# Macro-Financial Spillovers

John Cotter
Graduate School of Business, University College Dublin

Mark Hallam
Essex Business School, University of Essex

Kamil Yilmaz
Department of Economics, Koç University

# Macro-Financial Spillovers

John Cotter[1]        Mark Hallam[2]        Kamil Yilmaz[3]

17 April 2020

## Abstract

We analyse spillovers between the real and financial sides of the US economy allowing for differences in sampling frequency between financial and macroeconomic data. We find that financial markets are typically net transmitters of shocks to the real side of the economy, particularly during turbulent market conditions. Our macro-financial spillover measures are found to have significant predictive ability for future US macroeconomic conditions in both in-sample and out-of-sample forecasting environments. Furthermore, the predictive ability of our macro-financial measures frequently exceeds that of purely financial systemic risk measures previously employed in the literature for the same task. (*JEL* G10, G01, E30, E44, C13, C32)

**Keywords:** spillovers, connectedness, macro-financial, mixed-frequency, forecasting

[1]Graduate School of Business, University College Dublin, e-mail: john.cotter@ucd.ie
[2]Essex Business School, University of Essex, e-mail: mark.hallam@essex.ac.uk
[3]Department of Economics, Koç University, e-mail: kyilmaz@ku.edu.tr

# 1 Introduction

The 2008-2009 crisis highlighted the potential for adverse economic and financial shocks to spread across markets and countries. This has led to a significant increase in interest on topics of financial contagion, systemic risk and spillovers both in the academic community, and also from monetary authorities and financial regulatory bodies. One specific aspect of this topic attracting particular attention is the development of new quantitative measures and statistical tests for spillovers[1] and systemic risk. Some key examples including Allen et al. (2012), Billio et al. (2012), Diebold and Yilmaz (2014), Adams et al. (2014), Engle et al. (2015), Adrian and Brunnermeier (2016) and Brownlees and Engle (2017).

These previous studies have identified significant temporal and cross-sectional variation in the levels of financial spillovers and systemic risk. Substantial increases are typically observed during volatile periods and financial crises (see for example Adams et al., 2014, Diebold and Yilmaz, 2014). Further, some financial markets or institutions are found to be more central to the transmission of shocks than others. Such insights into the sources and transmission channels of spillovers and systemic risk can guide regulatory policy in an effort to minimise the spread of crises and limit their adverse effects.

The presence of macro-financial linkages between the financial and non-financial or real sides of the economy imply that significant adverse shocks to one sector will ultimately affect both sides of the economy. For example, a deterioration in financial conditions is expected to negatively impact the real side of the economy through a reduction in the willingness of financial firms to extend credit to corporate clients, which will in turn suppresses corporate investment (see Almeida and Campello, 2007, Ivashina and Scharfstein, 2010 and Cingano et al., 2016). Likewise, adverse macroeconomic shocks may feed back into financial markets, by increasing corporate defaults or reducing firm equity values.

The importance of this macro-financial dimension of spillovers and systemic risk is frequently noted in the literature (see, for example, Brunnermeier et al., 2011). Yet the existing work on spillover measurement has focused almost entirely on the financial sector, largely ignoring the real or non-financial side of the economy. The few exceptions that address this topic include Baur (2012), Claessens et al. (2012) and Dungey et al. (2013). In these examples the real side of the economy enters only via financial data for non-financial firms, rather than through the

---

[1]Across the finance and econometrics literatures the terms 'spillovers' and 'connectedness' have both been used to describe effects of the type studied here and are used interchangeably in the current work.

macroeconomic series of interest. In a similar manner both Allen et al. (2012) and Brownlees and Engle (2017) create non-financial variants of their respective systemic risk indices by utilising data for non-financial firms, though it should be noted that their objective is not to study the interaction between the two sides of the economy.

By working with purely financial data, such approaches sidestep the practical issue that most financial series are available at much higher sampling frequencies than the monthly or quarterly frequency of most macroeconomic series. Given that the vast majority of econometric methods assume series to be sampled at a single common frequency, employing a combination of macroeconomic and financial series typically requires financial series to be aggregated to the lower frequency of the macroeconomic series, thus discarding potentially relevant high-frequency financial information. However, the obvious downside of employing financial data for non-financial firms in this manner is that they are likely to provide a less direct and narrower measure of conditions in the real sector of the economy than true macroeconomic series.

Given these gaps in the literature, we perform a detailed empirical investigation of macro-financial spillovers in the US economy. To achieve this, we develop a new methodology for estimating spillovers that is designed specifically for the macro-financial context and avoids the limitations outlined above. Furthermore, we investigate the predictive ability of our macro-financial spillover measures for future macroeconomic conditions. This study adds to and complements existing work in several fields of the finance literature. Most directly we add to the existing literature on the measurement of spillovers and contagion. The work also contributes to the literature on macro-financial interaction and linkages more generally, particularly the use of financial and macro-financial predictors of macroeconomic conditions.

Our approach employs the Diebold-Yilmaz spillover measures developed by Diebold and Yilmaz (2014) for the quantitative measurement of financial spillovers. The Diebold-Yilmaz (DY) approach provides a set of spillover measures at various levels of aggregation that are directional in nature, thus allowing a detailed analysis of spillover structure. However, unlike previous work employing these spillover measures, we work in a mixed-frequency environment in which the series employed are sampled at different frequencies, such as monthly and daily. As discussed above, this is highly relevant for the macro-financial datasets of interest in this context, given the differences in sampling frequency at which macroeconomic and financial series are typically available. More specifically, we employ the approach for modelling mixed-frequency data of Ghysels (2016) to permit spillovers to be estimated directly from macro-financial datasets

in which the series of interest may be sampled at different frequencies. We can thus avoid the loss of potentially relevant information induced by the data aggregation required when using a common-frequency modelling approach.

We focus on equity and bond markets for the financial side of the economy and a broad measure of economic conditions for the real side of the economy. Both the dynamics and magnitude of estimated US macro-financial spillovers obtained from our new mixed-frequency extension of the DY approach differ significantly from those obtained from a simpler, but otherwise equivalent, common-frequency modelling approach that discards additional high-frequency financial information. Perhaps most notably, the magnitude of spillovers estimated by our new mixed-frequency DY approach is typically substantially higher than that implied by the analogous common-frequency approach. This finding suggests that the use of a common-frequency modelling approach that discards additional high-frequency financial information results in estimated macro-financial connectedness being lower on average. We also find that financial markets are usually net transmitters of shocks to the real economy. This is particularly evident when markets face turbulence, most notably during the 2008-2009 crisis. These findings are also shown to hold more generally, carrying through to a group of five other advanced economies.

In the literature on systemic risk, work such as Allen et al. (2012), Giglio et al. (2016) and Brownlees and Engle (2017) has found empirical evidence that systemic risk measures have forecasting ability for future macroeconomic conditions. Motivated by these studies, we perform a detailed empirical analysis to see if the same is true for our macro-financial spillover measures. A key advantage of our DY-based approach in this context is that it directly provides a set of pairwise directional spillover measures, rather than just a single higher-level numerical measure. This allows us to straightforwardly construct combination forecasts from these individual pairwise measures that are found to exhibit particularly consistent and strong predictive ability for future macroeconomic conditions.

We find that not only do our measures have significant predictive ability for macroeconomic conditions, but that they provide more consistently strong performance across a range of forecast horizons than existing systemic risk measures including the CATFIN and SRISK measures proposed by Allen et al. (2012) and Brownlees and Engle (2017) respectively. This outperformance over existing systemic risk measures is especially pronounced in an out-of-sample context, or when forecasting more specific aspects of macroeconomic conditions, rather than aggregate macroeconomic conditions. Furthermore, the improvements in predictive accuracy provided by

our measures is particularly large during the 2008-2009 crisis.

The remainder of the paper is organised as follows. Section 2 introduces the methodology we develop for quantitatively measuring the strength and structure of macro-financial spillovers. Section 3 applies our methodology to perform a detailed empirical analysis of macro-financial spillovers in the US and a more concise analysis for a set of five other advanced economies. Section 4 examines the predictive ability of our macro-financial spillover measures for future macroeconomic conditions and finally Section 5 concludes.

## 2 Macro-Financial Spillover Estimation

Our quantitative measures of spillovers are based on the established DY spillover methodology of Diebold and Yilmaz, (2012, 2014). This approach has been applied extensively to study financial spillovers, but has not been employed in a macro-financial context. The DY approach relies on the forecast error variance decomposition from a vector autoregressive (VAR) model, in which all series are observed at a common sampling frequency. However, the combinations of macroeconomic and financial time series of interest here will typically contain series at different sampling frequencies, with financial series generally available at much higher sampling frequencies than macroeconomic series.

The traditional solution would be to aggregate all high-frequency financial time series to the sampling frequency of the lowest frequency macroeconomic time series employed, before applying the standard DY methodology to the transformed data. Whilst simple, the obvious drawback of such an approach is the potentially relevant information lost when aggregating the higher frequency financial series. Instead, we avoid these issues by replacing the standard common-frequency VAR model with a mixed-frequency VAR (MF-VAR) model. This allows us to employ high and low frequency series together, estimating our spillover measures directly from a mixed-frequency macroeconomic and financial dataset.

### 2.1 The Mixed-Frequency VAR Model

We follow Ghysels (2016) for the specification and estimation of the MF-VAR. To illustrate the approach we assume for simplicity that there are only two distinct sampling frequencies (high-frequency and low-frequency). We also assume that the number of high-frequency time periods is the same in each low-frequency period. Both of these are true for the empirical analysis

here, however it should be noted that the methodology is applicable generally and that these assumptions can be relaxed at a cost of more complex notation and implementation.

Formally, we observe a $K$-dimensional mixed-frequency vector process, which contains $K_L < K$ low-frequency macroeconomic series and $K_H = K - K_L$ high-frequency financial series. In terms of low-frequency time periods, which we index by $\tau_L$, the low-frequency macro series are observed once per period and collected in the $K_L$-dimensional vector process $x_L(\tau_L)$. Each high-frequency financial series is observed $m$ times every low-frequency time period. We group the high-frequency observations within each low-frequency period by series, giving $K_H$ $m$-dimensional vectors $x_{H,1}(\tau_L), \ldots, x_{H,K_H}(\tau_L)$.

We create a stacked vector for each low-frequency time period that contains both the $K_L$-dimensional vector of low-frequency macroeconomic observations, $x_L(\tau_L)$, and the $K_H$ $m$-dimensional vectors $x_{H,1}(\tau_L), \ldots, x_{H,K_H}(\tau_L)$ with all the high-frequency financial data observed during the same low-frequency time period. The resulting stacked vector is denoted by $\underline{x}(\tau_L)$ and is of dimension $K_x$, where $K_x \equiv (mK_H + K_L)$:

$$\underline{x}(\tau_L) \equiv \left[ x_{H,1}(\tau_L)', \ldots, x_{H,K_H}(\tau_L)', x_L(\tau_L)' \right]'$$

Following Ghysels (2016), we then specify a standard VAR model for the stacked mixed-frequency vector $\underline{x}(\tau_L)$. The general form of the $p$-th order MF-VAR is thus given by:

$$\underline{x}(\tau_L) = A_0 + \sum_{j=1}^{p} A_j \underline{x}(\tau_L - j) + \underline{\varepsilon}(\tau_L) \tag{2.1}$$

where $A_0$ is an $K_x$-dimensional parameter vector, $A_j, j = 1, \ldots p$ are $(K_x \times K_x)$ parameter arrays and $\underline{\varepsilon}(\tau_L)$ is an $K_x$-dimensional vector of errors. Despite the somewhat non-standard composition of the vector $\underline{x}(\tau_L)$, the model is mathematically equivalent to a standard VAR. As such, standard methods for estimation and analysis of VAR models can be employed.

In addition to the stacked mixed-frequency vector process $\underline{x}(\tau_L)$ introduced above, we will also consider the associated $K$-dimensional low-frequency vector process denoted by $\overline{x}(\tau_L)$, which contains both the $K_L$ low-frequency macroeconomic series and the $K_H$ high-frequency financial series appropriately aggregated down to the lower frequency of the macroeconomic series:

$$\overline{x}(\tau_L) \equiv \left[ x_{HtL}(\tau_L)', x_L(\tau_L)' \right]'$$

where $x_{HtL}(\tau_L)$ is used to denote the set of high-frequency series aggregated to the lower frequency in the $\tau_L$-th time period. In the current work we primarily employ return levels as financial series, with return volatilities also considered in the Appendix. As such, the high-frequency financial series contained in the vectors $x_{H,1}(\tau_L), \ldots, x_{H,K_H}(\tau_L)$ will be weekly returns (or return volatilities) and those in the aggregated low-frequency vector $x_{HtL}(\tau_L)$ are monthly returns (or return volatilities). We also specify a standard common-frequency VAR model for the low-frequency vector process $\overline{x}(\tau_L)$, which we will refer to as the common-frequency VAR (CF-VAR).

It should be noted that whilst the MF-VAR and CF-VAR are both technically specified at the lower sampling frequency, the MF-VAR also incorporates higher frequency information available within each low frequency time period that is not used by the CF-VAR. We use a combination of monthly macroeconomic series and weekly financial series, allowing the MF-VAR to incorporate potentially relevant intra-month information on financial market behaviour at the weekly frequency. In the case of the CF-VAR, this high-frequency information is discarded when the financial series are aggregated down to the monthly frequency. Ghysels (2016), Schorfheide and Song (2015) and others have previously shown that the use of mixed-frequency methods may provide gains in accuracy for both estimation and forecasting in the context of VAR models relative to a common-frequency approach. During our empirical analysis we will directly compare the DY spillover measures obtained from the traditional CF-VAR with those from the MF-VAR to investigate the impact of including this additional high-frequency information.

## 2.2 Forecast Error Variance Decomposition and Spillover Measures

After the specified VAR model has been estimated, the next step is to compute the forecast error variance decomposition (FEVD) arrays which are then used to compute the various DY spillover measures. Following Diebold and Yilmaz (2014), we employ the approach of Pesaran and Shin (1998) to compute generalised FEVD values. This approach is widely employed in the literature and so details are relegated to Appendix A.1.

For a generic $K$-dimensional VAR, the FEVD arrays are of dimension $(K \times K)$ and of the form:

$$\begin{bmatrix} \phi_{11}(H) & \cdots & \phi_{1K}(H) \\ \vdots & \ddots & \vdots \\ \phi_{K1}(H) & \cdots & \phi_{KK}(H) \end{bmatrix} \quad \text{for} \quad H = 1, 2, \dots \tag{2.2}$$

where $\phi_{kl}(H)$ for $k, l = 1, \dots, K$ is the fraction of the $H$-step-ahead error variance in forecasting series $k$ that is attributable to shocks in series $l$. The FEVD array elements thus have a clear interpretation as possible measures of spillovers between the series in the system. More specifically, the pairwise DY spillover from series $i$ to series $j$ is given by:

$$S_{ij}(H) = \frac{100}{K} \cdot \phi_{ji}(H) \tag{2.3}$$

Multiplying the relevant FEVD element $\phi_{ji}(H)$ by the factor $100/K$ ensures that each pairwise spillover value is expressed as a percentage of the total forecast error variance across all series in the VAR[2].

The DY spillover measures are complementary to alternative systemic risk measures as tools for monitoring market conditions. It is worth emphasising that the DY pairwise spillover measures are directional in nature in the sense that $S_{ij} \neq S_{ji}$ for $i \neq j$, which is made possible through the use of the generalised VAR approach to computing the FEVD discussed above. This is a key theoretical advantage compared to most common measures of pairwise association such as correlation, which are non-directional in nature and thus can measure only the strength of association between two series. Indeed this is also an important difference between the DY spillover measures and other established measures of systemic risk employed in the finance literature, such as the CATFIN measure of Allen et al. (2012) or the SRISK measure of Brownlees and Engle (2017).

The fact that the DY approach naturally produces a set of pairwise spillover measures rather than simply a single numerical measure also emerges as an important advantage when we employ the measures to forecast macroeconomic conditions in Section **??**. As demonstrated later, the fact that each of the directional pairwise spillover measures contains different information concerning the structure of macro-financial spillovers allows us exploit the use of combination forecasts, which have had much success in the previous literature on predictability.

---

[2]This follows because, as discussed in Appendix A.1, the sum of the elements in each row of the FEVD array equals one, giving a sum over all elements equal to $K$.

Whilst the pairwise spillover measures of equation (2.3) permit a detailed analysis of the direction and structure of spillovers, more aggregated measures may also be useful to concisely quantify the overall strength of spillovers. We therefore employ both the disaggregated pairwise measures above and the most aggregated measure proposed by Diebold and Yilmaz (2014). This is referred to as the total spillover index and provides a single numerical measure of the overall level of spillovers between the series included in the underlying VAR. It is computed as:

$$S(H) = \frac{100}{N} \sum_{\substack{i,j=1 \\ i \neq j}}^{K} \phi_{ji}(H) \tag{2.4}$$

The total spillover index gives the percentage of the total $H$-step-ahead forecast error variance for all series that is attributable to shocks *across* series i.e. excluding the direct effect of shocks to each given series on itself.

## 2.3 Transformation of Mixed-Frequency Forecast Error Variance Decomposition

The FEVD arrays for the MF-VAR model are computed as in the common-frequency case. However, they will have a non-standard structure arising from the non-standard composition of the stacked vector, $\underline{x}(\tau_L)$. Whilst the standard DY spillover measures can be computed directly from these mixed-frequency FEVD arrays, the interpretation of the measures obtained will differ from the standard common-frequency case.

The relevant issues and concepts are best illustrated using a simple example, for which we use a bivariate model with one low-frequency monthly macroeconomic series and one high-frequency financial weekly series. We thus have $m = 4, K_L = 1$ and $K_H = 1$, giving a stacked mixed-frequency vector of dimensions $K_x = 5$, with the form $\underline{x}(\tau_L) = [x_H(\tau_L, 1), \ldots, x_H(\tau_L, 4), x_L(\tau_L)]'$. For the corresponding common-frequency VAR, we have a $(2 \times 1)$ vector process $\overline{x}(\tau_L) = [x_{HtL}(\tau_L), x_L(\tau_L)]'$. This results in $(5 \times 5)$ FEVD arrays for the MF-VAR and $(2 \times 2)$ arrays for the CF-VAR, given respectively by:

$$\begin{bmatrix} \theta_{11}(H) & \ldots & \theta_{15}(H) \\ \vdots & \ddots & \vdots \\ \theta_{51}(H) & \ldots & \theta_{55}(H) \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \phi_{11}(H) & \phi_{12}(H) \\ \phi_{21}(H) & \phi_{22}(H) \end{bmatrix} \quad \text{for} \quad H = 1, 2, \ldots \tag{2.5}$$

9

with the differences in notation used only to distinguish the FEVD elements for the MF-VAR and CF-VAR.

It is clear that the FEVD arrays for the MF-VAR in (2.5) will be larger than those for the corresponding CF-VAR, since $K_x > K$. This arises because the weekly high-frequency series observed in each low-frequency monthly time period are treated mathematically as separate series when estimating the MF-VAR, but enter the CF-VAR as a single monthly series. As a result, in the common-frequency case a single FEVD element completely characterises the directional spillovers at the chosen forecast horizon between a given pair of macroeconomic or financial series, whereas in the mixed-frequency case it will generally be characterised by multiple FEVD array elements.

We therefore develop an approach for transforming the FEVD arrays obtained from the MF-VAR to produce new FEVD arrays with the same structure and dimensions as those for the corresponding CF-VAR. We can then compute DY spillover measures from these transformed arrays that are directly comparable to those in the standard common-frequency case. The basic intuition of the transformation approach is outlined here, with mathematical details found in Appendix A.2.

Intuitively we exploit the correspondence between the elements of the FEVD arrays for the mixed-frequency and common-frequency cases. More specifically, for the current example we group the MF-VAR FEVD elements in (2.5) into sub-arrays as follows:

$$\begin{bmatrix} \Theta_{11}(H) & \Theta_{12}(H) \\ \Theta_{21}(H) & \Theta_{22}(H) \end{bmatrix} \quad \text{for} \quad H = 1, 2, \ldots \tag{2.6}$$

where:

$$\Theta_{11}(H) \equiv \begin{bmatrix} \theta_{11}(H) & \ldots & \theta_{14}(H) \\ \vdots & \ddots & \vdots \\ \theta_{41}(H) & \ldots & \theta_{44}(H) \end{bmatrix} \qquad \Theta_{12}(H) \equiv \begin{bmatrix} \theta_{15}(H) \\ \vdots \\ \theta_{45}(H) \end{bmatrix}$$

$$\Theta_{21}(H) \equiv \begin{bmatrix} \theta_{51}(H) & \ldots & \theta_{54}(H) \end{bmatrix} \qquad \Theta_{22}(H) \equiv \theta_{55}(H)$$

Each of the sub-arrays $\Theta_{kl}(H)$ in (2.6) can be viewed as a mixed-frequency analogue of the corresponding scalar element $\phi_{kl}(H)$ from the CF-VAR FEVD array in (2.5). For example, the

$(4 \times 1)$ sub-vector $\Theta_{12}(H)$ characterise the effects of shocks to the monthly low-frequency series (series 2) on the weekly high-frequency series (series 1). Specifically, $\theta_{i5}$ for $i = 1, \ldots, 4$ measures the fraction of the $H$-step-ahead error variance in forecasting the high-frequency series in week $i$ of the month that is attributable to shocks in the low-frequency series. The scalar element $\phi_{12}(H)$ for the common-frequency case describes the same directional pairwise relationship for the case where both series are observed at the lower monthly frequency.

The approach detailed in Appendix A.2 transforms each of the sub-arrays $\Theta_{kl}(H)$ in (2.6) into a scalar value, whose interpretation is directly comparable with the corresponding element $\phi_{kl}(H)$ in the standard common-frequency case. This comparability of the values is ensured by directly basing the transformation used on the mathematical definition of the generalised FEVD elements.

# 3    Macro-Financial Spillovers

We focus primarily on the estimation and analysis of macro-financial spillovers for the United States, given the central role that the country plays in global economic and financial markets. However, in Section 3.3 we also more briefly examine macro-financial spillovers for five additional advanced economies, namely Canada, France, Germany, Japan and the United Kingdom, in order to explore whether the empirical findings for the US are replicated more widely.

Section 3.1 begins by describing the data and implementation of the approach employed to estimate spillovers. Section 3.2 performs a graphical analysis of the US macro-financial spillover measures obtained using our approach. In this subsection we also include measures obtained from the existing common-frequency approach, to allow us to assess the practical effects of ignoring the high-frequency financial data when using a more traditional modelling approach. The international analysis of macro-financial spillovers for other advanced economies follows in Section 3.3.

## 3.1    Data and Spillover Estimation

On the real side of the economy our series of interest is the Chicago Fed National Activity Index (CFNAI), which is frequently used in empirical work to provide a single numerical measure of US macroeconomic activity that is broader and less noisy than specific series such as industrial production (see for example Allen et al., 2012). The CFNAI is a composite index derived from 85

underlying macroeconomic indicators, which are grouped into four categories: production and income, employment unemployment and hours, consumption and housing, and sales, orders and inventories. The CFNAI is employed at its regular monthly frequency in the current empirical analysis. The CFNAI is constructed to have a mean value of zero, with positive (negative) values corresponding to growth above (below) its historical trend. We thus work with the level of the CFANI series, given that this corresponds in a broad sense to the change in macroeconomic conditions.

On the financial side we focus on equity and bond markets, represented by the S&P500 equity index and the 10-year US Treasury Note respectively. The strength and structure of spilllovers between bond markets and the real economy may vary with bond maturity and type (as suggested by Brenner et al., 2009). For simplicity we restrict our attention to US sovereign bonds, specifically the 10-year Treasury Note.

We focus primarily on return levels for the financial series, under the assumption that changes in macroeconomic conditions will be linked to changes in asset values. However, we do repeat the core analysis of Section 3.2 using return volatilities for the sake of completeness, with the results found in Appendix C.1. The key empirical findings follow through mostly unchanged, notwithstanding some differences, such as the dynamics of the indexes during the recent crisis.

For the non-US analysis of Section 3.3, we use month-on-month industrial production growth as our macroeconomic variable for the real side of the economy given the lack of an equivalent index to the CFNAI for other countries. To provide results that are entirely comparable across countries, we also repeat the US analysis using monthly IP growth in place of the CFNAI. Similar to the US. for the financial side of the economy we use return levels for the major national equity index for each country[3] and generic national 10-year government bond indexes obtained from Bloomberg.

For the US analysis, our sample period spans 1975:01 to 2018:04, thus including many significant economic and financial events of recent decades. Due to limited data availability for the non-US government bond index series, the sample period for the international analysis begins in 1990.

The raw data for the two financial series consist of daily closing prices, from which we

---

[3]S&P/TSX for Canada, CAC40 for France, DAX for Germany, Nikkei 225 for Japan, FTSE100 for the UK and as before the S&P500 for the US.

produce a closing price series at a weekly frequency. To sidestep the practical issues caused by the variation in the number of weeks per month, we employ a data pre-processing and transformation approach to produce weekly series with a constant four weeks per month[4]. These weekly closing prices are used to compute weekly returns. Further details of the data processing approaches, together with plots of all the US series employed, can be found in Appendix B.

Our interest lies in obtaining dynamic estimates of macro-financial spillovers, rather than static full-sample estimates, to investigate how the strength and structure of spillovers has varied over time. To achieve this we use a standard rolling window estimation approach in which the parameters of the MF-VAR and the connectedness measures are re-estimated for each window. A window length of 60 months is employed, since it appears to offer a good balance between providing a sufficient sample size to estimate the parameters of the underlying MF-VAR to an appropriate level of accuracy, and allowing dynamics of connectedness to be captured. We have however checked the robustness of our results to reasonable changes in the window length and there is no qualitative impact on the results.

When computing the spillover measures we primarily considered forecast horizons of 3, 6 and 12 months, consistent with most previous studies. However, we found that the estimates did not show significant sensitivity to the choice of forecast horizon, and so report results only for the horizon of 3 months. It should however be noted that the connectedness measures obtained for different choices of forecast horizon are simply measuring different aspects of the connectedness structure between the chosen series. The choice should not therefore be viewed as a model specification issue with a single optimal value.

## 3.2 US Macro-Financial Spillovers

We begin in Figure 1 by plotting the total spillover indexes obtained from both the mixed-frequency and common-frequency DY approaches. It can be seen that the estimated total spillover indexes obtained from the two approaches show broadly the same movements over the sample period. However, despite the high correlation between the indexes, the level of total connectedness implied by the new mixed-frequency approach is, with one or two exceptions, consistently higher than that obtained from the common-frequency approach. For example, the average value of total connectedness for the mixed-frequency and common-frequency DY

---

[4]This is not strictly necessary, since the MF-VAR framework can accommodate deterministic time variation in the number of high-frequency observations per low-frequency period (as discussed by Ghysels, 2016), but significantly simplifies the exposition and implementation.

approaches are 24.79% and 16.38% respectively, representing the proportion of the total forecast error variance in the entire system that is due to shocks across series[5]. Thus by aggregating the financial data to monthly frequency and ignoring the additional intra-monthly information it contains, one obtains substantially lower estimates of the level of connectedness across the real and financial sectors.
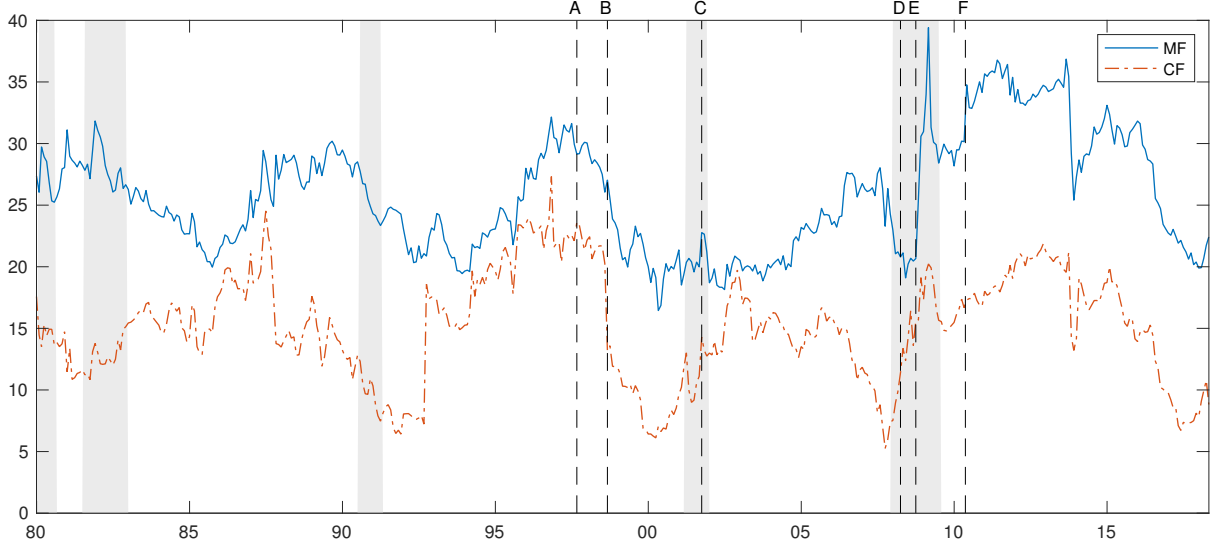


**Figure 1: Total spillover indexes between the financial and real economy series**
Total spillover indexes for mixed-frequency (denoted MF) and common-frequency (denoted CF) approaches are presented for the sample period 1980:01 to 2018:04. Return levels are employed for the financial S&P500 and 10-year Treasury Note series, and levels for the real economy CFNAI series. Values are computed using a 3-month forecast horizon and a 60-month rolling window. Points marked are as follows. A: Asian financial crisis, Jul '97, B: Russian financial crisis and LTCM collapse, Aug to Sept '98, C: September 11, Sept '01, D: collapse of Bear Stearns, Mar '08, E: Lehman Brothers collapse, AIG bailout and Fannie Mae and Freddie Mac being placed in government conservatorship, May '09, F: start of the EU debt crisis in April '10 and flash crash of May '10. Shaded areas correspond to NBER US recession dates.

A possible explanation for this finding of higher average spillover levels for the mixed-frequency case can be found in the previous literature on the effects of macroeconomic announcements on financial markets. Studies such as Andersen et al. (2003) and Green (2004) have employed high-frequency intraday financial data to study the effects of announcements over short time periods and have found significant intraday effects. However, they note that the use of lower frequency daily data prevents these effects from being observed and thus may bias estimates of the response in the financial markets downwards. Intuitively an analogous explanation can be applied in the current analysis, in which shocks to either the financial or real series may result in significant within-month spillovers that are visible through the use of weekly data for some series in the mixed-frequency case, but either ignored completely or

---

[5]It is interesting to note that the size of the total spillover index for our macro-financial application is smaller than the total spillover index typically observed in previous studies using purely financial datasets, which often reach values of 75% or 80%. This stems from the lower levels of connectedness present between financial markets and the real economy compared to between financial markets.

underestimated when using purely monthly data in the common-frequency approach.

Considering briefly the dynamics of the total spillover indexes, we see substantial fluctuations over the sample period, many of which coincide with major economic or financial events. Some notable examples are marked on Figure 1, with the largest spikes in spillovers occurring during the recent global financial crisis, particularly around the collapse of Lehman Brothers, the bailout of AIG and Fannie Mae and Freddie Mac being placed in government conservatorship.

In addition, with the exception of the global financial crisis of the last quarter of 2008 and April-May 2010, the movements in the mixed-frequency spillover measure are typically substantially smoother than the common-frequency spillovers, which frequently exhibits significant upward or downward jumps. This likely results from incorporation of intra-month financial information in the mixed-frequency analysis, which results in the effects of sustained shocks to financial series being spread across consecutive weeks and incorporated into the spillover index gradually. In the common-frequency case on the other hand, only the accumulated shock is observed at the end of the month, leading to a more significant jump in the index when this information is incorporated into the new value of the index.

The total spillover index provides an informative but highly aggregated measure that potentially hides many interesting details of the structure of macro-financial spillovers. We decompose the total spillover measures of Figure 1 into their component pairwise directional spillover measures. We first represent this information in the form of spillover decomposition plots in Figure 2. Given that the total spillover index equals the sum of all pairwise measures, the top of the complete shaded area corresponding to the relevant total spillover index (as previously plotted in Figure 1) and the shaded areas beneath representing the contribution of each individual pairwise spillover. These spillover decomposition plots thus provide an effective way to visualise the relative importance of each of the directional spillover channels to the total spillover index over time.

For the mixed-frequency case in panel (a), it is immediately clear that the vast majority of the total macro-financial spillover index is comprised of spillovers originating in financial markets, represented by the bottom four shaded areas of the spillover decomposition plots. The contribution of the real side of the economy (the sum of the areas represented by the yellow and orange colours) accounts for only a small part of the total spillovers between the sides of the economy in this context. For the common-frequency case in panel (b), spillovers from the real side of the economy are larger than in panel (a), most notably during the late 1980's
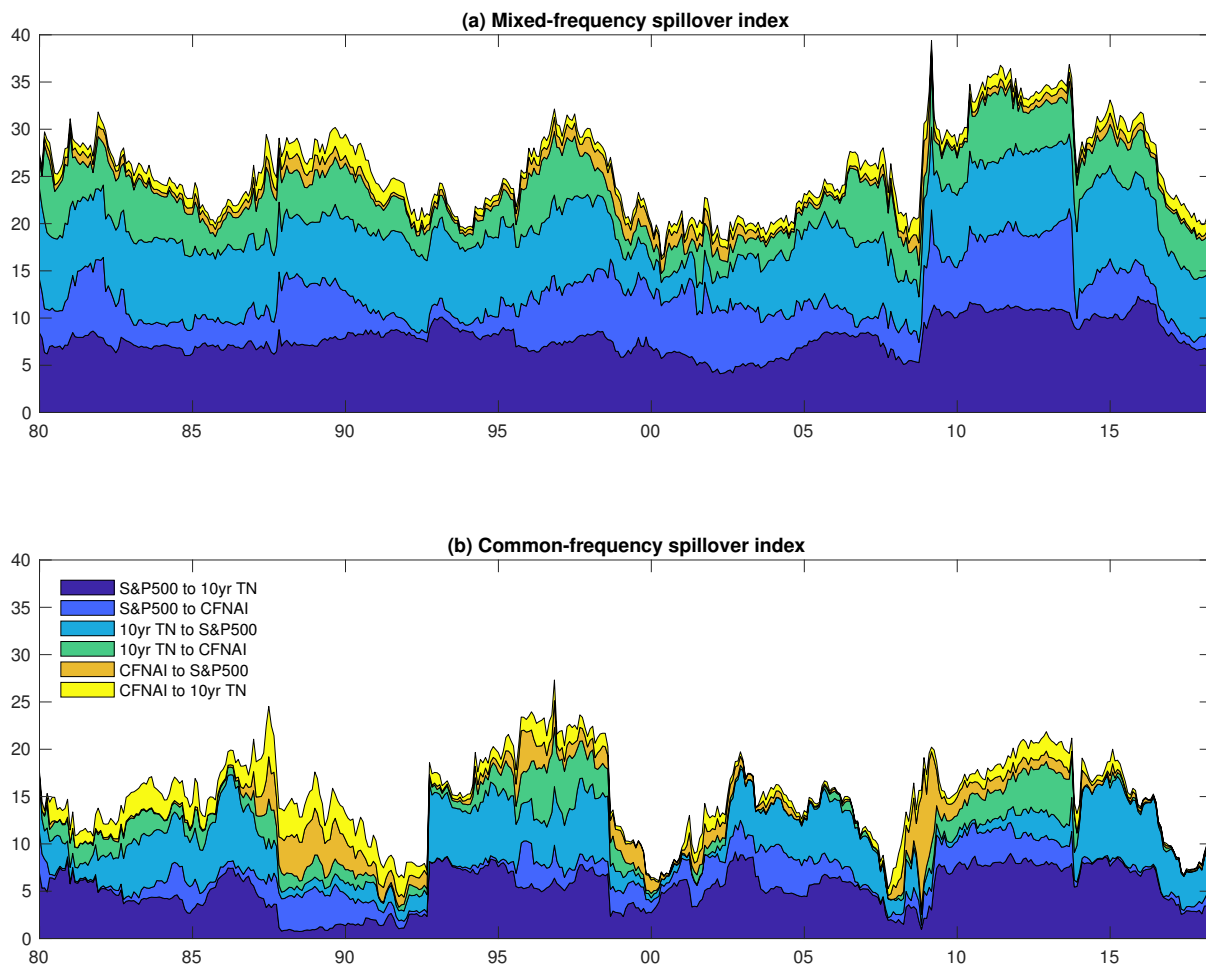
15

**Figure 2: Decomposition of total macro-financial spillover indexes into pairwise components**
The figure presents area plots in which the top of the complete shaded area corresponds to the relevant total spillover index and each shaded area below representing the contribution of each specific pairwise spillover to the value of the total spillover index. The relevant spillover measures for the mixed-frequency case are plotted in panel (a) and those for the common-frequency case in panel (b). Return levels are used for the financial S&P500 and 10-year Treasury Note series and levels for the real economy CFNAI series. Values are computed using a 3-month forecast horizon and a 60-month rolling window.

and during the 2008-2009 crisis. This difference in the relative contribution of the real side of the economy is also observed to a lesser extent in the early 1980s and in the late 1990's. As discussed above, we hypothesise that this difference is due to the common-frequency index not incorporating higher frequency financial information, and thus the estimates obtained suggest a relatively smaller role for financial markets.

Furthermore, a comparison of panels (a) and (b) of Figure 2 shows that mixed-frequency directional spillover measures are more stable over time compared to the common-frequency directional spillover measures. While the contributions of the pairwise directional spillover measures to total spillovers typically changes gradually in the mixed-frequency case, they often change drastically over time in the common-frequency case in a way that is difficult to meaningfully analyse or connect to known historical events.

We next represent a subset of the pairwise spillover measures as standard line plots in Figure 3. Panel (a) plots financial to financial pairwise spillover measures and panel (b) plots pairwise spillovers from each of the two financial markets to the real economy. Given the relative unimportance of spillovers from the real sector to financial markets highlighted above, we do not analyse these remaining two pairwise measures further. Figure 3 complements the spillover decomposition plots of Figure 2, with the latter being more suited to analysing the relative contribution of the individual spillover measures to the total level, and the former more suited to analysing the dynamics of specific spillover measures in absolute terms. In particular, Figure 3 allows us to more effectively examine how changes in pairwise spillovers relate to key historical events as we did previously for the total spillover index in Figure 1.
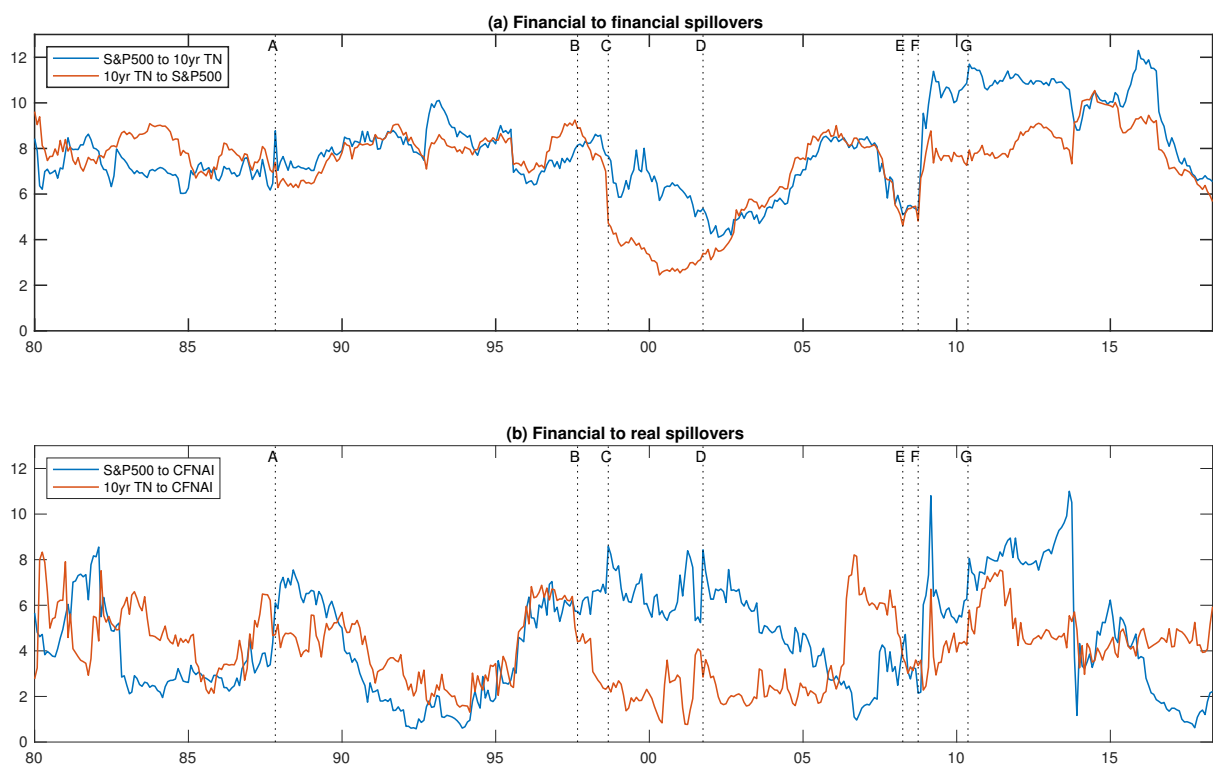


**Figure 3: Pairwise macro-financial spillover measures**
The figure presents plots of a subset of the pairwise macro-financial spillover measures for the sample period 1980:01 to 2018:04. All measures plotted are for the mixed-frequency case. Return levels are employed for the financial S&P500 and 10-year Treasury Note series, and levels for the real economy CFNAI series. Values are computed using a 3-month forecast horizon and a 60-month rolling window. Points marked are as follows. A: Black Monday, Oct '87, B: Asian financial crisis, Jul '97, C: Russian financial crisis and LTCM collapse, Aug to Sept '98, D: September 11, Sept '01, E: collapse of Bear Stearns, Mar '08, F: Lehman Brothers collapse, AIG bailout and Fannie Mae and Freddie Mac being placed in government conservatorship, May '09, G: start of the EU debt crisis in April '10 and flash crash of May '10.

There are several instances of increases in total spillovers being driven by increased spillovers from the financial sector to the real side of the economy. During the 1980 to 1982 inflationary period spillovers from the bonds markets to the real economy increased substantially and reached local peaks in three instances: March 1980, December 1980 and February 1982, all of which

coincide with increases in the federal funds target rate. These tightenings of monetary policy by the Volcker Fed led to higher interest rates, generating rapid contractionary impacts on the real economy, as captured by the spillovers from bond markets to the real economy reaching 8%. They subsequently declined gradually as inflation was brought under control leading to lower interest rates. These events are also visible from the mixed-frequency spillover measures in panel (a) of Figure 2, but not the equivalent common-frequency measures in panel (b).

The Federal Reserve implemented another rate hike cycle between December 1986 and September 1987, with the federal funds target rate rising from 5.75% to 7.25% over this period. As a consequence, spillovers from bond markets to the real economy increased from 3.2% in November 1986 to as high as 6.5% in May 1987. When the Federal Reserve lowered its policy rate back to 6.75% in response to Black Monday, spillovers from bonds to the real economy declined to 4% as of December 1987. However, the Federal Reserve returned to its rate hike cycle in March 1988 and increased the fed funds target rate all the way to 9.75% by the end of 1988 and kept it there for another year. As a result, spillovers the bond market to the real economy stayed above 4 percentage points until October 1990.

Increases in spillovers from equity markets to the real economy are visible after the Black Monday stock market crash in October 1987. The same occurred in the mid and late 1990s due to the build-up to the Mexican Tequila crisis beginning at the end of 1994, followed by the Asian and Russian financial crises, and the collapse of LTCM in 1997 and 1998. During this period spillovers from both the stock and bond markets to the real economy increased from around 2 percentage points to 6 percentage points. Both measures also increase in May 2000. when the Federal Reserve increased its policy rate from 6% to 6.5% despite the bursting of the dotcom bubble in the first half of 2000.

Following this, spillovers from equities and bonds to the real economy declined, though from mid-2006 both recorded a sharp increase from around 2 to 8 percentage points when the Fed caught the markets off-guard by increasing the federal funds target rate in both May and June. With the start of the 2008-2009 financial crisis spillovers from both bond and equity markets to the real economy increased sharply, with the latter being the largest contributor to the sudden increase in the total spillover index throughout this crisis.

### 3.3 International Evidence

In order to examine the structure of macro-financial spillovers in other countries, we perform a condensed version of the preceding US analysis for a set of advanced economies consisting of Canada, France, Germany, Japan and the United Kingdom. The analysis is performed country-by-country for each economy in isolation, although a true multi-country analysis to study international macro-financial spillovers would be an interesting avenue for future work. We present the mixed-frequency and common-frequency total spillover indexes for the six countries in Figure 4 and the corresponding decomposition plots in Figure 5. To conserve space we do not include spillover decomposition for the common-frequency case in Figure 5, however similar patterns are observed to those in Figure 2 for the US. Aside from the small differences in choice of macroeconomic series and sample period outlined in Section 3.1, the approach used to estimate macro-financial spillovers is identical to that used for the US in the preceding subsection.

Let us briefly discuss the total spillover index plots in Figure 4. Similar to the case with the US, the spillover index based on the mixed-frequency analysis is almost always higher than the corresponding index based on the common-frequency approach[6]. While in each country the two indices tend to follow similar trajectories over time, the MF-based index typically captures the most important global financial crisis of 2008 in all six countries, whereas a similar statement cannot be made for the CF-based index. Furthermore, the MF-based index in the European countries (France, Germany and the UK) tend to follow similar trajectories over time indicating the presence of similar forces at play. Likewise, the MF-based index for Canada is quite similar to that of the US.

Next we focus on the decomposition of the MF-based spillover index in all six countries in Figure 5. As it is the case with the US, the vast majority of the total macro-financial spillover indexes in other developed economies are comprised of spillovers originating in financial markets, represented by the bottom four sections of the spillover decomposition plots. The real sector accounts for only a small contribution to the total spillovers between the two sides of the economy.

---

[6]The only exception is the case of Canada in 2000.
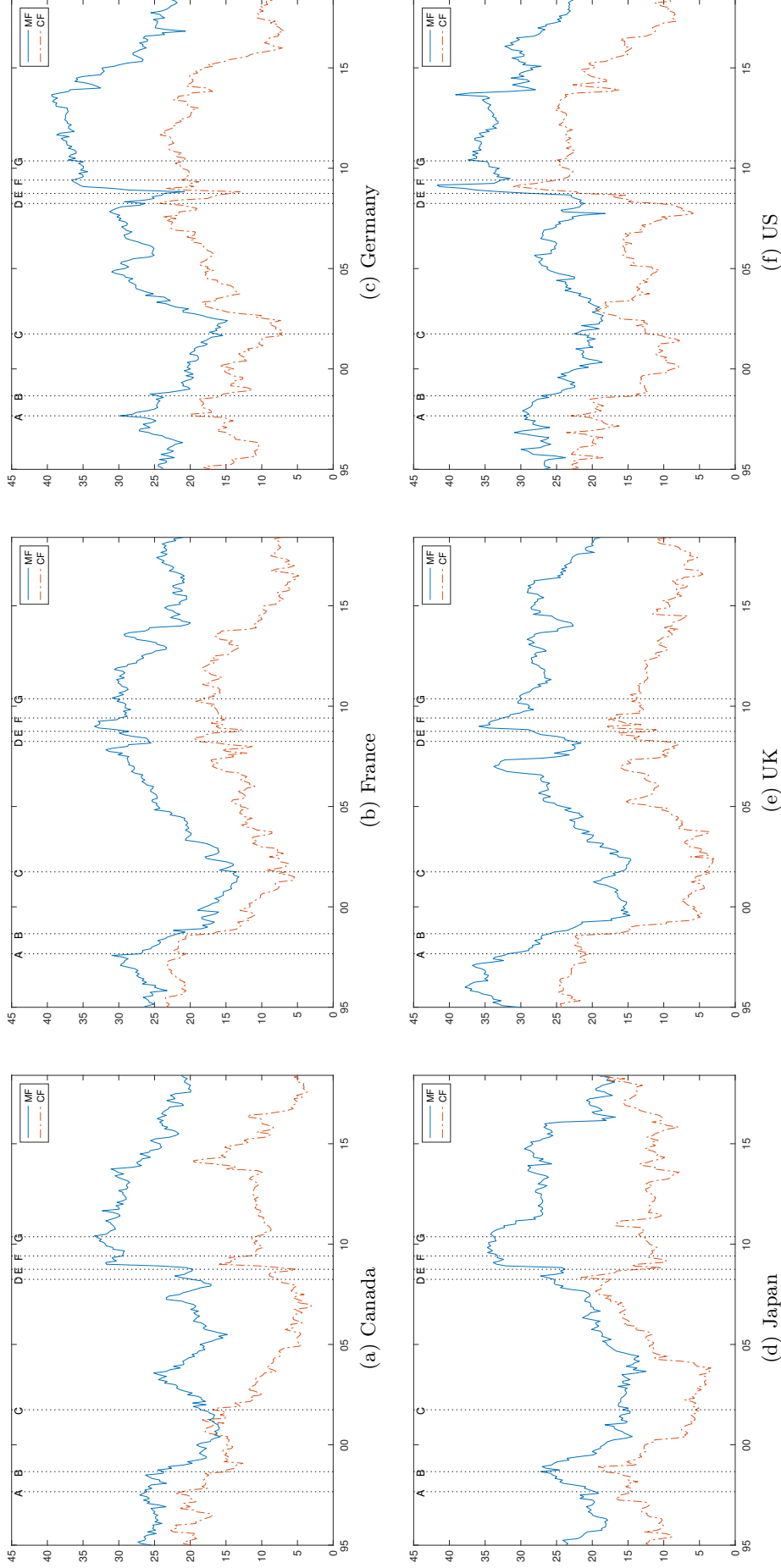
**Figure 4: International total spillover indexes between the financial and real economy series**

The figure presents total spillover indexes for mixed-frequency (denoted MF) and common-frequency (denoted CF) approaches for the sample period 1995:01 to 2018:04. Return levels are employed for the financial series, using a major domestic equity index (S&P/TSX for Canada, CAC40 for France, DAX for Germany, Nikkei 225 for Japan, FTSE100 for the UK and S&P500 for the US) and a generic 10-year domestic government bond index. For the real economy series, month-on-month growth rates of industrial production are employed for all countries. Spillover index values are computed using a 3-month forecast horizon and a 60-month rolling window. Points marked are as follows. A: Asian financial crisis, Jul '97, B: Russian financial crisis and LTCM collapse, Aug to Sept '98, C: September 11, Sept '01, D: collapse of Bear Stearns, Mar '08, E: Lehman Brothers collapse, AIG bailout and Fannie Mae and Freddie Mac being placed in government conservatorship, May '09, F: start of the EU debt crisis in April '10 and flash crash of May '10. Shaded areas correspond to NBER US recession dates.

20
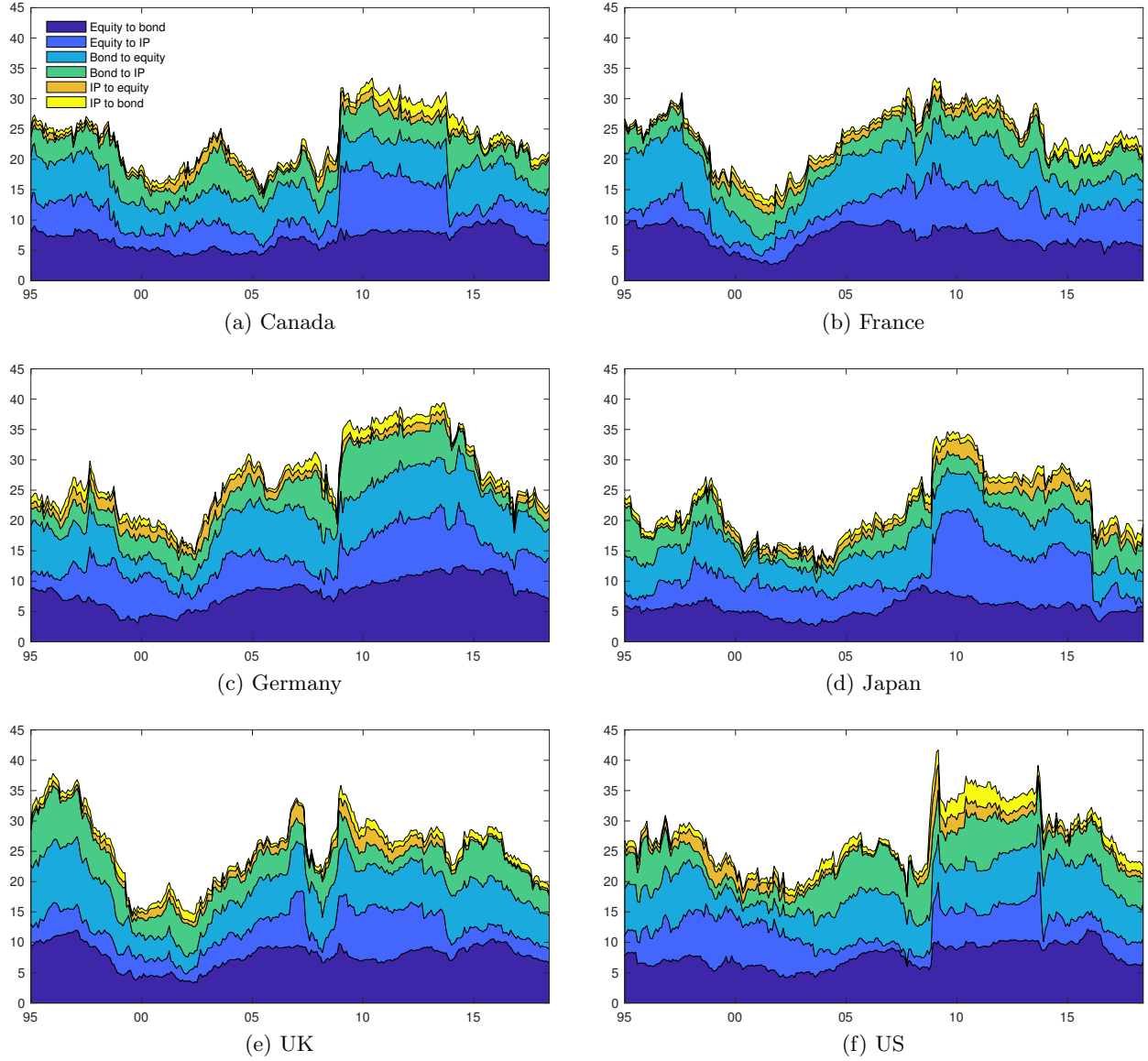
**Figure 5: Decomposition of international total macro-financial spillover indexes into pairwise components**

The figure presents area plots in which the height of the complete shaded area corresponds to the relevant mixed-frequency total spillover index (common-frequency results are available upon request), with each shaded area below representing the contribution of each specific pairwise spillover to the value of total spillover index. Return levels are employed for the financial series, using major domestic equity index (S&P/TSX for Canada, CAC40 for France, DAX for Germany, NIKKEI 225 for Japan, FTSE100 for the UK and S&P500 for the US) and a generic 10-year domestic government bond. For the real economy series, month-on-month growth rates of industrial production are employed for all countries. Spillover index values are computed using a 3-month forecast horizon and a 60-month rolling window.

# 4 Macro-Financial Spillovers as Predictors of Future Macroeconomic Conditions

A reoccurring question of interest in the literature on systemic risk is whether the quantitative measures developed have predictive ability for future macroeconomic conditions. In particular, it is often suggested (see, for example Allen et al., 2012 and references therein) intuitively that an increase in systemic risk may have a negative impact on current and future economic conditions, primarily through a reduction in lending from banks to the non-financial sector. Allen et al. (2012), Giglio et al. (2016) and Brownlees and Engle (2017) amongst others have found empirically that various quantitative measures of systemic risk do have forecasting ability for key macroeconomic series. Motivated by these findings, we now perform a similar exercise to examine whether the current level of macro-financial spillovers also exhibits predictive ability for future macroeconomic conditions.

## 4.1 Forecasting Environment and Predictive Regressions

Our empirical approach closely follows those of Allen et al. (2012) and Brownlees and Engle (2017) and we focus on the problem of forecasting macroeconomic series, specifically the level of the CFNAI and its subcomponents, using predictive regressions of the form:

$$y_{t+n} = \alpha_n + \beta_n s_t + \sum_{i=0}^{q} \gamma_{n,i} y_{t-i} + \delta_n X_t + \epsilon_t \tag{4.1}$$

where $y_t$ is the value of the macroeconomic series of interest, $s_t$ is the value of the chosen macro-financial spillover measure and $X_t$ is a vector of financial control variables commonly used in the literature as simple predictors of future macroeconomic conditions. The vector of control variables $X_t$ consists of the current values of the default spread, the term spread and the return on the S&P 500 equity index. The forecast horizon is denoted by $n$ and we consider forecast horizons one month ($n = 1$) up to a maximum of 12 months ($n = 12$).

Given that both Allen et al. (2012) and Brownlees and Engle (2017) comprehensively investigated the ability of the CATFIN and SRISK measures respectively to forecast future macroeconomic conditions, we also consider forecasting models which include CATFIN or SRISK in place of our macro-financial spillover measures in the relevant predictive regressions[7]. In con-

---

[7]Data for SRISK were kindly provided by the NYU V-Lab (https://vlab.stern.nyu.edu) and those for CATFIN were obtained from Turan Bali's personal website (https://sites.google.com/a/georgetown.edu/turan-bali).

trast to our macro-financial measures that incorporate information from both the financial and real sides of the economy, both CATFIN and SRISK are more traditional financial systemic risk measures. The former measures the aggregate level of systemic risk in the financial system estimated via value-at-risk or expected shortfall, whereas SRISK utilises market and balance sheet data to estimate the expected capital shortfall of financial firms subject to the occurrence of a systemic event.

Following the empirical analysis in Allen et al. (2012) and Brownlees and Engle (2017) respectively, CATFIN enters the predictive regressions in level form and SRISK as a log first difference or growth rate. Data availability constraints for the CATFIN and SRISK measures shorten our sample period for the forecasting exercise relative to the preceding empirical analysis, with forecast performance compared over the period spanning June 2000 to December 2017.

To examine the incremental value of our macro-financial spillover measures and that of the existing SRISK and CATFIN measures for forecasting macroeconomic conditions, we focus on evaluating forecast performance attainable when including these measures relative to that of an otherwise identical benchmark forecasting model that excludes the spillover or systemic risk measures. The form of the predictive regressions for the benchmark model is thus given by (4.1) with the $\beta_n s_t$ term excluded.

## 4.2   Evaluation of Predictive Ability

To formally evaluate forecast performance we use the commonly employed test for equal predictive accuracy of Clark and West (2007), henceforth CW07, which provides a test of the null hypothesis that the forecasts obtained from two (possibly nested) forecasting models perform equally. In all cases we compare the performance of forecasts produced by the models augmented with the various macro-financial spillover and financial systemic risk measures against those from the benchmark model described above. In all cases, rejection of the null hypothesis of equal predictive accuracy implies that the simpler benchmark model is outperformed by the extended model that incorporates spillover or systemic risk measures.

We begin with an in-sample forecasting environment, in which all forecasting models are estimated using data spanning the complete evaluation period of June 2000 to December 2017 and the predictive accuracy of the resulting in-sample forecasts over this period are evaluated. Sample p-values for the CW07 test are reported in Table 1 for both our total and pairwise macro-

**Table 1: In-sample predictive accuracy for macroeconomic conditions**

| Horizon ($n$) | Total | S&P to TNX | S&P to NAI | TNX to S&P | TNX to NAI | NAI to S&P | NAI to TNX | PW mean | PW med | SRISK | CATFIN |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $n = 1$ | 0.409 | 0.251 | 0.450 | 0.199 | 0.489 | 0.009 | 0.023 | 0.000 | 0.002 | 0.178 | 0.172 |
| $n = 2$ | 0.272 | 0.285 | 0.157 | 0.221 | 0.428 | 0.002 | 0.026 | 0.000 | 0.004 | 0.362 | 0.013 |
| $n = 3$ | 0.116 | 0.089 | 0.144 | 0.039 | 0.405 | 0.012 | 0.021 | 0.001 | 0.011 | 0.467 | 0.018 |
| $n = 4$ | 0.137 | 0.086 | 0.212 | 0.039 | 0.464 | 0.008 | 0.012 | 0.001 | 0.003 | 0.383 | 0.007 |
| $n = 5$ | 0.113 | 0.112 | 0.119 | 0.049 | 0.478 | 0.013 | 0.024 | 0.006 | 0.012 | 0.161 | 0.014 |
| $n = 6$ | 0.087 | 0.087 | 0.090 | 0.040 | 0.485 | 0.057 | 0.023 | 0.010 | 0.007 | 0.313 | 0.030 |
| $n = 7$ | 0.086 | 0.086 | 0.084 | 0.036 | 0.384 | 0.232 | 0.029 | 0.018 | 0.010 | 0.139 | 0.007 |
| $n = 8$ | 0.091 | 0.081 | 0.069 | 0.040 | 0.201 | 0.434 | 0.027 | 0.023 | 0.027 | 0.228 | 0.024 |
| $n = 9$ | 0.102 | 0.076 | 0.068 | 0.040 | 0.103 | 0.464 | 0.023 | 0.018 | 0.015 | 0.170 | 0.010 |
| $n = 10$ | 0.134 | 0.090 | 0.077 | 0.053 | 0.049 | 0.374 | 0.026 | 0.022 | 0.019 | 0.039 | 0.002 |
| $n = 11$ | 0.178 | 0.114 | 0.053 | 0.077 | 0.025 | 0.429 | 0.025 | 0.022 | 0.017 | 0.021 | 0.002 |
| $n = 12$ | 0.194 | 0.116 | 0.050 | 0.092 | 0.019 | 0.344 | 0.023 | 0.022 | 0.023 | 0.094 | 0.028 |

The table reports sample p-values for the CW07 test of equal predictive accuracy applied to in-sample $n$-step-ahead forecasts for the level of the CFNAI over the period 2000:06-2017:12. Forecasts are obtained using predictive regressions of the form given in equation (4.1) containing the current value of a single spillover or systemic risk measure, with the exception of the combination forecasts (labelled PW mean and PW med.) that are constructed as discussed in the main text. For pairwise spillovers, equities, bonds and the real economy are denoted by S&P500, TNX and NAI respectively. The null hypothesis in each case is that the forecasts for the relevant model and the benchmark model have equal predictive accuracy, with rejection of the null implying that the extended model has superior predictive accuracy to the benchmark model.

financial spillover measures, and the existing SRISK and CATFIN systemic risk measures. We also include two combination forecasts obtained from the set of pairwise macro-financial spillover measures (columns 'PW mean' and 'PW med'), which are constructed as described below.

Beginning with the total macro-financial spillover index, it can be seen that the overall level of macro-financial spillovers has somewhat mixed predictive ability for macroeconomic conditions as proxied by the CFNAI. It is able to provide an increase in predictive accuracy over the benchmark model that is statistically significant at the 10% level for the intermediate forecast horizons of 6, 7 and 8 months (and borderline significant at $n = 9$), but not for other horizons. The predictive accuracy of the various pairwise spillover measures in the following 6 columns is typically stronger, with a large number of increases in predictive accuracy over the benchmark model that are statistically significant at the 10% or 5% levels. The smaller forecasting gains from the total index relative to the individual pairwise measures indicates that whilst the total spillover index provides a useful single summary measure of spillovers that is conceptually more similar to existing measures in the literature, it is the ability to decompose it into pairwise directional spillovers that makes our DY-based approach valuable in this context. However, the gains in predictive accuracy provided by the pairwise measures clearly vary from one measure to another and across forecast horizons. Typically for the case of pairwise spillovers originating from financial series (the first four of the pairwise measures), gains in predictive ability appear to be concentrated in the intermediate and longer forecast

horizons, the converse is true for pairwise spillovers from the real side of the economy to the equity market, and finally spillovers from the real side of the economy to the bond market provide highly statistically significant gains in predictive accuracy over the benchmark model at all horizons.

The variation in predictive accuracy gains observed across forecasting horizons for the six pairwise spillover measures suggests that there are potential performance improvements to be obtained by combining the informational content of the individual spillover measures. One way to achieve this would be to extend the predictive regression in (4.1) to simultaneously include multiple spillover measures. However, it is frequently found in the literature that this approach of combining multiple predictors results in poorer performing forecasts than combining the forecasts obtained from distinct forecasting models containing different predictors. Examples from other areas of the finance literature include Rapach et al. (2010) who consider the problem of forecasting the equity premium, and Paye (2012) who consider forecasts for equity market volatility using macroeconomic variables. The former also contains a concise discussion of the possible reasons for the strong empirical performance of combination forecasts.

The construction of combination forecasts is made possible by the fact that the DY spillover approach naturally provides a set of directional pairwise measures, each measuring different aspects of the spillover structure. This is in contrast to the majority of existing systemic risk measures, which in their standard form provide only higher level summary measures. Therefore we include two simple combination forecasts obtained as the mean and median of the forecasts[8] obtained from the six pairwise spillover measures, which are denoted by 'PW mean' and 'PW med' respectively in Table 1. Both of these combination forecasts constructed from the set of pairwise spillover measures perform strongly across all forecast horizons, providing improvements in predictive accuracy over the benchmark model that are statistically significant at either the 5% or 1% levels. Given their consistently strong performance, the ability to easily construct these combination forecasts is a significant advantage of the DY approach in this context.

Finally moving on to the predictive ability of the existing SRISK and CATFIN measures, the former fails to produce statistically significant gains over the benchmark model except for the longest horizons of 10, 11 and 12 months. The CATFIN measure on the other hand provides gains in predictive accuracy across all forecast horizons except the shortest one-month horizon

---

[8]More complicated forecast combination methods are possible, such as those that weight the individual forecasts based on past performance. However, in practice these are frequently found to perform similarly to the simple combinations such as the mean (see e.g. Rapach et al., 2010).

that are significant at either the 5% or 1% levels. Although CATFIN has consistently strong in-sample forecasting performance, the p-values for the mean combination forecast constructed from our set of pairwise spillover measures are lower for 7 of the 12 horizons. Therefore, whilst the gains in predictive accuracy over the benchmark model from both SRISK and our pairwise spillover combination forecasts are highly significant, the latter has a small overall advantage in an in-sample forecasting environment.

We next evaluate the out-of-sample forecasting performance of the various models relative to the same benchmark model used above. In all cases out-of-sample forecasts are produced using a standard rolling-window approach with a fixed window length of 60 months. Data within the window are used to estimate the parameters of the relevant predictive regression and produce the $n$-step-ahead forecast, with the model parameters re-estimated each time the window is rolled forward. The out-of-sample combination forecasts are computed as the mean or median of the out-of-sample forecasts obtained from the pairwise spillover measures. Such a pseudo-out-of-sample environment forecasting environment arguably better represents how these measures would potentially be used in practice by an individual attempting to forecast future economic conditions in real time and, as argued by Diebold (2015), is thus an effective approach to assess forecasting performance during different historical periods.

To evaluate out-of-sample forecasting performance we again employ the CW07 test for equal predictive accuracy, but supplement this with the out-of-sample $R^2$ measure of Campbell and Thompson (2008). For a series of out-of-sample forecasts produced for periods $t = 1, \ldots, S$, the out-of-sample $R^2$ measure is computed as:

$$R_{OS}^2 = 1 - \frac{\sum_{t=1}^{S} (y_t - \hat{y}_t)^2}{\sum_{t=1}^{S} (y_t - \hat{y}_t^b)^2}$$

where $y_t$ is the actual value of the series to be predicted, and $\hat{y}_t$ and $\hat{y}_t^b$ are the forecasted values from the model under consideration and the benchmark model respectively. As such, positive (negative) values of $R_{OS}^2$ imply that the forecasting model under consideration has a lower (higher) mean-squared prediction error (MSPE) than the benchmark model.

Out-of-sample forecasting results are presented in Table 2. To conserve space we exclude results for the predictive regressions containing each of the individual pairwise macro-financial

**Table 2: Out-of-sample $R^2$ values and tests of equal predictive accuracy**

| Horizon ($n$) | Total | | PW mean | | PW median | | SRISK | | CATFIN | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $R^2$ | CW | $R^2$ | CW | $R^2$ | CW | $R^2$ | CW | $R^2$ | CW |
| n = 1 | -0.069 | 0.814 | 0.000 | 0.294 | 0.011 | 0.162 | -0.020 | 0.675 | 0.006 | 0.069 |
| n = 2 | -0.017 | 0.135 | 0.037 | 0.034 | 0.029 | 0.060 | -0.020 | 0.797 | 0.000 | 0.105 |
| n = 3 | 0.056 | 0.028 | 0.138 | 0.026 | 0.132 | 0.030 | -0.039 | 0.910 | 0.013 | 0.080 |
| n = 4 | 0.051 | 0.033 | 0.136 | 0.022 | 0.127 | 0.016 | -0.008 | 0.546 | 0.019 | 0.099 |
| n = 5 | 0.087 | 0.030 | 0.147 | 0.029 | 0.112 | 0.017 | -0.006 | 0.357 | -0.012 | 0.165 |
| n = 6 | 0.075 | 0.025 | 0.134 | 0.025 | 0.093 | 0.015 | 0.006 | 0.063 | -0.024 | 0.200 |
| n = 7 | 0.102 | 0.027 | 0.179 | 0.032 | 0.161 | 0.033 | 0.006 | 0.129 | -0.001 | 0.160 |
| n = 8 | 0.104 | 0.019 | 0.211 | 0.035 | 0.207 | 0.042 | 0.010 | 0.054 | -0.035 | 0.362 |
| n = 9 | 0.069 | 0.013 | 0.156 | 0.024 | 0.168 | 0.035 | 0.039 | 0.022 | 0.025 | 0.107 |
| n = 10 | 0.064 | 0.006 | 0.201 | 0.017 | 0.187 | 0.024 | 0.110 | 0.012 | -0.023 | 0.189 |
| n = 11 | 0.073 | 0.004 | 0.228 | 0.009 | 0.196 | 0.013 | 0.105 | 0.031 | -0.024 | 0.114 |
| n = 12 | 0.073 | 0.004 | 0.228 | 0.011 | 0.180 | 0.009 | 0.074 | 0.057 | -0.041 | 0.258 |

The table reports Campbell and Thompson (2008) out-of-sample $R^2$ values and sample p-values for the CW07 test of equal predictive accuracy from predictive regressions for out-of-sample $n$-step-ahead forecasts for the level of the CFNAI over the period 2000:06-2017:12. Forecasts are obtained using predictive regressions of the form given in equation (4.1) containing the current value of a single spillover or systemic risk measure, with the exception of the combination forecasts (labelled PW mean and PW median) that are constructed as discussed in the main text. Out-of-sample forecasts for all models are obtained using a standard rolling window approach with a window length of 60 months. The Campbell and Thompson (2008) out-of-sample R-squared values are computed for each model relative to the benchmark model that excludes spillover or systemic risk measures. For the Clark and West (2007) test the null hypothesis in each case is that the forecasts for the relevant model and the benchmark model have equal predictive accuracy, with rejection of the null implying that the extended model has superior predictive accuracy to the benchmark model.

spillover measures[9] and include only the combination forecasts obtained from these pairwise forecasts.

Beginning with the Campbell and Thompson (2008) out-of-sample $R^2$ values, in almost all cases the forecasts obtained using macro-financial spillover measures result in lower mean squared prediction errors than the benchmark model, the only exceptions being total spillovers for the shortest 1 and 2 month horizons. Most notably, the strong performance of the combination forecasts from the pairwise spillovers continues in the out-of-sample case, dominating the benchmark model at all horizons and displaying larger out-of-sample $R^2$ values than those for the total spillover index in all cases. These predictive gains relative to the benchmark model typically increase as the forecast horizon increases. At the longest 11 and 12 month horizons, the pairwise mean combination forecast results in a mean-squared prediction error 23% lower than that for the benchmark model. In contrast, the predictive performance of SRISK and CATFIN is mixed. Forecasts incorporating SRISK have a larger MSPE than the benchmark model for shorter forecast horizons of less than 6 months but outperform it for longer horizons, whereas for CATFIN the opposite is observed. However, in cases where the MSPE values from the forecasts incorporating SRISK or CATFIN are lower than those from the benchmark model,

---

[9]These results are broadly similar to the previous in-sample case, with the pairwise measures providing statistically significant gains in predictive accuracy over the benchmark model in many cases, but performance varying across horizons and spillover measures.

they are higher than those from the combination forecasts from the pairwise macro-financial spillover measures.

The out-of-sample results for the CW07 test reinforce the strong in-sample predictive accuracy of the pairwise mean and median combination spillover forecasts, with gains over the benchmark model that are significant at the 5% level in all cases except for the shortest 1-month forecast horizon. Interestingly, compared to the previous in-sample forecasting environment, the out-of-sample forecasting performance of the total macro-financial spillover index relative to the benchmark model improves substantially, producing gains that are statistically significant at the 5% level for all forecast horizons longer than 2 months. Indeed, forecasts utilising the total spillover index have lower out-of-sample p-values than any of the other methods for horizons of 10 to 12 months, including the two combination forecasts. Turning finally to the forecasts obtained using the SRISK and CATFIN measures, the CW07 test results largely confirm the general patterns discussed previously in the context of the out-of-sample $R^2$ values. SRISK exhibits statistically significant gains in predictive ability over the benchmark model at longer forecast horizons, being typically competitive with the pairwise combination forecasts, but falling marginally behind forecasts employing the total spillover index. The CATFIN-based forecasts on the other hand provide statistically significant gains only at forecast horizons of 1, 3 and 4 months. This is in contrast to the previous in-sample environment where highly statistically significant improvements over the benchmark model were observed at almost all horizons.

In a similar manner to Giglio et al. (2016) we next investigate the ability of the various spillover and systemic risk measures to forecast the four disaggregated subcomponents of the CFNAI index, which consist of production and income (PI), employment, unemployment and hours (EUH), consumption and housing (CH), and sales, orders and inventories (SOI). This allows us to investigate whether the relative predictive ability of the measures varies according to the specific aspects of macroeconomic conditions that are being forecasted. Table 3 presents p-values for the CW07 test for both in-sample and out-of-sample forecasts constructed using the same approach as employed above. To conserve space only results for a subset of the previous forecasting horizons are presented.

Beginning with the in-sample results, the strong predictive accuracy of the combination spillover forecasts for the aggregate CFNAI index in Table 1 carries through for the subcomponents of the index, with highly statistically significant gains over the benchmark in almost all

**Table 3: In-sample and out-of-sample predictive accuracy for alternative measures of macroeconomic conditions**

| Horizon ($n$) | (1) In-sample | | | | | (2) Out-of-sample | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Total | PW mean | PW med | SRISK | CAT-FIN | Total | PW mean | PW med | SRISK | CAT-FIN |
| | | | | | (a) Production and income (PI) | | | | | |
| n = 1 | 0.394 | 0.001 | 0.003 | 0.165 | 0.108 | 0.869 | 0.050 | 0.045 | 0.508 | 0.042 |
| n = 2 | 0.330 | 0.000 | 0.059 | 0.345 | 0.084 | 0.461 | 0.278 | 0.493 | 0.585 | 0.138 |
| n = 3 | 0.144 | 0.003 | 0.022 | 0.378 | 0.101 | 0.041 | 0.029 | 0.054 | 0.942 | 0.134 |
| n = 6 | 0.073 | 0.015 | 0.014 | 0.292 | 0.402 | 0.017 | 0.017 | 0.028 | 0.040 | 0.111 |
| n = 9 | 0.086 | 0.044 | 0.032 | 0.249 | 0.218 | 0.010 | 0.016 | 0.013 | 0.043 | 0.129 |
| n = 12 | 0.257 | 0.057 | 0.051 | 0.088 | 0.307 | 0.017 | 0.002 | 0.001 | 0.078 | 0.061 |
| | | | | | (b) Employment, unemployment and hours (EUH) | | | | | |
| n = 1 | 0.248 | 0.000 | 0.005 | 0.213 | 0.234 | 0.633 | 0.005 | 0.031 | 0.840 | 0.111 |
| n = 2 | 0.023 | 0.000 | 0.000 | 0.467 | 0.091 | 0.026 | 0.005 | 0.012 | 0.876 | 0.158 |
| n = 3 | 0.006 | 0.000 | 0.000 | 0.276 | 0.307 | 0.009 | 0.006 | 0.015 | 0.772 | 0.105 |
| n = 6 | 0.002 | 0.001 | 0.000 | 0.329 | 0.167 | 0.016 | 0.015 | 0.021 | 0.900 | 0.138 |
| n = 9 | 0.002 | 0.001 | 0.001 | 0.180 | 0.181 | 0.007 | 0.008 | 0.016 | 0.060 | 0.100 |
| n = 12 | 0.001 | 0.002 | 0.009 | 0.073 | 0.297 | 0.005 | 0.004 | 0.008 | 0.031 | 0.132 |
| | | | | | (c) Consumption and housing (CH) | | | | | |
| n = 1 | 0.128 | 0.000 | 0.006 | 0.365 | 0.121 | 0.789 | 0.206 | 0.593 | 0.625 | 0.242 |
| n = 2 | 0.454 | 0.001 | 0.002 | 0.322 | 0.031 | 0.243 | 0.017 | 0.037 | 0.524 | 0.213 |
| n = 3 | 0.299 | 0.007 | 0.008 | 0.447 | 0.026 | 0.030 | 0.006 | 0.009 | 0.913 | 0.031 |
| n = 6 | 0.199 | 0.001 | 0.001 | 0.477 | 0.117 | 0.002 | 0.001 | 0.003 | 0.835 | 0.124 |
| n = 9 | 0.197 | 0.001 | 0.000 | 0.164 | 0.075 | 0.002 | 0.001 | 0.002 | 0.870 | 0.044 |
| n = 12 | 0.285 | 0.000 | 0.001 | 0.303 | 0.261 | 0.008 | 0.001 | 0.001 | 0.596 | 0.086 |
| | | | | | (d) Sales, orders and inventories (SOI) | | | | | |
| n = 1 | 0.324 | 0.002 | 0.094 | 0.385 | 0.122 | 0.635 | 0.301 | 0.426 | 0.976 | 0.053 |
| n = 2 | 0.114 | 0.006 | 0.034 | 0.021 | 0.109 | 0.140 | 0.079 | 0.109 | 0.442 | 0.172 |
| n = 3 | 0.060 | 0.006 | 0.018 | 0.493 | 0.049 | 0.011 | 0.017 | 0.035 | 0.864 | 0.331 |
| n = 6 | 0.070 | 0.016 | 0.011 | 0.400 | 0.136 | 0.025 | 0.023 | 0.039 | 0.907 | 0.318 |
| n = 9 | 0.114 | 0.040 | 0.033 | 0.127 | 0.248 | 0.029 | 0.041 | 0.048 | 0.218 | 0.279 |
| n = 12 | 0.175 | 0.035 | 0.022 | 0.162 | 0.296 | 0.006 | 0.008 | 0.006 | 0.109 | 0.324 |

The table reports sample p-values for the CW07 test of equal predictive accuracy applied to in-sample (columns (1)) and out-of-sample (columns (2)) $n$-step-ahead forecasts for the various subcomponents of the CFNAI index over the period 2000:06-2017:12. CFNAI subcomponents are production and income (PI), employment, unemployment and hours (EUH), consumption and housing (CH), and sales, orders and inventories (SOI). Forecasts are obtained using predictive regressions of the form given in equation (4.1) containing the current value of a single spillover or systemic risk measure, with the exception of the combination forecasts (labelled PW mean and PW med.) that are constructed as discussed in the main text. Out-of-sample forecasts for all models are obtained using a standard rolling window approach with a window length of 60 months. The null hypothesis for the Clark and West (2007) test in each case is that the forecasts for the relevant model and the benchmark model have equal predictive accuracy, with rejection of the null implying that the extended model has superior predictive accuracy to the benchmark model.

cases. In the case of forecasts based on the total macro-financial spillover index, the performance of forecasts for the EUH subcomponent is consistently stronger than that for the aggregate CF-NAI index, that for the CH subcomponent is generally weaker than for the aggregate index and that for the PI and SOI subcomponents display generally similar patterns of performance to those observed previously. This suggests that in the case of some predictors, the relative gains in predictive accuracy depend not only on the forecast horizon, but also on the specific aspect of macroeconomic conditions to be forecasted.

Both SRISK and CATFIN fail to exhibit consistent in-sample predictive power for the CFNAI subcomponents, providing statistically significant improvements over the benchmark approach in only a minority of cases across the various forecast horizons and index subcomponents. For the case of CATFIN, this contrasts sharply to its consistently strong in-sample performance when forecasting the aggregate CFNAI index, particularly given that the pairwise spillover combination forecasts manage to preserve their previous performance for the case of the CFNAI subcomponents.

From columns (2) of Table 3 there is strong evidence of out-of-sample gains in predictive accuracy for the four CFNAI subcomponents using either the total spillover index or the two forecast combination schemes. It is only at the shortest 1 and 2 month forecast horizons that the various forecasts incorporating macro-financial spillover measures fail to outperform the benchmark model at the 5% level, and for the EUH subcomponent the gains from the combination forecasts are statistically significant at all horizons. As with the in-sample case, CATFIN and SRISK only provide statistically significant improvements in out-of-sample predictive accuracy over the benchmark model for the CFNAI in a small number of cases, with their overall performance lagging substantially behind that of the forecasts based on macro-financial spillover measures for all of the CFNAI subcomponents.

## 4.3   Time-variation in Predictive Accuracy

Whilst the CW07 test and the Campbell and Thompson (2008) out-of-sample $R^2$ measures are both employed widely in the literature to compare predictive accuracy between competing forecasting models, they do not directly reveal anything about how the relative performance of the alternative models may vary over the forecast evaluation period. In particular, it may be the case that specific predictors of future macroeconomic conditions have greater predictive ability during some time periods than others, depending on the state of the economy.

In order to shed light on this and other related issues, we construct plots of the cumulative squared prediction errors from predictive regressions for the $n$-period-ahead values of the CFNAI relative to that of the benchmark model. Such plots are a standard tool in the forecasting literature, with two examples including Rapach et al. (2010) and Paye (2012). We subtract the values of the cumulative squared prediction errors obtained for the predictive regression of interest over the in-sample or out-of-sample period from those obtained from the previous benchmark model that excludes all spillover and systemic risk measures. Values larger (smaller) than one imply that the forecasting model in question has a lower (higher) cumulative squared prediction error than the benchmark model up to that point in time. An increase (decrease) during a specific time period implies that the relevant model is currently outperforming (underperforming) the benchmark model.

Figure 6 plots the difference between the cumulative squared prediction errors (CSPEs) for the in-sample forecasts obtained using the various spillover and systemic risk measures and those from the benchmark model. To conserve space we only include results for the total spillover index, the two combination forecast constructed from the pairwise spillovers, SRISK and CATFIN, and include only a subset of the forecasting horizons. For consistency with the preceding analysis of predictive accuracy, plots are produced for the same forecast evaluation periods as used above. However, it should again be noted that the vertical position or height of each line at a given point is affected by the chosen start date for each plot (given that all must equal zero at the start of the evaluation period). The shape or gradient of each line is however unaffected by this choice.

For the 1 and 2 month horizons, the forecasting performance of the total spillover index remains close to that of the benchmark model throughout the evaluation period, deviating only slightly from zero. For horizons of greater than 3 months forecasts obtained from the total spillover index perform worse than the benchmark model in early parts of the evaluation period, with the relative CSPE values becoming steadily more negative until around 2006. However, after the start of the financial crisis in 2008 there is a large and sharp increase in the relative CSPE until the peak of the crisis had subsided in around 2010, suggesting large gains in predictive accuracy for the total spillover index during the crisis. Even more substantial gains in predictive accuracy are typically observed for the pairwise mean combination forecasts during the crisis period at the 3 to 12 month forecast horizons. However, unlike the forecasts obtained from the total spillover index, they typically outperform the benchmark over the pre-crisis
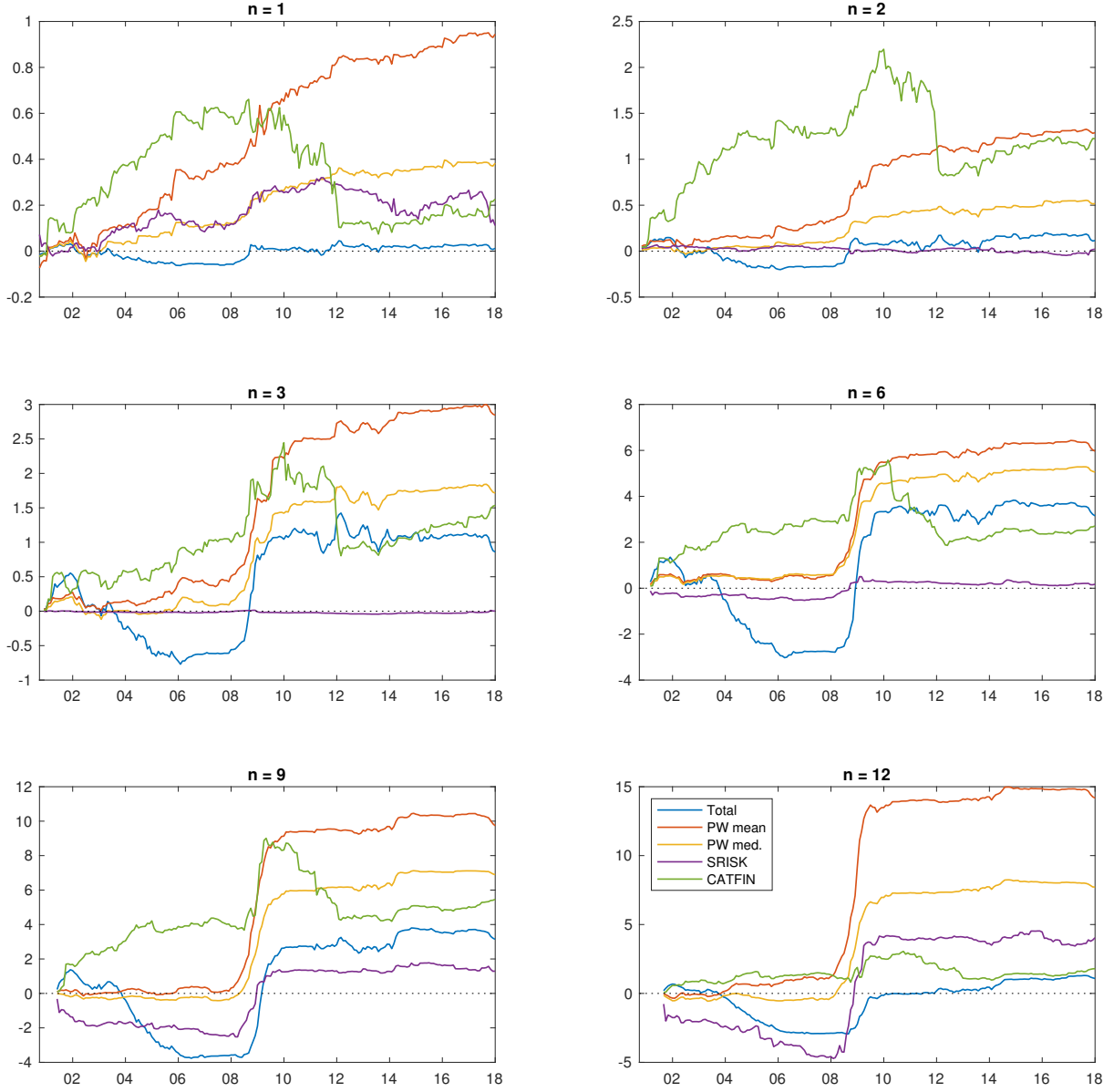
**Figure 6: In-sample cumulative squared forecast errors relative to the benchmark model**
The figure plots cumulative squared in-sample forecast errors of the benchmark predictive regression model minus the cumulative squared forecast errors of the predictive regression models that include the total macro-financial spillover index, the combination forecasts computed as the mean and median of the pairwise spillover index forecasts, SRISK or CATFIN. Larger values correspond to stronger performance of the relevant extended predictive regression model relative to the benchmark, with values above (below) zero implying a smaller (larger) cumulative squared forecast error than the benchmark model. The dependent variable to be forecasted in all cases is the $n$-period-ahead level of the CFNAI.

period too, resulting in substantially higher relative CSPEs when considered over the complete evaluation period. Finally, the forecasts obtained from the pairwise median combination forecast typically lies in between that of the the total spillover index and the mean combination forecast.

The performance of the SRISK-based forecasts is nearly identical to that of the benchmark model throughout the evaluation period for the intermediate 2, 3 and 6 month forecast horizons. At the longer 9 and 12 month horizons, the dynamics of forecast performance are somewhat similar to those of the forecasts based on the total spillover index, with performance inferior to the benchmark in the pre-crisis period, superior to it during the crisis and comparable to it in the post-crisis period. Considering the CATFIN-based forecasts, we generally observe strong performance at the start of the evaluation period followed by a further increase during the crisis period. However, this peak coinciding with the 2008 crisis is substantially less pronounced than that exhibited by the forecasts employing macro-financial spillover measures, and furthermore is typically followed by a decrease in performance relative to the benchmark model.

Figure 7 presents equivalent plots for the out-of-sample case, with all out-of-sample forecasts produced using the same rolling window approach discussed above. In the out-of-sample forecasting environment the differences in cumulative squared forecast errors for the augmented models compared to the benchmark model tend to be quite small at the beginning of the evaluation period, however from 2008 onwards their performance diverges substantially. In most cases the forecasts based on the total spillover index, the pairwise forecast combinations and those based on CATFIN again exhibit clear increases in predictive accuracy relative to the benchmark model during the 2008 crisis, though the absolute and relative sizes of these increases varies substantially across forecast horizons. In the case of the combination forecasts these gains continue to accumulate gradually to the end of the sample period. By contrast, in the case of forecasts based on the total spillover index and CATFIN, some or all of the gains in cumulative forecast accuracy over the benchmark model attained during the crisis are generally lost in the post-crisis period.

## 5   Conclusion

We estimate and analyse the structure of macro-financial spillovers between equities, bonds and the real side of the economy. For this purpose we develop a new methodology for estimating macro-financial spillovers that combines established quantitative measures of financial spillovers
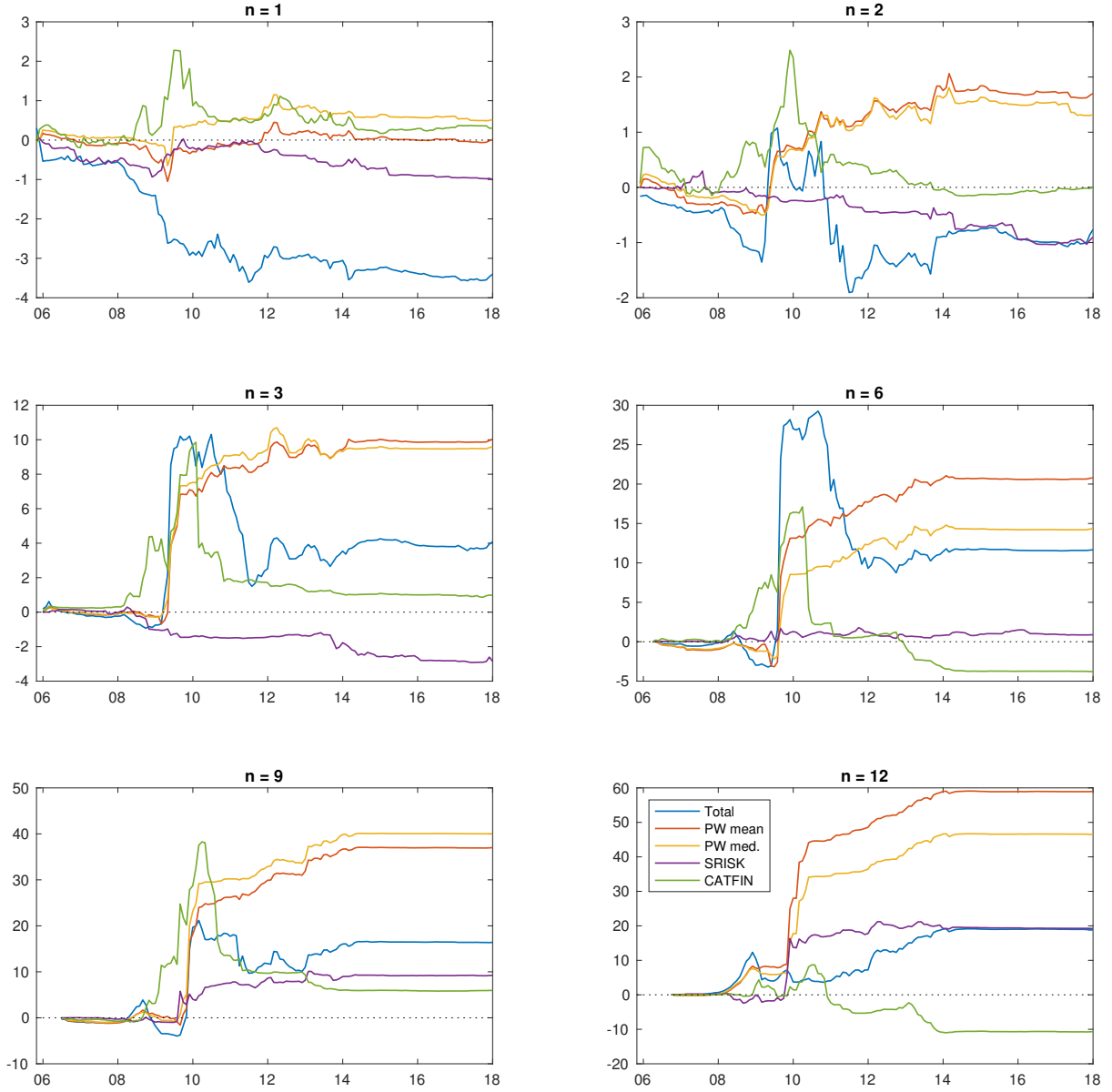
**Figure 7: Out-of-sample cumulative squared forecast errors relative to the benchmark model**
The figure plots cumulative squared out-of-sample forecast errors of the benchmark predictive regression model minus the cumulative squared forecast errors of the predictive regression models that include the total macro-financial spillover index, the combination forecast computed as the mean and median of the pairwise spillover index forecasts, SRISK or CATFIN. Larger values correspond to stronger performance of the relevant extended predictive regression model relative to the benchmark, with values above (below) zero implying a smaller (larger) cumulative squared forecast error than the benchmark model. The dependent variable to be forecasted in all cases is the $n$-period-ahead level of the CFNAI.

with mixed-frequency econometric methods. Our approach permits the use of mixed-frequency macro-financial datasets without the need to aggregate the higher frequency financial series down to the lower frequency as the macroeconomic series. The methodology produces a set of different macro-financial spillover measures that take into account of the direction of spillovers. The directionality of the measures obtained permits more detailed analysis of market linkages than other approaches that measure only association.

In our analysis of macro-financial spillovers in the US economy from 1975 to 2018 we find that the magnitude of the mixed-frequency spillovers are substantially greater than those obtained from an analogous common-frequency approach. This suggests that the loss of high-frequency information incurred by the use of a common-frequency modelling approach results in the financial and real sides of the economy appearing less connected. Furthermore, the preservation of additional high-frequency information by our mixed-frequency approach results in spillover measures that appear more consistent with key events that occurred during our sample period. The same empirical findings are also evident for 5 other advanced economies.

The directional nature of our spillover measures allows us to decompose overall macro-financial spillovers into pairwise subcomponents and demonstrate that the largest magnitude of spillovers originates from the financial, rather than the real, side of the economy. This decomposition also clearly shows that the relative importance of each financial market has changed over time. Fluctuations in the pairwise spillovers from equity and bond markets to the real side of the economy are shown to be consistent with significant events occurring during the sample period analysed. Again, these empirical findings too carry through to other advanced economies.

Motivated by existing work analysing the predictive ability of financial systemic risk measures for future macroeconomic series, we explore whether our macro-financial spillover measures can be employed to forecast US macroeconomic conditions. We find that forecasts produced using our spillover measures predict both broad measures of overall macroeconomic conditions and also measures representing more specific aspects of the state of the macroeconomy. Forecasting performance is particularly strong for the case of simple combination forecasts obtained as the mean or median of the individual forecasts obtained from our set of pairwise macro-financial spillover measures. In an in-sample forecasting environment, these combination forecasts marginally outperform existing systemic risk measures when forecasting aggregate macroeconomic conditions. However, they substantially outperform them when forecasting dis-

aggregated measures representing more specific aspects of economic conditions. When moving to an out-of-sample forecasting environment the gains in predictive accuracy of our macro-financial spillover are typically even larger. When examining the dynamics of forecasting performance, we find that the gains in predictive accuracy provided by our spillover measures are especially large during the 2008-2009 crisis, both relative to our chosen benchmark model and also relative to those provided by existing financial systemic risk measures.

# References

Adams, Z., R. Füss, and R. Gropp. 2014. Spillover Effects among Financial Institutions: A State-Dependent Sensitivity Value-at-Risk Approach. *Journal of Financial and Quantitative Analysis* 49:575–598.

Adrian, T., and M. K. Brunnermeier. 2016. CoVaR. *American Economic Review* 106:1705–1741.

Allen, L., T. G. Bali, and Y. Tang. 2012. Does Systemic Risk in the Financial Sector Predict Future Economic Downturns? *Review of Financial Studies* 25:3000–3036.

Almeida, H., and M. Campello. 2007. Financial Constraints, Asset Tangibility, and Corporate Investment. *Review of Financial Studies* 20:1429–1460.

Andersen, T., T. Bollerslev, F. Diebold, and C. Vega. 2003. Micro effects of macro announcements: Real-time price discovery in foreign exchange, typescript. *American Economic Review* 93:38–62.

Baur, D. G. 2012. Financial contagion and the real economy. *Journal of Banking and Finance* 36:2680–2692.

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

Brenner, M., P. Pasquariello, and M. Subrahmanyam. 2009. On the Volatility and Comovement of U.S. Financial Markets around Macroeconomic News Announcements. *Journal of Financial and Quantitative Analysis* 44:1265.

Brownlees, C. T., and R. F. Engle. 2017. SRISK: A Conditional Capital Shortfall Measure of Systemic Risk. *Review of Financial Studies* 30:48–79.

Brunnermeier, M. K., G. Gorton, and A. Krishnamurthy. 2011. Risk topography. *NBER Macroeconomics Annual* 26:149–176.

Campbell, J. Y., and S. B. Thompson. 2008. Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average? *Review of Financial Studies* 21:1509–1531.

Cingano, F., F. Manaresi, and E. Sette. 2016. Does Credit Crunch Investment Down? New Evidence on the Real Effects of the Bank-Lending Channel. *Review of Financial Studies* 29:2737–2773.

Claessens, S., H. Tong, and S.-J. Wei. 2012. From the financial crisis to the real economy: Using firm-level data to identify transmission channels. *Journal of International Economics* 88:375–387.

Clark, T. E., and K. D. West. 2007. Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics* 138:291–311.

Diebold, F. X. 2015. Comparing Predictive Accuracy, Twenty Years Later: A Personal Perspective on the Use and Abuse of Diebold–Mariano Tests. *Journal of Business and Economic Statistics* 33:1–9.

Diebold, F. X., and K. Yilmaz. 2014. On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics* 182:119–134.

Dungey, M., M. Luciani, and D. Veredas. 2013. Googling SIFIs. *Working Paper* .

Engle, R. F., E. Jondeau, and M. Rockinger. 2015. Systemic Risk in Europe. *Review of Finance* 19:145–190.

Ghysels, E. 2016. Macroeconomics and the reality of mixed frequency data. *Journal of Econometrics* 193:294–314.

Giglio, S. W., B. Kelly, and S. Pruitt. 2016. Systemic risk and the macroeconomy: An empirical evaluation. *Journal of Financial Economics* 119:457–471.

Green, T. C. 2004. Economic news and the impact of trading on bond prices. *The Journal of Finance* 59:1201–1233.

Ivashina, V., and D. Scharfstein. 2010. Bank lending during the financial crisis of 2008. *Journal of Financial Economics* 97:319–338.

Parkinson, M. 1980. The Extreme Value Method for Estimating the Variance of the Rate of Return. *The Journal of Business* 53:61–65.

Paye, B. S. 2012. 'Déjà vol': Predictive regressions for aggregate stock market volatility using macroeconomic variables. *Journal of Financial Economics* 106:527–546.

Pesaran, M. H., and Y. Shin. 1998. Generalized impulse response analysis in linear multivariate models. *Economics Letters* 58:17–29.

Rapach, D. E., J. K. Strauss, and G. Zhou. 2010. Out-of-Sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy. *Review of Financial Studies* 23:821–862.

Schorfheide, F., and D. Song. 2015. Real-Time Forecasting With a Mixed-Frequency VAR. *Journal of Business and Economic Statistics* 33:366–380.

# Appendices

## A   Forecast Error Variance Decomposition

### A.1   Generalised Forecast Error Variance Decomposition

We denote the generalised FEVD values by $\theta_{ij}(H)$, where $\theta_{ij}(H)$ measures the fraction of the total $H$-step-ahead error variance in forecasting series $i$ attributable to shocks in series $j$. Following Pesaran and Shin (1998), the generalised forecast error variance decomposition values are computed for any given forecast horizon $H = 1, 2, \ldots$ as:

$$\theta_{ij}(H) = \frac{\sigma_{jj} \sum_{h=0}^{H-1} \left(e_i' B_h \Sigma e_j\right)^2}{\sum_{h=0}^{H-1} \left(e_i' B_h \Sigma B_h' e_i\right)} \quad \text{for } i, j = 1, \ldots, K_x \tag{A.1}$$

where $\Sigma$ is the covariance matrix of the error vector $\underline{\varepsilon}(\tau_L)$, $\sigma_{jj}$ is the $j$-th diagonal element of $\Sigma$ and $e_j$ is the $K_x$-dimensional selection vector with a 1 in the $j$-th element and zeros elsewhere. The arrays $B_i, i = 1, \ldots$ are the coefficient arrays from the infinite order moving average (MA) representation of the MF-VAR in equation (2.1). These can be obtained from the coefficient

arrays of the standard representation of the VAR via a simple recursion (see Diebold and Yilmaz, 2014 for details).

It is worth noting that unlike other common approaches to computing the FEVD that rely on orthogonalisation to account for potential correlation between shocks, such as the Cholesky decomposition, the values of the generalised FEVD arrays are not affected by the ordering of the series within the VAR. Instead, the approach accounts for potential correlation between shocks using the historical distribution of the errors.

## A.2 Transformation of MF-VAR FEVD Arrays

We continue to employ the simple bivariate example from Section 2.3 for illustration, repeating some of the key details here for convenience. With one low-frequency monthly series and one high-frequency weekly series we obtain a $(5 \times 1)$ vector process with the form $\underline{x}(\tau_L) = [x_H(\tau_L, 1), \dots, x_H(\tau_L, 4), x_L(\tau_L)]'$ for the MF-VAR, and a $(2 \times 1)$ vector process $\overline{x}(\tau_L) = [x_{HtL}(\tau_L), x_L(\tau_L)]'$ for the corresponding CF-VAR. This results in $(5 \times 5)$ and $(2 \times 2)$ FEVD arrays, given respectively by:

$$
\begin{bmatrix} \theta_{11}(H) & \dots & \theta_{15}(H) \\ \vdots & \ddots & \vdots \\ \theta_{51}(H) & \dots & \theta_{55}(H) \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \phi_{11}(H) & \phi_{12}(H) \\ \phi_{21}(H) & \phi_{22}(H) \end{bmatrix} \quad \text{for} \quad H = 1, 2, \dots \qquad \text{(A.2)}
$$

We argued previously that the elements of the MF-VAR FEVD arrays can be grouped into sub-arrays as:

$$
\begin{bmatrix} \Theta_{11}(H) & \Theta_{12}(H) \\ \Theta_{21}(H) & \Theta_{22}(H) \end{bmatrix} \quad \text{for} \quad H = 1, 2, \dots \qquad \text{(A.3)}
$$

where:

$$
\Theta_{11}(H) \equiv \begin{bmatrix} \theta_{11}(H) & \dots & \theta_{14}(H) \\ \vdots & \ddots & \vdots \\ \theta_{41}(H) & \dots & \theta_{44}(H) \end{bmatrix} \qquad \Theta_{12}(H) \equiv \begin{bmatrix} \theta_{15}(H) \\ \vdots \\ \theta_{45}(H) \end{bmatrix}
$$

$$
\Theta_{21}(H) \equiv \begin{bmatrix} \theta_{51}(H) & \dots & \theta_{54}(H) \end{bmatrix} \qquad \Theta_{22}(H) \equiv \theta_{55}(H)
$$

such that each of the sub-arrays $\Theta_{kl}(H)$ in (2.6) can be viewed as a mixed-frequency analogue

of the corresponding scalar element $\phi_{kl}(H)$ from the CF-VAR FEVD array. Our transformation approach produces new FEVD arrays from the MF-VAR FEVD arrays with the same structure and dimensions as those for the corresponding CF-VAR. We denote a generic element of the new transformed FEVD arrays by $\psi_{kl}(H)$ for $k, l = 1, \ldots, K$.

The key is to perform the transformation such that the value and interpretation of each element $\psi_{kl}(H)$ is directly comparable with the corresponding element $\phi_{kl}(H)$. This relies on the correspondence between the elements of the mixed and common frequency arrays discussed above and the mathematical definition of the generalised FEVD elements in equation (A.1). For ease of notation, we denote the numerator and denominator of (A.1) more compactly as:

$$\theta_{ij}(H) = \frac{\lambda_{ij}(H)}{\mu_i(H)} \quad \text{for } i, j = 1, \ldots, K_x \tag{A.4}$$

where:

$$\lambda_{ij}(H) \equiv \sigma_{jj} \sum_{h=0}^{H-1} \left( e_i' B_h \Sigma e_j \right)^2 \quad \text{and} \quad \mu_i(H) \equiv \sum_{h=0}^{H-1} \left( e_i' B_h \Sigma B_h' e_i \right)$$

The denominator $\mu_i(H)$ corresponds to the total $H$-step-ahead forecast error variance for series $i$ and the numerator is the forecast error variance for series $i$ due to shocks in series $j$ (normalised such that the shock is one standard deviation in size).

We compute each element $\psi_{kl}(H)$ in the transformed FEVD array as:

$$\psi_{kl}(H) = \frac{\sum_{i \in \mathcal{I}_k, \, j \in \mathcal{J}_l} \lambda_{ij}(H)}{\sum_{i \in \mathcal{I}_k} \mu_i(H)} \quad k, l = 1, \ldots K, \ H = 1, 2, \ldots \tag{A.5}$$

where $\mathcal{I}_k$ and $\mathcal{J}_l$ are sets containing the row and column indexes respectively for the elements in the MF-VAR FEVD array that correspond to the element $\phi_{kl}(H)$ in the sense discussed above.

For the previous bivariate example, the elements in the MF-VAR FEVD array that correspond to $\phi_{kl}(H)$ are those contained in the sub-array $\Theta_{kl}(H)$ in equation (A.3). For example, for $k = 1, l = 1$, we have $\mathcal{I}_1 = \{1, \ldots, 4\}$, $\mathcal{J}_1 = \{1, \ldots, 4\}$ (the elements of $\Theta_{11}(H)$) and thus:

$$\psi_{11}(H) = \frac{\sum_{i \in \mathcal{I}_1, \, j \in \mathcal{J}_1} \lambda_{ij}(H)}{\sum_{i \in \mathcal{I}_1} \mu_i(H)} = \frac{\sum_{i=1, j=1}^{4} \lambda_{ij}(H)}{\sum_{i=1}^{4} \mu_i(H)} \quad H = 1, 2, \ldots$$

Likewise, for $k = 2, l = 1$ we find $\mathcal{I}_2 = \{5\}$, $\mathcal{J}_1 = \{1, \ldots, 4\}$ (the elements of $\Theta_{21}(H)$) and thus:

$$\psi_{21}(H) = \frac{\displaystyle\sum_{i \in \mathcal{I}_2,\, j \in \mathcal{J}_1} \lambda_{ij}(H)}{\displaystyle\sum_{i \in \mathcal{I}_2} \mu_i(H)} = \frac{\displaystyle\sum_{j=1}^{4} \lambda_{5j}}{\mu_5(H)} \qquad H = 1, 2, \ldots$$

# B   Data Appendix

## B.1   Construction of Weekly Financial Series

The MF-VAR approach of Ghysels (2016) is applicable in situations where the number of high-frequency time periods per low-frequency period varies deterministically over time. This does however somewhat complicate the implementation of the method. As discussed in the main text, we employ a combination of monthly macroeconomic and weekly financial time series for the empirical analysis. Clearly if working with monthly and weekly time intervals in a traditional sense, the number of weeks per calendar month varies from month to month.

To avoid the complications introduced by deterministic time variation in the number of weeks per calendar month, we pre-process the data and work with what we term 'pseudo weeks' rather than standard calendar weeks. This approach is possible because we observe all financial series at a daily frequency that is higher than the final desired weekly frequency. These pseudo weeks are constructed by dividing the trading days within each month into 4 sub-periods whose lengths vary, but are as close as possible to being equal. For example, months with 20 trading days are divided into four 5-day sub-periods, those with 19 trading days are divided into three 5-day periods and a 4-day period, those with 22 days are split into two 5-day periods and two 6-day periods and so on. The vast majority of pseudo-weeks contain either 5 or 6 trading days, however Februarys or months with an unusually large number of weekday non-trading days due to holidays may contain one or more weeks with 4 trading days.

With the exception of months containing exactly 20 trading days (that are always split into 4 weeks of equal length), the way in which the weeks are ordered within a given month will clearly influence the values (e.g. returns) obtained. For example, months with 21 working days can be split into four pseudo-weeks with lengths 5-5-5-6, with lengths 5-5-6-5, with lengths 5-6-5-5 or with lengths 6-5-5-5, each of which will produce different final values for prices, returns and return volatilities. Therefore to avoid this issue we compute values over all possible split orders for a given month and then average the resulting values. Finally, while computing pseudo-

weekly returns or return volatilities for each period we also adjust the return and volatility values obtained to account for the fact that the length of the return period actually differs slightly from one pseudo-week to another (4, 5 or 6 trading days).

## B.2 Plots of Data Series

The US financial and macroeconomic series used to estimate spillover measures are plotted for the full sample period in Figure A.1. Major economic and financial events during the sample period are clearly visible in the plots, either as substantial increases in financial volatility or large changes in the CFNAI; examples include the 1980-1981 recession, the Asian and Russian financial crisis and the collapse of Long-Term Capital Management in the late 1990's, the dotcom bubble and 9/11 in the early 2000's and the recent global financial crisis in the late 2000's and early 2010's.
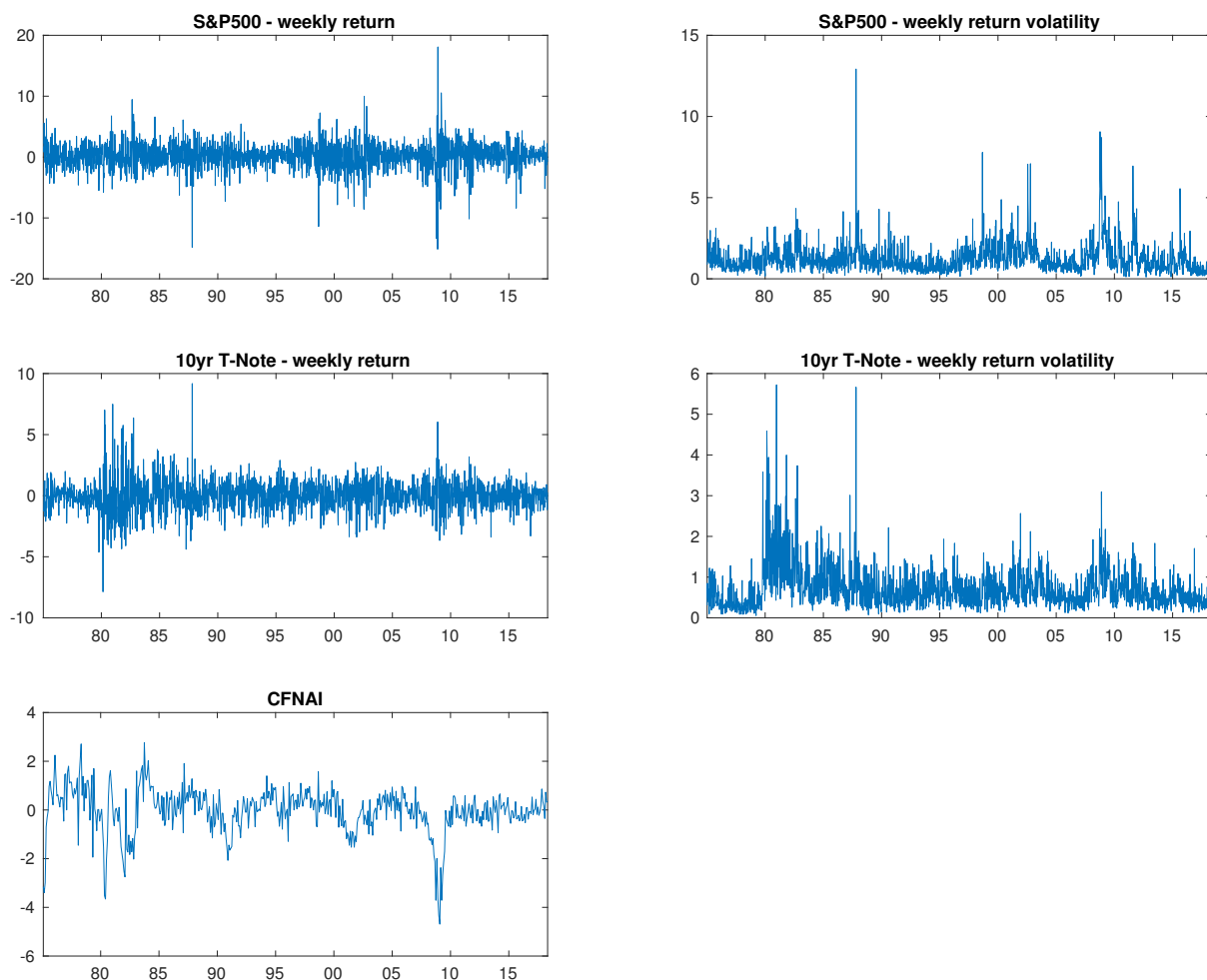
**Figure A.1: Time series plots of financial and real economy series**
The financial (S&P500 and 10-year Treasury Bond) and real economy (CFNAI) series are plotted for the full sample period 1975:01 to 2015:09. Returns and return volatilities are expressed in percentage terms for the weekly frequency, with standard deviations plotted for the latter constructed using a range-based approach detailed in Appendix C.1. The CFNAI series is plotted in level form.

# C   Supplementary Empirical Results

## C.1   Total Spillover Indexes for Logarithmic Return Volatilities

We also computed our mixed-frequency macro-financial spillover index using logarithmic return volatilities in place of return levels for the financial series as in Diebold and Yilmaz (2014). These previous studies employ the range-based estimator of return volatility proposed by Parkinson (1980), which estimates return volatility from the high and low prices during the chosen return period. As daily data on high and low prices are not available during the earlier parts of the sample period we approximate this estimator by replacing the high and low prices with the highest and lowest daily closing prices observed during each week. For the later parts of the sample period where both estimators can be computed, we confirm that our volatility measure using only close prices is highly correlated with the Parkinson (1980) estimator, with a correlation coefficient of just over 0.9 for the S&P500 and just under 0.9 for the 10-year Treasury Note series.

Figure A.2 presents an equivalent plot to Figure 1, for the case where logarithmic return volatilities are employed instead of return levels for financial series. The total spillover index series display broadly similar dynamics to those observed for the case of return levels. For the current case of return volatilities the spillover indexes typically display slightly more pronounced spikes than for return levels around adverse financial or economic events (see in particular the Asian and Russian financial crises). This is due to the fact that financial volatility spillovers typically rises in turbulent times and drops in tranquil times, whereas return spillovers may increase in both situations. Finally, we again observe similar differences in the level of macro-financial spillovers between the mixed-frequency and common-frequency cases, with the former implying a higher average level of spillovers over the sample period.
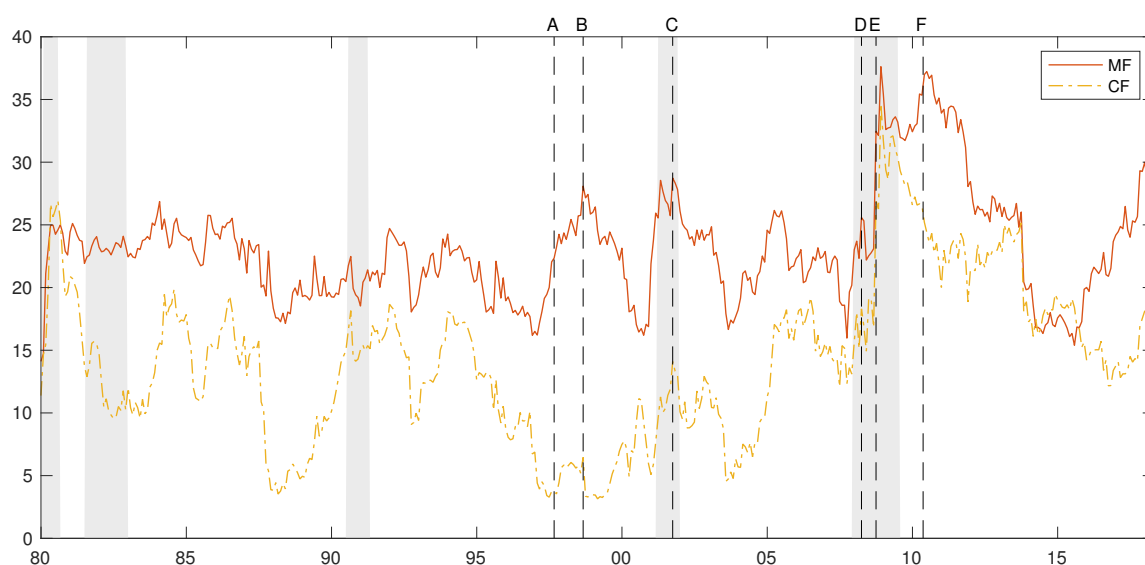
**Figure A.2: Total spillover indexes between the financial and real economy series obtained from logarithmic return volatiities**

Total spillover indexes for mixed-frequency (denoted MF) and common-frequency (denoted CF) approaches are presented for the sample period 1980:01 to 2018:04. Logarithmic return volatilities are employed for the financial S&P500 and 10-year Treasury Note series, levels for the real economy CFNAI series. Values are computed using a 3-month forecast horizon and a 60-month rolling window. Points marked are as follows. A: Asian financial crisis, Jul '97, B: Russian financial crisis and LTCM collapse, Aug to Sept '98, C: September 11, Sept '01, D: collapse of Bear Stearns, Mar '08, E: Lehman Brothers collapse, AIG bailout and Fannie Mae and Freddie Mac being placed in government conservatorship, May '09, F: start of the EU debt crisis in April '10 and flash crash of May '10. Shaded areas correspond to NBER US recession dates.