimoniously specified structural quantile VAR (QVAR) as proposed in Chavleishvili and Manganelli (2019). We use simulations to show how financial stress contributed to the major recessions in two recent economic crises, the GFC and the COVID-19 pandemic.

6.1. Direct growth forecasting with quantile regressions

We first employ the direct multistep, single-equation forecasting approach applied in the seminal paper by Adrian et al. (2020). In our case, we regress future real GDP growth on current values of financial indicators over expanding cumulative growth horizons from 1 month to 12 months ahead. We do not consider horizons further ahead, as we do not expect financial stress to have longer-term cyclical effects on the economy. The models are estimated as quantile regressions to allow for nonlinear relationships in the conditional joint distribution of financial stress and economic growth.

In the first set of regressions, the Purchasing Managers' Index (PMI) is included as a forecast variable, along with the CISS and lagged real GDP growth, to control for general business conditions at the forecast origin. Indeed, Jarocinski and Mackowiak (2017) finds strong predictive power of the PMI for macroeconomic activity variables in both the euro area and the US, and figure 15 in Appendix C illustrates the generally strong comovement between annual real GDP growth and the PMI over the euro area business cycle. The following quantile regression equation is estimated for monthly data for the euro area:[22]

$$\Delta y_{t+h} = \beta_0^\theta + \beta_1^\theta CISS_t + \beta_2^\theta PMI_t + \beta_3^\theta \Delta y_{t-1} + \epsilon_{t+h}^\theta, \tag{13}$$

where Δy_{t+h} denotes the annualised real GDP growth from month t to month $t+h$ and superscript θ indicates the regression quantiles $\theta = 5\%, 10\%, ..., 95\%$; ϵ_{t+h}^θ is the forecast error. Representative results for the 3-month forecast horizon ($h = 3$) are graphically summarised in Figure 9.[23]

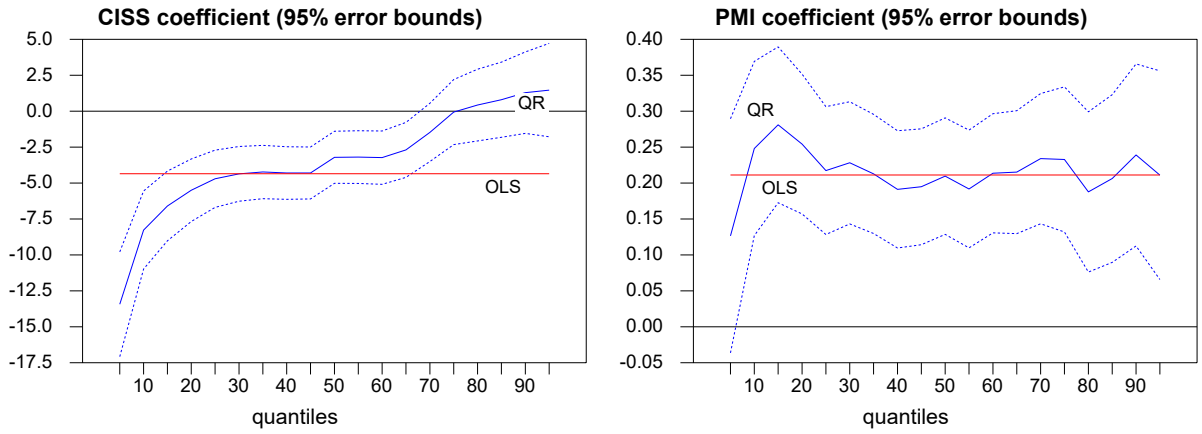

Fig. 9. Predictive quantile regression coefficients for the euro area. This figure plots the estimated quantile-specific coefficients β_1^θ on the CISS (left-hand panel) and β_2^θ on the PMI (right-hand panel) of the quantile regression (13). The variable to be predicted is real GDP growth 3 months ahead ($h = 3$). The quantile regression coefficients are plotted as solid blue lines with 95% error bounds as blue dashed lines for each quantile $\theta = 5\%, 10\%, ..., 95\%$. For comparison, the OLS coefficients are plotted as red horizontal lines. Regressions are estimated using monthly euro area data from 1998M7 to 2019M12.

Two main facts stand out: First, the predictive coefficient of the CISS, plotted from the

[22] All data are described in Appendix C.

[23] Results for the US case are very similar to those for the euro area and are available on request.

lowest (5%) to the highest (95%) regression quantile, shows a pronounced asymmetric pattern. In the lower tail of the growth distribution, the CISS has strong predictive power, with a negative coefficient of around −13 at the 5% quantile. The predictive coefficient then gradually increases to exceed the OLS estimate of around −5 in the middle part of the distribution; in the upper tail, the coefficient becomes zero and statistically insignificant. Second, the quantile regression coefficients on the PMI are stable throughout the growth distribution, close to the OLS estimate of around 0.2. Thus, the relationship between the PMI and future real GDP growth is essentially linear. These general results confirm our expectation that systemic stress is important for economic growth mainly in bad states of the economy. Capturing this asymmetry may be important because linear empirical approaches would tend to underestimate the effects of financial stress on short-term economic growth during financial crisis episodes and overestimate such effects in normal economic times.

In a second exercise, we apply this predictive framework to assess the lower-tail predictive power of the CISS vis-à-vis alternative FSI designs (the special cases from Section 4.7) and other well-known financial business-cycle indicators.[24] Predictive power is assessed in terms of tick loss, which is defined as the asymmetrically-weighted sample average of prediction errors:

$$\mathcal{T}_\theta = \frac{1}{T}\sum_{t=1}^{T} \epsilon_{t+3}^{\theta}(\theta - \mathcal{I}(\epsilon_{t+3}^{\theta} < 0)). \tag{14}$$

The lower the tick loss, the better the forecasting performance. For each indicator and each quantile $\theta \in \{10\%, 20\%, 50\%, 90\%\}$, we compute the percentage gain in tick loss ($\mathcal{T}_\theta$-gain) over the tick loss of a simple quantile autoregression, averaged over all forecast horizons from 1 to 12 months.[25] We report the average tick loss gain along with the average rank of an indicator in all horizon-specific performance rankings (with the first rank given to the indicator with the highest tick loss gain).

Table 2 presents the results of the comparison between the CISS and the six alternative FSI designs based on data for the euro area (panel a) and the US (panel b). In general, the CISS is found to predict short-term real GDP growth particularly well at the lower quantiles. In the euro area, the predictive power of the CISS is second only to the FSI based on the dynamic factor model (FSI-DF-std). The average tick loss gain of the CISS is about 22% compared to 34% for the dynamic factor model. The predictive power of the CISS remains relatively strong at the 20% quantile, but progressively deteriorates at the higher quantiles. In the US case, the CISS is even the best predictor at the 10% and 20% quantiles. The tick loss gain of about 13% at the 10% quantile is lower than for the euro area, probably reflecting the larger data sample in the US case (see the notes section of table 2). However, the dynamic factor model performs relatively poorly when applied to US data. Looking at the design characteristics of the FSIs, it is noticeable that the predictive performance of the CISS is generally quite similar to that of the FSIs based on z-score transformed components (FSI-average-std, FSI-PCA-std and FSI-DF-std). In contrast, the other FSIs based on cdf-transformed components (which remove

[24] We drop the PMI from the list of regressors without much impact on the quantitative and qualitative results.

[25] The tick loss gain for each FSI is computed as $100 \cdot (\mathcal{T}_\theta^{AR} - \mathcal{T}_\theta^{FSI})/\mathcal{T}_\theta^{AR}$, where $\mathcal{T}_\theta^{FSI}$ is the FSI tick loss and $\mathcal{T}_\theta^{AR}$ is the autoregression tick loss.

the cardinal scale of the input series) perform much worse at the lower quantiles. Thus, in terms of short-term growth-at-risk forecasting properties, the time-varying correlation weighting embedded in the CISS design seems to compensate for the loss in forecasting performance when using cdf-transformed index components.

Panel c of Table 2 compares the CISS with alternative financial indicators based on US data.[26] The CISS again emerges as the best predictor in the lower tails of the growth distribution (10% and 20%). However, the differences in the tick loss gain with most other indicators are not very large. The closest competitors are the term spread of the government bond yield curve and the Gilchrist-Zakrajšek corporate bond spread net of the excess bond premium (GZ-netspread). The NFCI performs worst in the lower tails, but becomes the best financial predictor in the upper tail (90%). Nevertheless, and as a general pattern observed in all three exercises, none of the financial indicators improves the predictive power of a simple autoregression in the upper tail by more than a mere 7% at best. The overall results of these horse-race exercises are robust to the addition of more lags (for all forecast variables) and to the inclusion of the PMI as a control variable. However, when the 1973-1979 period is added to the sample of US data underlying the third exercise, the predictive power of the NFCI rises very sharply, surpassing that of the CISS and all other financial indicators. The ranking of the CISS relative to the other indicators remains basically stable in the larger data sample.

6.2. Growth forecasts from a quantile VAR

In this section we use a quantile VAR to examine the dynamic interlinkages between systemic financial stress and the real economy. The analysis uses euro area data, but US data yield very similar results.[27] The QVAR includes three endogenous variables measured at monthly frequency: the CISS, the PMI Composite Output Index, and the year-on-year change in log real GDP. The series are described in Appendix C. 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 for a given month is published as a flash estimate towards the end of that month and as a final estimate at the beginning of the following month. Against this background, we consider the PMI to be a timely indicator of business activity and sentiment that should help us to better identify shocks to real GDP growth, given the rather long publication lags of real GDP (and industrial production, the variable that informs the interpolation of quarterly real GDP into monthly observations). Year-on-year real GDP growth can be seen as a measure of the cyclical component of real GDP (see Hamilton (2018)). Its time series properties fit well with those of the PMI and the CISS, which appear to follow stationary, though rather persistent stochastic processes. The QVAR is estimated with two lags over the period October 1998 to February 2020.[28] As proposed in Chavleishvili and Manganelli (2019) for QVARs, we identify the structural shocks to the model

[26] For a very similar exercise using euro area data, see Figueres and Jarociński (2020). In that paper, the CISS emerges as the best growth-at-risk predictor compared with several corporate bond spreads, the TED spread, a bank lending rate spread, the term spread, an average sovereign bond yield spread, stock market volatility (VSTOXX) and the first principal component of all these financial indicators excluding the CISS.

[27] Results are available on request.

[28] Standard information criteria uniformly suggest an optimal lag order of two in a linear VAR with the same model variables.

Table 2: Relative performance of different financial indicators in quantile growth predictions

	$\theta = 10\%$		$\theta = 20\%$		$\theta = 50\%$		$\theta = 90\%$	
Predictor X_t	$\mathcal{T}_\theta$-gain	rank	$\mathcal{T}_\theta$-gain	rank	$\mathcal{T}_\theta$-gain	rank	$\mathcal{T}_\theta$-gain	rank
a) Euro Area - Special cases								
CISS	**21.74**	3.25	16.01	3.50	4.85	5.50	2.89	5.33
FSI-average-cdf	15.49	4.92	14.48	4.50	5.76	4.17	4.03	2.92
FSI-PCA-cdf	15.05	5.75	13.81	5.75	5.42	5.08	4.07	**2.50**
FSI-average-std	20.96	3.83	**16.64**	**3.25**	**7.64**	**1.75**	**4.24**	**2.17**
FSI-PCA-std	21.45	**3.17**	16.56	3.67	7.37	2.92	**4.15**	3.08
FSI-DF-std	**34.31**	**1.00**	**23.97**	**1.00**	**7.67**	**1.58**	2.93	5.25
Turbulence	11.72	6.08	8.87	6.33	2.59	7.00	0.26	6.75
b) US - Special cases								
CISS	**12.79**	**1.17**	**7.26**	**2.08**	**1.73**	3.08	**0.33**	**3.17**
FSI-average-cdf	10.89	3.83	6.03	4.42	1.07	3.92	0.06	5.08
FSI-PCA-cdf	10.71	4.92	5.58	5.42	0.69	6.67	0.10	3.92
FSI-average-std	11.67	4.08	7.14	2.42	**1.81**	**1.83**	0.09	3.83
FSI-PCA-std	**11.79**	**3.08**	**7.16**	**2.17**	1.72	**3.00**	0.06	5.75
FSI-DF-std	6.46	6.92	4.09	6.83	0.80	6.17	**1.91**	**1.75**
Turbulence	11.68	4.00	5.81	4.67	1.71	3.33	0.09	4.50
c) *US - Standard financial indicators*								
CISS	**12.79**	**1.17**	**7.26**	**2.00**	**1.73**	3.00	**0.33**	**3.08**
EBP	10.74	3.83	5.95	4.33	1.05	3.92	0.06	5.25
GZ-spread	10.52	4.92	5.47	5.42	0.67	6.58	0.10	3.75
GZ-netspread	11.51	4.00	7.04	2.58	**1.78**	**1.83**	0.09	4.00
Term spread	**11.63**	**2.92**	**7.07**	**2.17**	1.69	**2.92**	0.06	5.67
NFCI	6.34	6.92	4.02	6.83	0.78	6.17	**1.94**	**1.75**
SP500-DY	11.33	4.25	5.65	4.67	1.67	3.58	0.09	**4.50**

Notes: The numbers in this table are the results of quantile regressions $\Delta y_{t+h} = \beta_0^\theta + \beta_1^\theta X_t + \beta_2^\theta \Delta y_{t-1} + \epsilon_{t+3}^\theta$, where Δy_{t+h} is the annualised real GDP growth from month t to month $t+h$, forecast horizon $h = 1, ..., 12$, and θ the regression quantiles. $\mathcal{T}_\theta$-gain denotes the average tick loss (see Equation (14)) gain over all h regressions against the tick loss from a simple quantile autoregression without any X_t. Positive numbers indicate lower losses relative to the autoregression. Columns labelled "rank" show the average rank of each predictor over all 12 prediction horizons. EBP and GZ-spread are the excess bond premium and the corporate bond spread from Gilchrist and Zakrajšek (2012); GZ-netspread is the GZ-spread minus the EBP; term spread is the yield differential between 10-year and 3-month government bonds; NFCI is the National Financial Conditions Index from Brave and Butters (2012); SP500-DY is the dividend yield on the S&P 500 stock price index. Effective regression samples of monthly data: 1999M1-2019M12 for the euro area and 1980M1-2019M12 for the US.
Sources: ECB, FRED, Haver, Board of Governors of the Federal Reserve System.

variables using the recursive Cholesky factorisation. The CISS is ordered first in the QVAR, which implies that it does not respond to simultaneous shocks to the PMI and real GDP growth. This assumption can be justified from an information delay perspective, as argued in Inoue et al. (2009) and Hóllo et al. (2012). Since the PMI and real GDP are published with a lag, their realisations in a given month are not directly observable by financial market participants. We therefore assume that the CISS cannot react to contemporaneous innovations in the PMI and real GDP growth. Similarly, we order the PMI ahead of real GDP growth, so that growth can react to contemporaneous shocks in the PMI, but not vice versa. Nevertheless, the qualitative results of our empirical exercises remain robust to a reverse ordering of the real variables and the CISS.

Our structural QVAR can be sketched as follows:[29]

$$y_t = \phi^\theta + \sum_{l=1}^{2} \Phi_l^\theta y_{t-j} + \Gamma^\theta y_t + \varepsilon_t^\theta, \tag{15}$$

where $y_t \in \mathbb{R}^3$ is a vector collecting the endogenous variables in the order CISS, PMI and real GDP growth; ϕ^θ denotes the vector of intercepts for a given quantile $\theta \in (0,1)$; Φ_l^θ are the two quantile-specific 3×3 matrices containing the lag coefficients; and Γ^θ is 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 $\mathcal{F}_{1,t} = \left(y'_{t-1}, y'_{t-2}\right)'$ and $\mathcal{F}_{i,t} = \left(\mathcal{F}'_{i-1,t}, y_{i-1,t}\right)'$ for $i = 2,3$ be the recursive information set, then the conditional quantile functions can be used to explain the relationships between the variables across the entire conditional distribution and can be obtained by solving the following conditional quantile restrictions:

$$\Pr(y_{i,t} < Q_{y_{i,t}}(\theta|\mathcal{F}_{i,t}) = \theta, \quad i = 1,2,3, \tag{16}$$

with

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

denoting the conditional quantile functions identified through restrictions (16).

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

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

where $\beta_i(\theta) = \left(\left(\phi^\theta\right)_i, \left(\Phi_1^\theta\right)_{i,.}, \ldots, \left(\Phi_p^\theta\right)_{i,.}, \left(\Gamma^\theta\right)_{i,.}\right)'$ is 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 and $I(u < 0)$ the indicator function.

All direct and indirect, lagged and contemporaneous, state-dependent interactions between the model variables can be summarised in terms of quantile impulse response functions (QIRFs). We compute the QIRFs using the simulation algorithm in Chavleishvili and Manganelli (2019). The idea of a QIRF is to explore whether a particular part of the conditional distribution of

[29] For a detailed exposition see Chavleishvili and Manganelli (2019).

a variable of interest responds differently from other parts to an unpredicted change in one of the model variables, as opposed to an IRF from a linear VAR which only estimates conditional mean effects. For a horizon $h \geq 1$, a QIRF is defined as

$$Q_{y_{t+h}}(\theta|\mathcal{F}_{t+1}, \delta_i) - Q_{y_{t+h}}(\theta|\mathcal{F}_{t+1}), \quad \theta \in (0,1)^3, \tag{17}$$

where δ_i denotes the value at time $t+1$ value of the variable $i \in (1,2,3)$, $y_{i,t+1}$, after being hit by a certain structural shock. It should be noted that θ is now defined as a three-dimensional vector with elements $\theta_i \in (0,1)$, which means that the QIRF can be computed for cases where the model variables follow paths along different quantiles of their distributions.[30]

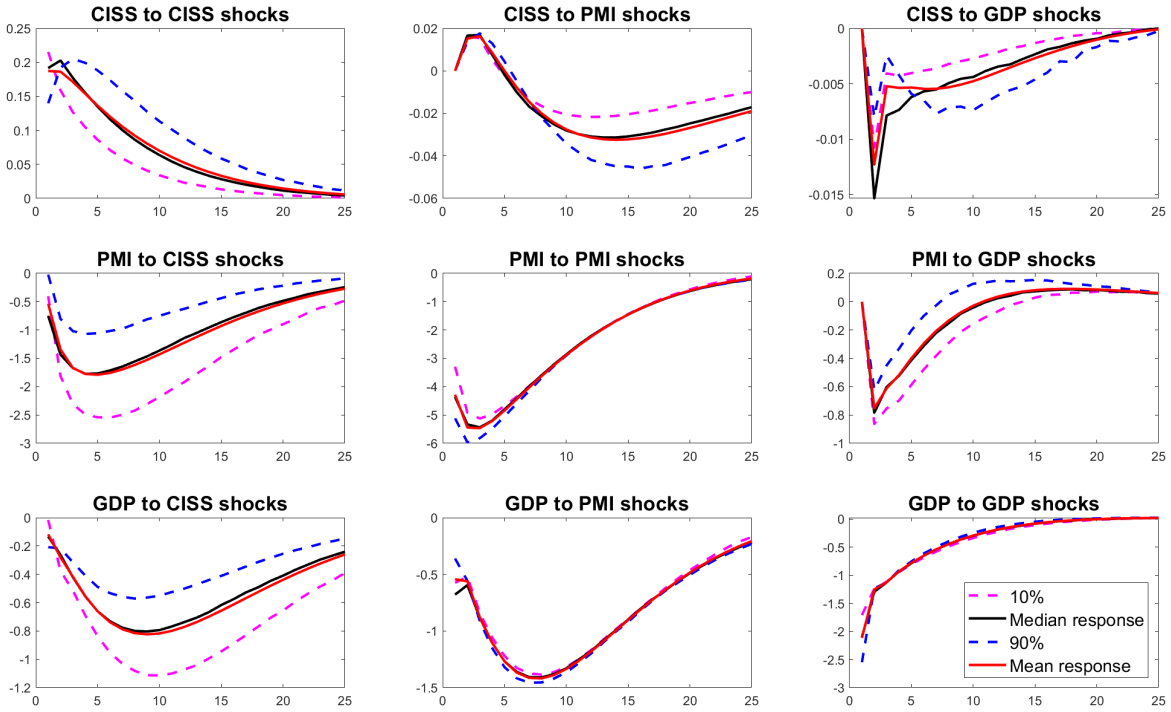

Fig. 10. The figure plots structural quantile impulse response functions (QIRFs) for the euro area QVAR. The QIRFs are shown for adverse shocks in each variable, with the absolute size being one standard deviation of the median shocks. An adverse shock to the CISS is an increase, while an adverse shock to the PMI and real GDP growth is a decrease. The responses of the CISS, the PMI and real GDP growth are displayed in the first, second and third rows respectively. Correspondingly, the responses to shocks to the CISS, the PMI and real GDP growth are shown in the first, second and third columns respectively. Each panel shows the QIRFs for the 10%, 50% and 90% quantiles of the conditional distribution of the response variable.

Figure 10 shows QIRFs from simulations in which each variable, in turn, experiences an adverse shock equal to one standard deviation at the median of its unconditional empirical distribution. The respective post-shock values δ_i at time $t+1$ for the CISS, the PMI and real GDP growth are set accordingly to $\delta = (0.2196, 47.3143, -0.9631)$. The QIRFs are calculated for horizons $h = 1, \ldots, 24$ months and for three different quantiles $\theta \in (0.1, 0.5, 0.9)$, i.e., 10%,

[30] This property of conditional quantile forecasting makes a QVAR particularly suitable for macro stress testing exercises as demonstrated in Chavleishvili and Manganelli (2019).

50% and 90%.[31] The responses of a variable to the three different shocks are shown row by row, so that the responses of each of the three variables to a particular shock are shown column by column.

The QIRFs in the first column and row, all associated with shocks to or responses of the CISS, display some distinct asymmetries. Most interesting in our context are the much stronger effects of CISS shocks on the PMI and real GDP growth at the 10% quantile compared to the 90% quantile and the median responses. For example, a CISS shock of +0.18 lowers annual real GDP growth by a maximum of 1.1% over a roughly one-year horizon at the 10% quantile of the growth distribution, compared with −0.8% and −0.5% for the median and the 90% quantile, respectively. In addition, as can be seen from figure 11, the growth responses to a CISS shock at the 10% quantile are statistically different from the responses at the median and the 90% quantile at horizons around the peak effects, and responses of all three quantiles are statistically different from zero. The differences across quantiles are similarly large for the PMI responses to the same CISS shock, with the PMI falling by 2.5 points at the 90% quantile, 1.7 points at the median and 1.0 point at the 10% quantile. The responses of the CISS to own shocks show some temporary overshooting at the 90% quantile. This reflects a local non-stationarity 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 variables of economic activity reveal a linear behaviour. Accordingly, shocks to the PMI lead to similar changes in real GDP growth across all parts of the conditional growth distribution.

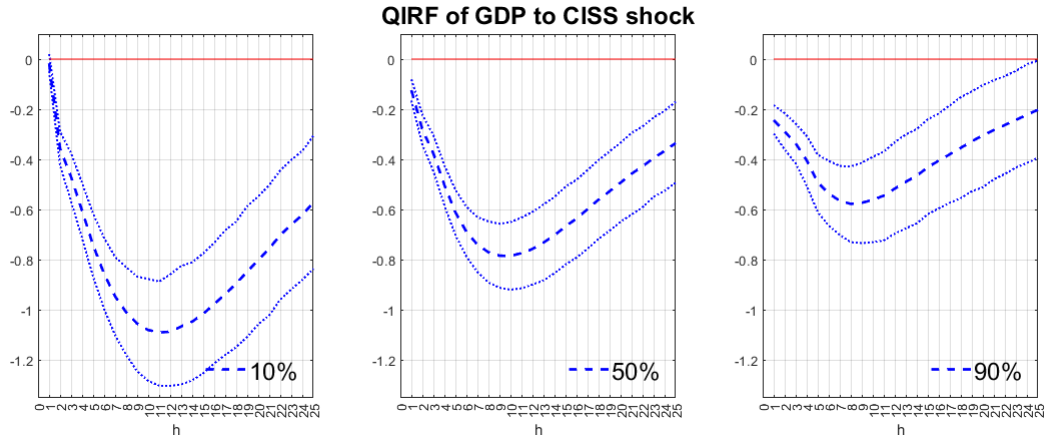

Fig. 11. The figure plots structural quantile impulse response functions (QIRFs) for euro area annual real GDP growth to a one standard deviation shock in the CISS for the 10% quantile (left panel), the 50% quantile (middle panel) and the 90% quantile (right panel). The QIRFs are plotted with 95% confidence intervals computed using the subsampling algorithm described in Chavleishvili and Mönch (2023).

We now use the QVAR to assess how well the CISS performs as an indicator of systemic financial crises. To this end, we run counterfactual scenarios covering two severe economic crisis

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

episodes, the 2008/09 GFC and the COVID-19 pandemic. In both episodes, economic activity collapsed with unprecedented severity and speed. However, conventional wisdom would suggest that financial stress was a key driver of the economic downturn only in the GFC and not in the pandemic.

The counterfactual scenarios are designed as conditional quantile forecasts of real GDP growth using the sampling approach suggested in Chavleishvili and Manganelli (2019). The forecast origin is set to August 2008 for the GFC exercise and to December 2019 for the simulation covering the COVID-19 crisis, and the forecast horizon is 24 months. Three different scenarios are calculated for each period. The first is an unconditional projection, which assumes that the system is not exposed to any structural shock over the forecast horizon. Accordingly, this scenario assumes that the two crises never happened. In figures 12 (GFC) and 13 (COVID-19) we show the median unconditional forecast for real GDP growth as a dashed turquoise line. In the second exercise, we restrict the CISS to follow its realised path over the projection horizon, implicitly assuming a sequence of structural shocks to the CISS quantiles that ensures the restriction is satisfied.[32] For this scenario, we plot a fan chart covering conditional quantiles from 1% to 99%. Comparing these forecast distributions conditional on a fixed CISS path with the unconditional median forecast on the one hand and realised growth on the other, provides an auxiliary means to quantify the contribution of CISS shocks to the two crisis episodes.[33] In the third scenario, both the CISS and the PMI are fixed at their realised paths over the projection horizon. The median conditional growth projection from this scenario is shown as a red dashed line in the figures below. Comparing the median projections from the second and third scenarios (the black and red dashed lines) provides a rough estimate of whether shocks to either financial stress or business sentiment are mainly responsible for the predictable part of the decline in real GDP growth.

In the context of the model, the GFC began with two large increases in the CISS of +0.17 and +0.33 in September and October 2008, respectively. The size of these increases was about three and five times larger, respectively, than the sample standard deviation of monthly changes in the CISS. In November of that year, the CISS climbed to its all-time high of 0.91, hovered around similar levels for several months, before gradually returning to its initial level by the end of the forecast horizon. Accordingly, in the second scenario with the realised CISS as the only forecast condition, the QVAR predicts unprecedented downside risks to real GDP growth in the near term (see the fan chart in figure 12). For example, in this scenario, the median conditional forecast projects a fall in real GDP of 3.4% by May 2009, with the 10% quantile predicting a "growth at risk" of -5.5%. Actual real GDP growth in May 2009 was around -5%, in the middle of the 10% and 25% forecast quantiles. As the median unconditional growth forecast is essentially flat over the first 12 months of the forecast horizon, the persistent rise

[32] Chavleishvili et al. (2023) describes in detail how conditional quantile forecasts with fixed conditioning paths are implemented in the QVAR framework of Chavleishvili and Manganelli (2019).

[33] We resort to counterfactual simulations to quantify the contribution of financial stress to the two economic crises, as historical decompositions for QVARs are not yet available. Chavleishvili et al. (2023) suggests a Shapley-value decomposition that quantifies the contribution of each QVAR variable to the projected path of any variable of interest, but only locally for a given quantile and not globally for the observed path of a variable. Historical decompositions and other tools for counterfactual analysis in linear VARs are discussed in Kilian and Lütkepohl (2017).

in the CISS over this period could be seen as accounting for around 3 percentage points lower real GDP growth. Slightly larger effects emerge when comparing the lower quantiles of the unconditional forecast distribution with those of the fixed CISS scenario. The magnitude of these estimated output costs of financial stress is very large compared to standard downturns in the euro area business cycle. In the third scenario, i.e. when the forecasts are also conditioned on the realised path of the PMI, the median real GDP growth around the peak of the GFC is predicted to be about 1 percentage point lower than in the second scenario with the CISS as the only conditioning variable. The differences between the two scenarios become even smaller at the lower quantiles of the forecast distributions. Overall, it is probably fair to say that financial stress is likely to have played a major role in the collapse of the euro area economy during the GFC.

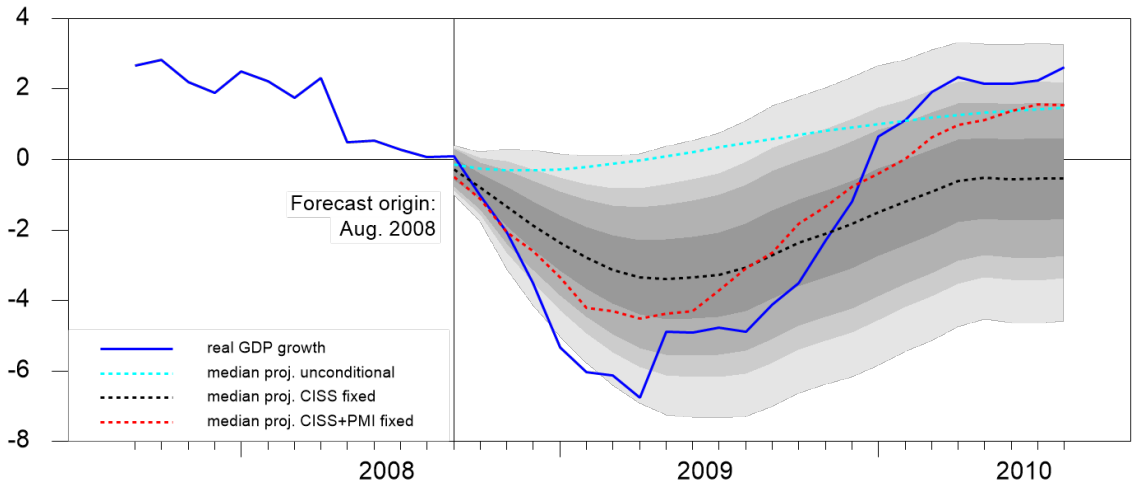

Fig. 12. The figure plots the monthly QVAR density forecasts of annual real GDP growth over a 2-year horizon covering the Great Financial Crisis of 2008/09. The forecast origin is August 2008. The dashed turquoise line represents the median of the unconditional forecast. The fan chart shows the forecast distribution conditional on the realised path of the CISS over the forecast horizon. The different grey shaded areas represent the ranges between the 1%, 5%, 10%, 25%, 75%, 90%, 95% and 99% quantiles. The dashed red line marks the median of the conditional projection with fixed paths of the CISS and the PMI as forecast conditions.

In contrast to the GFC, which was preceded by already elevated financial stress and declining economic growth associated with the spreading subprime mortgage market turmoil, the COVID-19 crisis hit the global economy unprepared. The arrival of the pandemic brought the world economy to a virtual standstill, and the pervasive uncertainty caused severe immediate stress in the financial system. The latter is captured in our QVAR by a sharp historical one-off rise in the CISS of 0.48 in March 2020, which is eight times the sample standard deviation of monthly CISS changes. However, this huge shock was largely reversed in the following months. This rapid reversal has been attributed in part to the improved resilience of the banking system, thanks to the regulatory and institutional overhaul following the GFC, as well as forceful policy interventions by fiscal, monetary and supervisory authorities at the national and international level (Giese and Haldane (2020)). As a result of this short-lived episode of

financial turmoil, growth-at-risk in the fixed CISS path scenario was never lower than around -2%, even at the most pessimistic 1% quantile over shorter horizons (see figure 13). At the median, the model predicts a maximum loss in output growth of about one percentage point relative to the median unconditional forecast. The picture changes radically when PMI shocks are added to the scenario. In March and April 2020, the PMI falls sharply, by -22 and -16 points respectively, which is eight and six times the historical standard deviation of monthly PMI changes. However, even these "horror" shocks to business sentiment are largely reversed in the subsequent two months and thus only temporary (see figure 15 in Appendix C). Conditioning the forecast scenario on fixed paths for both the CISS and the PMI, the QVAR predicts that real GDP falls by about 4% over the near term, which is around five percentage points lower than the unconditional projection. While these numbers are large by any standard, the estimated QVAR is still unable to capture the actual, very unusual dynamics of real GDP during the initial period of the pandemic. According to the interpolated data, real GDP in April 2020 was one fifth lower than a year earlier (see figure 13). Due to base effects in the annual growth rates, the rapid recovery of economic activity becomes visible in April 2021, with a reported growth rate of +18% year-on-year. Overall, the COVID-19 crisis simulations suggest that financial stress played only a minor role in causing the temporary collapse of the euro area economy.

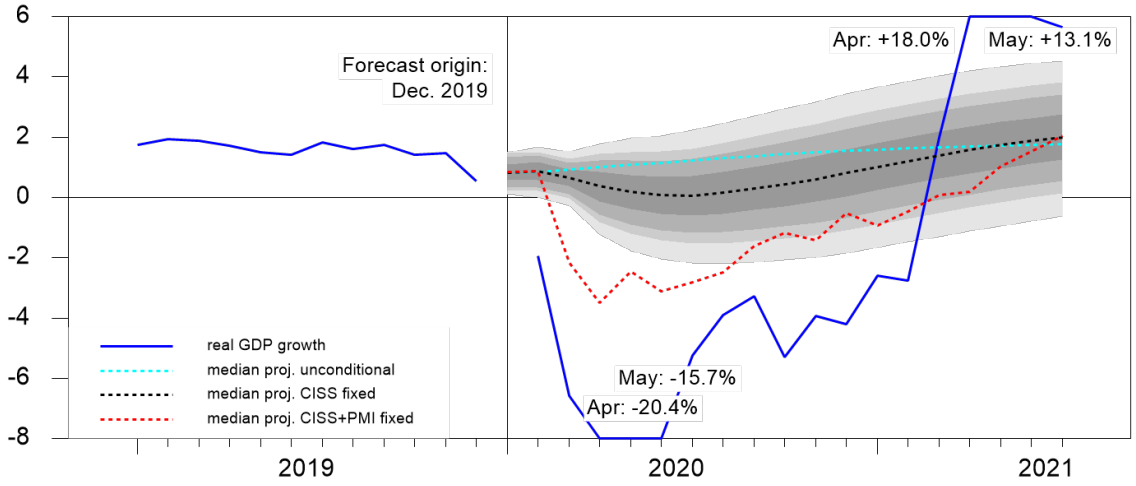

Fig. 13. The figure plots the monthly QVAR density forecasts of annual real GDP growth over a 2-year horizon covering the COVID-19 crisis. The forecast origin is December 2019. The dashed turquoise line represents the median of the unconditional forecast. The fan chart shows the forecast distribution conditional on the realised path of the CISS over the forecast horizon. The different grey shaded areas represent the ranges between the 1%, 5%, 10%, 25%, 75%, 90%, 95% and 99% quantiles. The dashed red line marks the median of the conditional projection with fixed paths of the CISS and the PMI as forecast conditions.

All in all, the counterfactual simulations support the narrative that the GFC was a truly systemic financial crisis, leading to a deep economic recession, and that such devastating financial strains can be well captured by the CISS. On the other hand, the model also supports the contention that the COVID-19 crisis is mainly a real phenomenon, representing a huge shock to the real economy that also affected the stability of the financial system, but only temporarily.

While it is likely that the initial financial strains somewhat propagated the large macro shocks, it appears that policymakers successfully intervened with decisive large-scale support measures that helped the financial system to maintain its crucial intermediation function in highly critical and uncertain times, thereby avoiding the (re)emergence of a vicious macro-financial cycle (Altavilla et al. (2020)).

7. Conclusions

Financial stress indices, such as the CISS, have become an integral part of the analytical surveillance toolkit of most financial stability authorities. In particular, systemic FSIs support macroprudential policy in its crisis management function by allowing policymakers to monitor realised systemic risk in more or less real time (Freixas et al. (2015)). Moreover, in typical growth-at-risk regression frameworks (Adrian et al. (2019), Adams et al. (2020), Boyarchenko et al. (2023)), the CISS is found to predict well short-term tail risks to real GDP and other macroeconomic variables of interest. This strong short-term predictive power of the CISS has proved helpful in counterfactual policy analysis, which can serve the preventive arm of macroprudential policy (Chavleishvili et al. (2021)). In a similar context, Chavleishvili et al. (2023) shows how financial stress indices can be used to quantify the short- and long-term costs and benefits of a monetary policy designed to lean against the financial winds. The CISS can also help to forecast macroeconomic variables under conditions of extreme uncertainty caused by extraordinary financial shocks.

In addition, our study raises questions that require further research. For example, why do composite indicators tend to outperform individual asset price indicators in forecasting standard macroeconomic variables, and which dimension reduction technique works best in this regard? And does this superiority vary with the design of the composite indicator, or does it simply reflect the fact that combined forecasts generally tend to improve forecast accuracy over individual forecasts?

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

Appendix A. Composition of the daily CISS

Table 3: 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: $\text{CMAX}_t = 1 - x_t / \max[x_i \in (x_{t-j} \vert j = 0, 1, \ldots, 520)]$
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 4: 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: $\text{CMAX}_t = 1 - x_t / \max[x_i \in (x_{t-j} \mid j = 0, 1, \ldots, 520)]$
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

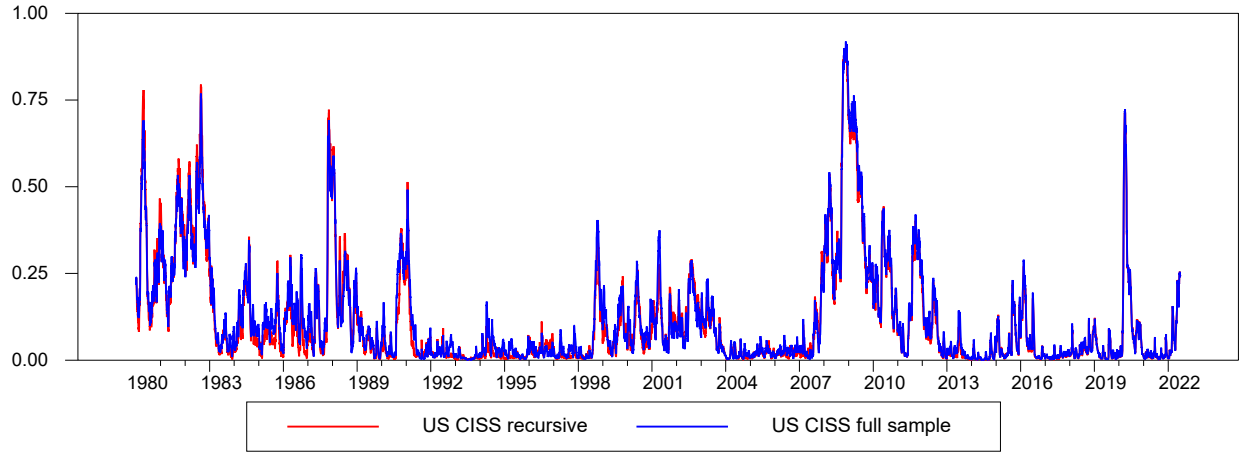

Fig. 14. The figure shows daily values of the standard US CISS computed from stress factors recursively transformed as from 1 January 2002 (red line) and a variant of the CISS using full-sample information at all times (blue line). Data is from 2 January 1980 to 29 May 2020.

Appendix C. VAR data

The CISS 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 are 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 (Doan (2016)). The procedure is very similar to the approach advocated by Stock and Watson (2010) and recently applied to derive a monthly series of euro area real GDP data in Jarocinski and Karadi (2020). Raw data is taken from the ECB's SDW and 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. Data source is Markit. Figure 15 plots the PMI along with the monthly series of annual real GDP growth.

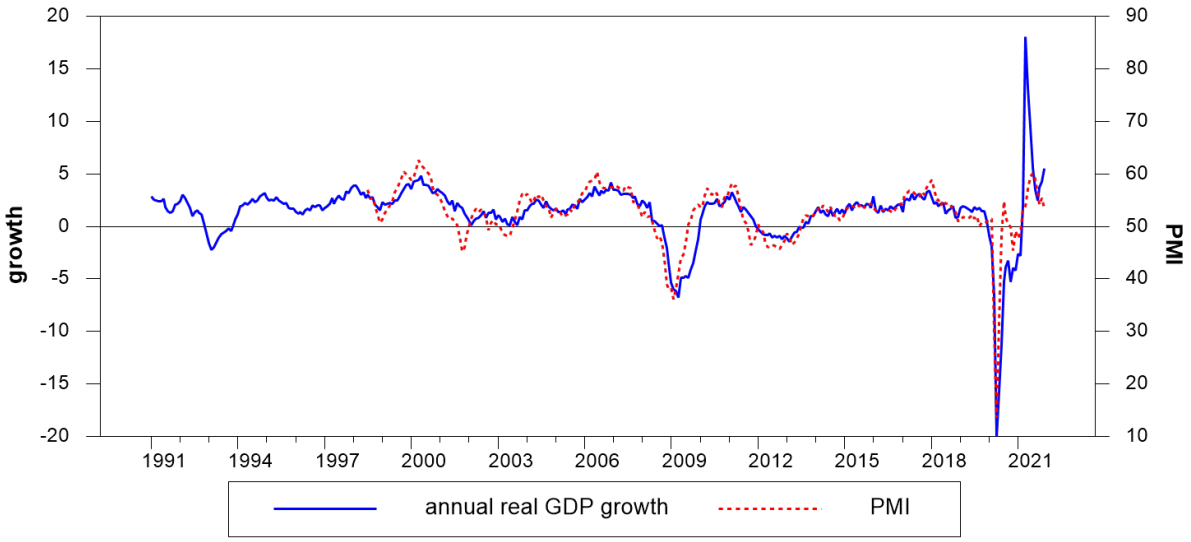

Fig. 15. The figure shows the year-on-year change in log real GDP and the PMI Composite Output Index for the euro area. A PMI reading above 50 indicates expansion in business activity and below 50 indicates contraction. Log real GDP is interpolated from quarterly to monthly frequency informed by the industrial production index. Data is monthly from January 1991 (real GDP) and July 1998 (PMI) to December 2021.

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.

Adrian, T., Boyarchenko, N., Giannone, D., 2019. Vulnerable growth. American Economic Review 109, 1263–1289.

Adrian, T., Boyarchenko, N., Giannone, D., 2020. Multimodality in macro-financial dynamics, Federal Reserve Bank of New York Staff Reports No. 903.

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.

Blix Grimaldi, M., 2010. Detecting and interpreting financial stress in the euro area, ECB Working Paper No. 1214, June.

Bloom, N., 2009. The impact of uncertainty shocks. Econometrica 77, 623–685.

Boyarchenko, N., Crump, R., Elias, L., Lopez Gaffney, I., 2023. What is "outlook-at-risk?", Federal Reserve Bank of New York Liberty Street Economics, February 15.

Boyarchenko, N., Crump, R. K., Kovner, A., Shachar, O., 2022. Measuring corporate bond market dislocations, Federal Reserve Bank of New York Staff Reports No. 957.

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.

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 No. 2565, June.

Chavleishvili, S., Kremer, M., Lund-Thomsen, F., 2023. Quantifying financial stability trade-offs for monetary policy: a quantile var approach, ECB Working Paper, forthcoming.

Chavleishvili, S., Manganelli, S., 2019. Forecasting and stress testing with quantile vector autoregression, ECB Working Paper No. 2330, November.

Chavleishvili, S., Mönch, E., 2023. Natural disasters as macroeconomic tail risks, mimeo.

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.

Doan, T., 2016. Disaggregate: Rats procedure to implement general disaggregation (interpolation/distribution) procedure, statistical Software Components RTS00050, Boston College Department of Economics.

Dovern, J., van Roye, B., 2014. International transmission and business-cycle effects offinancial 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.

Garcia-de-Andoain, C., Kremer, M., 2017. Beyond spreads: measuring sovereign market stress in the euro area. Economics Letters 159, 153–156.

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.

Groen, J. J. J., Nattinger, M. B., Noble, A. I., 2022. Measuring global financial market stresses, Federal Reserve Bank of New York Staff Reports No. 940.

Gurley, J. G., 1961. Review of "A Program for Monetary Stability, by Milton Friedman". Review of Economics and Statistics 43, 307–308.

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.

Inoue, A., Kilian, L., Kiraz, F., 2009. Do actions speak louder than words? Household expectations of inflation based on micro consumption data. Journal of Money, Credit and Banking 41, 1331–1363.

Jarocinski, M., Karadi, P., 2020. Deconstructing monetary policy surprises — the role of information shocks. American Economic Journal: Macroeconomics 12, 1–43.

Jarocinski, M., Mackowiak, B., 2017. Granger causal priority and choice of variable in vector autoregressions. Review of Economics and Statistics 99, 319–329.

Kilian, L., Lütkepohl, H., 2017. Structural vector autoregressive analysis. Cambridge University Press, Cambridge.

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.

Kliesen, K. L., Smith, D. C., 2010. Measuring financial market stress. Federal Reserve Bank of St. Louis Economic Synopses .

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.

Monin, P., 2017. The OFR financial stress index, OFR Working Paper 17-04.

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.

Stock, J. H., Watson, M. W., 2010. Research memorandum: Distribution of quarterly values of GDP/GDI across months within the quarter, mimeo.

Stuart, A., Ord, J., 1994. Distribution theory. In: *Kendall's Advanced Theory of Statistics, Vol. 1*, Arnold, London.

van Roye, B., 2014. Financial stress and economic activity in germany. Empirica 41, 101–126.

Vermeulen, R., Hoeberichts, M., Vasicek, B., Zigraiova, D., Smidkova, K., de Haan, J., 2015. Financial stress indices and financial crises. Open Economies Review 26, 383–406.

Wasserman, L., 2003. All of Statistics: A Concise Course in Statistical Inference. Springer, Berlin.

Acknowledgements

We thank Oliver Grothe, Thomas Kostka, Frederik Lund-Thomsen, Simone Manganelli, Vicenzo Quadrini, Bernd Schwaab, Alan Taylor and seminar participants at the ECB, the International Conference on Econometrics and Statistics in Hong Kong, the 1st Annual Workshop of the ESCB Research Cluster 3, the International Association for Applied Econometrics, the International Finance and Banking Society (IFABS), the Verein für Socialpolitik, Tsinghua University, the University of Rome III, the Bank of England, the Federal Reserve Board of Governors, and the American Economic Association for helpful comments and discussions. We also thank Paul Konietschke for excellent research assistance.

Sulkhan Chavleishvili
Aarhus University, Aarhus, Denmark; email: sulkhan.chavleishvili@econ.au.dk

Manfred Kremer
European Central Bank, Frankfurt am Main, Germany; email: manfred.kremer@ecb.europa.eu

Postal address 60640 Frankfurt am Main, Germany
Telephone +49 69 1344 0
Website www.ecb.europa.eu

PDF ISBN 978-92-899-6127-1 ISSN 1725-2806 doi:10.2866/25709 QB-AR-23-079-EN-N