Persistence in Financial Connectedness and Systemic Risk*

Jozef Baruník‡ and Michael Ellington†

September 20, 2023

Abstract

This paper characterises dynamic linkages arising from shocks with heterogeneous degrees of persistence. Using frequency domain techniques, we introduce measures that identify smoothly varying links of a transitory and persistent nature. Our approach allows us to test for statistical differences in such dynamic links. We document substantial differences in transitory and persistent linkages among US financial industry volatilities, argue that they track heterogeneously persistent sources of systemic risk, and thus may serve as a useful tool for market participants.

JEL Classifications: C10, C40, C55, C58, G00
Keywords: Finance, Network connections, Variance Decompositions, Persistence, Spectral Domain.

‡*Institute of Economic Studies, Charles University, Opletalova 26, 110 00, and The Czech Academy of Sciences, IITA, Pod Vodárenskou Věží 4, 182 08, Prague, Czech Republic.*
barunik@fsv.cuni.cz
†*University of Liverpool Management School, Chatham Building, Liverpool, L69 7ZH, UK.*
m.ellington@liverpool.ac.uk

*We thank Oliver Linton, Wolfgang Härdle, Melanie Schienle, Catherine Forbes, Lubos Hanus, Lukáš Vácha, Ryland Thomas, Chris Florackis, Costas Milas, Alex Kostakis and Charlie Cai for invaluable discussions and comments. We are grateful to Luboš Hanus for help in furnishing and converting estimation codes. We acknowledge insightful comments from numerous seminar presentations, such as: the 2019 and 2020 Society for Economic Measurement Conferences; the Danish National Bank; the 2019 STAT of ML conference; the 13th International Conference on Computational and Financial Econometrics; and many more. Jozef Baruník gratefully acknowledges support from the Czech Science Foundation under the 19-28231X (EXPRO) project. For estimation of dynamic horizon specific networks, we provide packages `DynamicNets.jl` in JULIA and `DynamicNets` in MATLAB. The packages are available at https://github.com/barunik/DynamicNets.jl and https://github.com/ellington/DynamicNets. **Disclosure Statement:** Jozef Baruník and Michael Ellington have nothing to disclose.

1 Introduction

Firms and economic units create connections through a variety of channels (see e.g. Richmond, 2019; Garvey et al., 2015; Herskovic et al., 2020).[1] These connections are dynamic and vary over time with changing states of an economy, as both stable and uncertain periods are associated with different shocks. At the same time, an increasing number of authors argue that economic variables are driven by shocks that influence their future value with heterogeneous levels of persistence (Bandi et al., 2021; Dew-Becker and Giglio, 2016). Connectedness and the subsequent network structures that emerge from these relationships are central to risk measurement and management, as well as to understanding macroeconomic risks that emerge over business cycles.

The main objective of this paper is to introduce measures capable of identifying the smoothly varying persistence structure of such linkages. As an interesting and important example, we provide an analysis of US sectors, with a focus on the financial sector, and identify a heterogeneous persistence structure of systemic risk.

There are a variety of ways to measure connectedness.[2] In particular, Diebold and Yilmaz (2014) provides a unifying framework for measuring connectedness and other network properties using variance decompositions from an approximate model. Variance decompositions track how shocks affect the future variation of variables within a system, and are therefore a natural choice for inferring network connectedness from data.[3] However, the time evolution of such measures typically involves the estimation of static models that require covariance stationarity and roll through a time series of data. More importantly, such measures aggregate shocks and mask the persistence structures of a network.

In this paper, we provide the tools to identify dynamic connectedness and other key network measures from variance decompositions. We argue that different levels of connectedness can form around transitory and persistent components of shocks within a system of financial data. To capture the time-varying nature of such transitory and persistent components of connectedness, we further consider time-varying variance decomposition matrices from vector autoregressions (VARs) as dynamic adjacency matrices. To identify the persistence structure of the network, we propose to use localised spectral decompositions[4] of variance error forecasts.

The main contribution of our paper is to provide a novel framework for measuring

[1] Richmond (2019) measures connections through consumption growth. Garvey et al. (2015) tracks connections that describe the supply chain, and Herskovic et al. (2020) studies network connections between firm volatilities.

[2] Some propose measures derived from correlations or coefficient estimates (see e.g. Engle and Kelly, 2012; Geraci and Gnabo, 2018; Calabrese and Osmetti, 2019), while others track links between individual companies and broader market/economic movements (see e.g. Acharya et al., 2012; Adrian and Brunnermeier, 2016; Acharya et al., 2017) .

[3] The subsequent literature has widely adopted the Diebold and Yilmaz (2014) approach to address a variety of issues in finance and economics (e.g. Yang and Zhou, 2017; Baruník et al., 2020).

[4] Note that frequency domain techniques are useful tools for denoising (Sun and Meinl, 2012; Haven et al., 2012) and forecasting (Sévi, 2014; Barunik et al., 2016) financial time series.

dynamic relationships that relate to different horizons of interest in multivariate time series models. We use a locally stationary Bayesian time-varying parameter VAR model, which is readily available in high-dimensional settings. We also develop a test for differences in connectedness over different horizons and show how to infer differences over time. We provide Monte Carlo evidence that our measures are able to reliably track connections from different data generating processes (DGPs), including those that are non-Gaussian. Finally, we make computationally efficient packages `DynamicNets.jl` in JULIA and `DynamicNets` in MATLAB that allows one to obtain our measures on data the researcher desires.[5]

Our approach provides a solution to the problems of using rolling windows (see e.g. Demirer et al., 2018) that does not suffer from dimensionality issues or inference problems. The Bayesian nature of our framework incorporates prior shrinkage and provides information about estimation uncertainty from the posterior distribution of the connectedness measures. This is in sharp contrast to conventional studies that only provide point estimates and rely on bootstrapping for confidence intervals. Our measures are also readily available for applications with large data systems. This extends the pairwise approach in Geraci and Gnabo (2018) to study linkages between firms.

The linkages that form over different horizons with heterogeneous persistence are important for a number of reasons. First, economic theory suggests that the marginal utility of agents' preferences depends on cyclical components of consumption (see e.g. Giglio et al., 2015; Bandi and Tamoni, 2017) and also on investment horizons in their risk attitudes (Dew-Becker and Giglio, 2016). Such behaviour can be observed, for example, under myopic loss aversion, where an agent's decision depends on the valuation horizon.

Second, unanticipated shocks or news have the capacity to alter these preferences and can therefore generate transitory and persistent linkages of different strengths. For example, a shock that has an impact at longer horizons may reflect permanent changes in expectations of future price movements. Such a shock may lead to a permanent change in a firm's future dividend payments (Balke and Wohar, 2002). Conversely, a shock that affects shorter horizons may suggest temporary changes in future price movements. For example, suppose that the shock relates only to a change in an upcoming dividend payment. This would likely result in a very short term change, reflecting the transitory nature of the news.

Third, firms have different short-run and long-run objectives, and investors view short-run and long-run risks differently (see e.g. Drechsler and Yaron, 2011; Gerrard et al., 2022). This behaviour motivates the long-run risk asset pricing literature pioneered by Bansal and Yaron (2004) and Bansal et al. (2010). The implication here is that investment horizons may be a source of systematic risk that investors demand compensation for (contributions on this topic include e.g. Brennan and Zhang, 2018; Chaudhuri and Lo, 2019).

Identifying network structures that form due to idiosyncratic shocks is also relevant because they can determine aggregate fluctuations (Acemoglu et al., 2012). These links

[5]The packages are available at `https://github.com/barunik/DynamicNets.jl` and `https://github.com/ellington/DynamicNets`

between individual or firm-level entities create systemic risks for sectors and the economy as a whole (Billio et al., 2012; Acemoglu et al., 2017). Such risks drive changes in uncertainty, which can be key factors in business cycles and financial distress (Bloom et al., 2018). Gabaix (2011) shows that sectoral co-movements are salient features of business cycles. Meanwhile, Atalay (2017) finds that the lion's share of variation in output growth is due to idiosyncratic industry-level shocks.

However, an understanding of the potential longevity of risks arising from these linkages and their importance in driving financial turmoil or business cycles is incomplete. We show how our approach provides measures of transitory and persistent linkages using the daily realised firm-level volatilities of S&P500 constituents. We classify the constituents into their eleven main sectors according to the Global Industry Classification Standard (GICS) and measure network connections from the transitory and persistent components of volatility shocks.

Our empirical results document substantial heterogeneities in transient and persistent measures of connectedness that reveal the nature of systemic risks arising from networks. Specifically, we document: i) spikes in persistent network connectedness when long-lasting financial and economic events occur; and ii) statistically significant differences between transitory and persistent network connectedness across sectors. Our network measures can serve as an online monitoring tool for sectoral uncertainty in markets of interest to macroprudential supervisors and investors alike.

The rest of the paper proceeds as follows. Section 2 derives our measures from locally stationary processes, discusses estimation, and proposes a test procedure for statistical differences in transitory and persistent connectedness. Section 3 provides Monte Carlo evidence that our measures are able to reliably track linkages and correctly identify statistical differences. In Section 4, we examine the links between firm-level volatilities of S&P500 sector constituents and assess the information content of sector connectedness measures beyond leading measures of uncertainty. Finally, section 5 concludes.

2 Measuring Transitory and Persistent Connections

Here we show how one can measure connectedness using the time-varying spectral decompositions. Our measures of connectedness are based on locally stationary processes. This assumes that the process is approximately stationary over a short time interval, which allows us to incorporate time variation into our analysis. This in turn allows us to construct our measures of frequency-dependent, time-varying network connectedness.

Formally, consider a doubly indexed N-variate time series $(\boldsymbol{X}_{t,T})_{1\leq t\leq T, T\in\mathbb{N}}$ with components $\boldsymbol{X}_{t,T} = (\boldsymbol{X}^1_{t,T}, \ldots, \boldsymbol{X}^N_{t,T})^\top$ that describe all variables in an economy. Here t refers to a discrete time index and T is an additional index indicating the sharpness of the local approximation of the time series $(\boldsymbol{X}_{t,T})_{1\leq t\leq T, T\in\mathbb{N}}$ by a stationary one. Coarsely speaking, we can consider $(\boldsymbol{X}_{t,T})_{1\leq t\leq T, T\in\mathbb{N}}$ to be a weakly locally stationary process if, for a large T, given a set S_T of sample indices such that $t/T \approx u$ over $t \in S_T$, the sample $(\boldsymbol{X}_{t,T})_{t\in S_T}$ approximates the sample of a weakly stationary time series depending on the rescaled

location u. Note that u is a continuous time parameter referred to as the rescaled time index, and T is interpreted as the number of available observation, hence $1 \leq t \leq T$ and $u \in [0,1]$, see Dahlhaus (1996) for further details.

We assume that the economy follows a locally stationary TVP-VAR of lag order p as

$$\boldsymbol{X}_{t,T} = \boldsymbol{\Phi}_1(t/T)\boldsymbol{X}_{t-1,T} + \ldots + \boldsymbol{\Phi}_p(t/T)\boldsymbol{X}_{t-p,T} + \boldsymbol{\epsilon}_{t,T}, \tag{1}$$

where $\boldsymbol{\epsilon}_{t,T} = \boldsymbol{\Sigma}^{-1/2}(t/T)\boldsymbol{\eta}_{t,T}$ with $\boldsymbol{\eta}_{t,T} \sim NID(0, \boldsymbol{I}_M)$ and $\boldsymbol{\Phi}(t/T) = (\boldsymbol{\Phi}_1(t/T), \ldots, \boldsymbol{\Phi}_p(t/T))^\top$ are the time varying autoregressive coefficients. In a neighborhood of a fixed time point $u_0 = t_0/T$, we approximate the process $\boldsymbol{X}_{t,T}$ by a stationary process $\widetilde{\boldsymbol{X}}_t(u_0)$ as

$$\widetilde{\boldsymbol{X}}_t(u_0) = \boldsymbol{\Phi}_1(u_0)\widetilde{\boldsymbol{X}}_{t-1}(u_0) + \ldots + \boldsymbol{\Phi}_p(u_0)\widetilde{\boldsymbol{X}}_{t-p}(u_0) + \boldsymbol{\epsilon}_t, \tag{2}$$

with $t \in \mathbb{Z}$ and under suitable regularity conditions $|\boldsymbol{X}_{t,T} - \widetilde{\boldsymbol{X}}_t(u_0)| = O_p(|t/T - u_0| + 1/T)$ which justifies the notation "locally stationary process." Crucially, the process has time varying VMA(∞) representation (Dahlhaus et al., 2009; Roueff and Sanchez-Perez, 2016)

$$\boldsymbol{X}_{t,T} = \sum_{h=-\infty}^{\infty} \boldsymbol{\Psi}_{t,T}(h)\boldsymbol{\epsilon}_{t-h} \tag{3}$$

where $\boldsymbol{\Psi}_{t,T}(h) \approx \boldsymbol{\Psi}(t/T, h)$ is a stochastic process satisfying $\sup_\ell ||\boldsymbol{\Psi}_t - \boldsymbol{\Psi}_\ell||^2 = O_p(h/t)$ for $1 \leq h \leq t$ as $t \to \infty$. Specifically, $\boldsymbol{\Psi}_{t,T}(h) = [\boldsymbol{\Phi}_{t,T}(h)]^{-1}$, which is key to understanding dynamics. Since $\boldsymbol{\Psi}_{t,T}(h)$ contains an infinite number of lags, we approximate the moving average coefficients at $h = 1, \ldots, H$ horizons (see the detailed discussion below). The network characteristics rely on variance decompositions, which are transformations of the impulse response functions, $\boldsymbol{\Psi}_{t,T}(h)$, and permit the measurement of the contribution of shocks to the system.

Since a shock to a variable in the model does not necessarily appear alone, i.e. orthogonally to shocks to other variables, an identification scheme is crucial in calculating variance decompositions. We adapt the generalized identification scheme in Pesaran and Shin (1998) to locally stationary processes. A natural way to disentangle connections that form over transitory and persistent components of shocks is to consider a spectral representation of the approximating model.[6] Hence instead of impulse responses, we propose to use the (local) frequency response of a shock. The building block of our measures consider a time-varying frequency response function $\boldsymbol{\Psi}_{t,T}(e^{-i\omega}) = \sum_h e^{-i\omega h}\boldsymbol{\Psi}_{t,T}(h)$ which we obtain from a Fourier transform of the coefficients with $i = \sqrt{-1}$.

Before introducing our network measures, we define the time varying spectral density of $\boldsymbol{X}_{t,T}$ at frequency ω which is locally the same as the spectral density of $\widetilde{\boldsymbol{X}}_t(u)$ at $u = t/T$ as a Fourier transform of VMA(∞) filtered series as

$$\boldsymbol{S}_{\boldsymbol{X}}(u,\omega) = \sum_{h=-\infty}^{\infty} \mathbb{E}\Big[\widetilde{\boldsymbol{X}}_{t+h}(u)\widetilde{\boldsymbol{X}}_t^\top(u)\Big]e^{-i\omega h} = \Big\{\boldsymbol{\Psi}(u, e^{-i\omega})\Big\}\boldsymbol{\Sigma}(u)\Big\{\boldsymbol{\Psi}(u, e^{+i\omega})\Big\}^\top. \tag{4}$$

[6] Baruník and Křehlík (2018) disentangle long-run and short-run unconditional network connections using standard VAR models.

The time-varying spectral density is a key quantity for understanding frequency dynamics. It describes the distribution of the time varying covariance of $\boldsymbol{X}_{t,T}$ over frequency components ω. The local spectral density captures the influence of the time-varying parameters through the impulse transfer functions $\boldsymbol{\Psi}(u, e^{-i\omega})$, $\boldsymbol{\Psi}(u, e^{+i\omega})$ above. Using the spectral representation for the local covariance that is associated with the local spectral density,

$$\mathbb{E}\left[\widetilde{\boldsymbol{X}}_{t+h}(u)\widetilde{\boldsymbol{X}}_{t}^{\top}(u)\right] = \int_{-\pi}^{\pi} \boldsymbol{S}_{\boldsymbol{X}}(u,\omega)e^{i\omega h}d\omega \tag{5}$$

we can naturally introduce time-varying frequency domain counterparts of variance decompositions. This is important since, as Diebold and Yilmaz (2014) note, we can view the variance decomposition matrix as an adjacency matrix forming asymmetric connections among a system of variables. In our case, this allows us to define dynamic adjacency matrices with different degrees of persistence.

2.1 A Route Towards Transitory and Persistent Connectedness

Variance decompositions are transformations of impulse responses $\boldsymbol{\Psi}_{t,T}(h)$ that allow us to measure the contribution of shocks to the system and thus to characterise the networks that form in response to shocks. While it may seem natural to choose a forecast horizon $h = 1, \ldots, H$ of interest, this choice is costly in terms of information aggregation and hence loss. In contrast to the cumulative information with increasing h, the spectral representation of the impulse responses $\boldsymbol{\Psi}_{t,T}(e^{-i\omega})$ contains much richer and more precise information. Switching to the frequency domain allows one to trace network connections arising from transitory and persistent components of shocks.

To illustrate, consider a simple bivariate system in which the links between two variables are of interest. Specifically, we are interested in how variable b responds to two different shocks to variable a.[7] The first is a *transitory shock* that causes the variable b to rise by one unit in period 1 and fall by one unit in period 2, before returning to zero from period onwards. Here we expect the proportion of error variation to be large at short horizons and small at long horizons due to the purely transitory effect.

The second is a *persistent shock* that results in a unit increase in the variable b and a gradual decrease to zero over the impulse horizon. In this case, the fraction of error variation is expected to be large at longer horizons and smaller at shorter horizons. How these quantities in the time and frequency domain capture the responses to these two shocks is the main motivation for our measures.

Figure 1 shows the impulse response functions of variable b to shocks to variable a in the left panel, the corresponding fraction of forecast error variances in the middle panel, and their corresponding spectral decompositions of forecast error variance fractions in the right panel. Note that one can hardly identify persistence of the shock from the

[7] In this example, we assume that the impact of own shocks on variable b increases by 1 unit and is persistent. This means that the impact of a 1 unit shock to variable b affects the value of b up to 20 horizons after we observe the shock. Note that our analysis in Figure 1 remains the same if the own shocks are transitory.

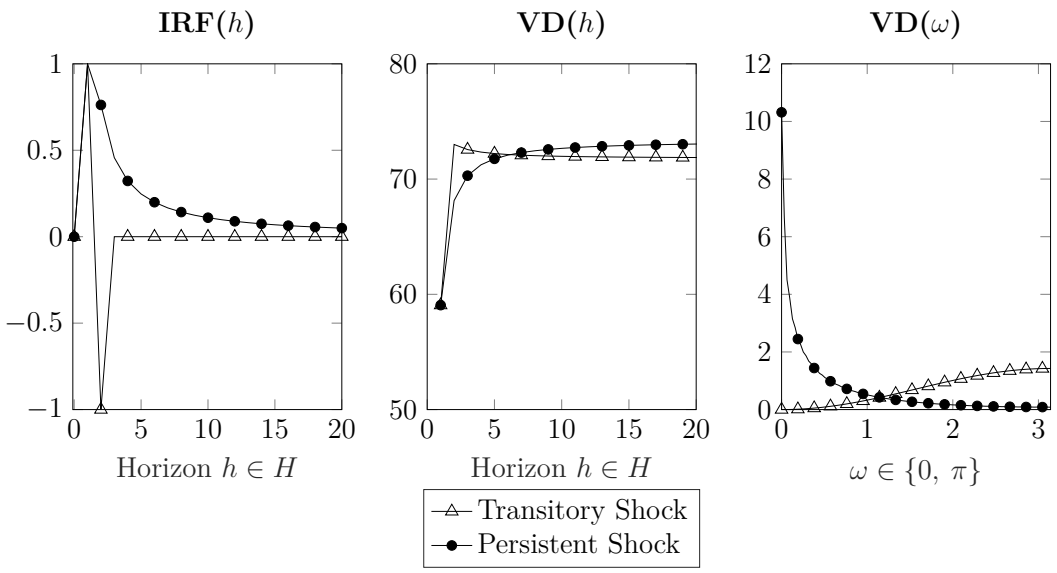

Figure 1: Impulse Response Functions and Variance Decompositions in time and frequency domain

Notes: This figure reports impulse response functions (left), variance decompositions in time domain (middle) and the corresponding spectral representation of the forecast error variance decompositions (right), of a variable b within a bivariate system with respect to a shock in the variable a. We consider two types of shocks, a transitory shock (triangles) and a persistent shock (dots) impacting horizon $h = 1, \ldots, H$ and frequency $\omega \in \{0, \ldots, \omega\}$.

almost indistinguishable forecast error variance shares (variance decompositions) in the time domain depicted by the middle panel of figure 1. A persistent shock results in a slightly larger value of forecast error variation relative to the variation due to a transitory shock. At the same time, if we were to estimate these quantities from the data and take into account the uncertainty of the estimates (the figure plots the theoretical values of a simple example), they would become statistically indistinguishable.

In contrast, the spectral representation of the forecast error variance shares in the right-hand panel of Figure 1 accurately captures the heterogeneous impact of the two shocks across frequencies. The transitory shock in the variable a has a negligible impact at low frequencies (i.e. close to $\omega = 0$), indicating that this shock has no importance for the long-run variation of the variable b, and larger weights at higher frequencies, showing the transitory nature of the link established by this shock. Conversely, the persistent shock affects low frequencies and correctly identifies a persistent link between the variables.

The main implication is that we can construct network measures that take into account the nature of the shocks that form such links. Thus, using spectral decompositions, we are able to identify transitory and persistent network links that are not apparent in the time domain.

2.2 Measuring Connectedness

The following proposition establishes the time-varying spectral representation of the variance decomposition of shocks from variable j to variable k. This is central to the existence of network measures in the time-frequency domain.[8]

Proposition 1 (Dynamic Adjacency Matrix)**.** *Suppose* $\mathbf{X}_{\mathbf{t},\mathbf{T}}$ *is a weakly locally stationary process with* $\sigma_{kk}^{-1}\sum_{h=0}^{\infty}\left|\left[\boldsymbol{\Psi}(u,h)\boldsymbol{\Sigma}(u)\right]_{j,k}\right| < +\infty, \forall j,k$*. Then the* ***time-frequency variance decompositions*** *of the* j*th variable at a rescaled time* $u = t_0/T$ *due to shocks in the* k*th variable on the frequency band* $d = (a,b) : a,b \in (-\pi,\pi), a < b$ *form a* ***dynamic adjacency matrix*** *defined as*

$$\left[\boldsymbol{\theta}(u,d)\right]_{j,k} = \frac{\sigma_{kk}^{-1}\int_a^b \left|\left[\boldsymbol{\Psi}(u,e^{-i\omega})\boldsymbol{\Sigma}(u)\right]_{j,k}\right|^2 d\omega}{\int_{-\pi}^{\pi}\left[\left\{\boldsymbol{\Psi}(u,e^{-i\omega})\right\}\boldsymbol{\Sigma}(u)\left\{\boldsymbol{\Psi}(u,e^{+i\omega})\right\}^{\top}\right]_{j,j} d\omega} \tag{6}$$

where $\boldsymbol{\Psi}(u,e^{-i\omega}) = \sum_h e^{-i\omega h}\boldsymbol{\Psi}(u,h)$ *is local impulse transfer function or frequency response function computed as the Fourier transform of the local impulse response* $\boldsymbol{\Psi}(u,h)$

Proof. See Appendix A. □

It is important to note that $\left[\boldsymbol{\theta}(u,d)\right]_{j,k}$ is a natural disaggregation of traditional variance decompositions to time-varying frequency bands. This is because a portion of the local error variance of the jth variable at a given frequency band due to shocks in the kth variable is scaled by the variance of the jth variable. Note that while the Fourier transform of the impulse response generally takes on complex values, the quantity in proposition (1) is the squared modulus of weighted complex numbers, thus producing a real quantity.

This relationship is an identity which means the integral is a linear operator, summing over disjoint intervals covering the entire range $(-\pi,\pi)$ recovers the time domain counterpart of the local variance decomposition for $h \to \infty$. The following remark formalizes this fact.

Remark 1 (Aggregation of Adjacency Matrix)**.** *Denote by* d_s *an interval on the real line from the set of intervals* D *that form a partition of the interval* $(-\pi,\pi)$*, such that* $\cap_{d_s\in D} d_s = \emptyset$*, and* $\cup_{d_s\in D} d_s = (-\pi,\pi)$*. Due to the linearity of integral and the construction of* d_s*, we have*

$$\left[\boldsymbol{\theta}(u)\right]_{j,k} = \sum_{d_s\in D}\left[\boldsymbol{\theta}(u,d_s)\right]_{j,k}.$$

[8]Note to notation: $[\boldsymbol{A}]_{j,k}$ denotes the jth row and kth column of matrix $\boldsymbol{A}$ denoted in bold. $[\boldsymbol{A}]_{j,\cdot}$ denotes the full jth row; $[\boldsymbol{A}]_{\cdot,j}$ denotes the full jth column. A $\sum A$, where A is a matrix that denotes the sum of all elements of the matrix A.

Remark (1) is important as it establishes the aggregation of network connectedness measures across different frequency bands to its time domain, *total* counterpart. Hence one can easily obtain time varying network measures across any horizon of interest using frequency bands that will always sum up to an aggregate time domain counterpart.

As the rows of the time-frequency network connectedness do not necessarily sum to one, we normalize the element in each by the corresponding row sum

$$\Big[\widetilde{\boldsymbol{\theta}}(u,d)\Big]_{j,k} = \Big[\boldsymbol{\theta}(u,d)\Big]_{j,k} \Big/ \sum_{k=1}^{N} \Big[\boldsymbol{\theta}(u)\Big]_{j,k} \tag{7}$$

Our notion that we can approximate well the process $\boldsymbol{X}_{t,T}$, by a stationary process $\widetilde{\boldsymbol{X}}_t(u)$ in a neighborhood of a fixed time point $u = t/T$, means that all associated local quantities approximate well their time varying counterparts. Following the arguments in Dahlhaus (1996), and using mild assumptions, one can easily see that local variance decompositions at a frequency band $\widetilde{\boldsymbol{\theta}}(u,d)$ approximate well the time-varying variance decompositions of the process $\boldsymbol{X}_{t,T}$.

Note that the local generalized variance decompositions form a dynamic adjacency matrix that defines a time-varying network at a given frequency band. Thus, we can directly use our measures as time-varying network characteristics that contain richer information in comparison to typical network analysis. In our notion, variance decompositions can be viewed as weighted links showing the strengths of connections. In addition, the links are directional, meaning that the j to k link is not necessarily the same as the k to j link, and hence the adjacency matrix is asymmetric. Even more important, the adjacency matrix is time-varying and frequency specific that allows the study of time-varying network characteristics at various frequency bands of the user's choice. The simplest is to measure transitory network connections over the short-run and persistent ones over the long-run.

Now we can define network connectedness measures that characterize a time-varying and frequency specific network. We define local network connectedness measures at a given frequency band as the ratio of the off-diagonal elements to the sum of the entire matrix

$$\mathcal{C}(u,d) = 100 \times \sum_{\substack{j,k=1 \\ j\neq k}}^{N} \Big[\widetilde{\boldsymbol{\theta}}(u,d)\Big]_{j,k} \Big/ \sum_{j,k=1}^{N} \Big[\widetilde{\boldsymbol{\theta}}(u)\Big]_{j,k} \tag{8}$$

This measures the contribution of forecast error variance attributable to all shocks in the system, minus the contribution of own shocks over frequency band d and infers system-wide connectedness over such frequency band. We can also define measures that reveal when an individual variable in the economy is a transmitter or receiver of shocks. Local directional connectedness measures how much of each variables's j variance is due to shocks in other variables $k \neq j$ in the economy over frequency band d is given by

$$\mathcal{C}_{j\leftarrow\bullet}(u,d) = 100 \times \sum_{\substack{k=1 \\ k\neq j}}^{N} \Big[\widetilde{\boldsymbol{\theta}}(u,d)\Big]_{j,k} \Big/ \sum_{j,k=1}^{N} \Big[\widetilde{\boldsymbol{\theta}}(u)\Big]_{j,k}, \tag{9}$$

defining the so-called FROM connectedness. One can precisely interpret this quantity as from-degrees (often called out-degrees in the network literature) associated with the nodes of the weighted directed network represented by the variance decompositions matrix generalized to a time-varying frequency specific quantity. Likewise, the contribution of variable j to variances in other variables is computed as

$$\mathcal{C}_{j\rightarrow\bullet}(u,d) = 100 \times \sum_{\substack{k=1 \\ k\neq j}}^{N} \left[\widetilde{\boldsymbol{\theta}}(u,d)\right]_{k,j} \Big/ \sum_{k,j=1}^{N} \left[\widetilde{\boldsymbol{\theta}}(u)\right]_{k,j} \tag{10}$$

and is the so-called TO connectedness. Again, one can precisely interpret this as to-degrees (often called in-degrees in the network literature) associated with the nodes of the weighted directed network represented by the variance decompositions matrix. These two measures show how other variables contribute to the variation of variable j, and how variable j contributes to the variation of others, respectively, in a time varying fashion at a chosen frequency band. We note here that taking the difference between TO connectedness and FROM connectedness summarizes information regarding directional connections in net-terms. Further, one can track pairwise connections over frequency bands in an analogous manner to the above as differences between the $j-k$th element and $k-j$th elements.

Importantly, the following proposition shows one can always reconstruct time domain network connectedness measures from our frequency-dependent networks.

Proposition 2 (Reconstruction of Dynamic Network Connectedness)**.** *Denote by d_s an interval on the real line from the set of intervals D that form a partition of the interval $(-\pi,\pi)$, such that $\cap_{d_s\in D} d_s = \emptyset$, and $\cup_{d_s\in D} d_s = (-\pi,\pi)$. We then have that*

$$\begin{aligned} \mathcal{C}(u) &= \sum_{d_s\in D} \mathcal{C}(u,d_s) \\ \mathcal{C}_{j\leftarrow\bullet}(u) &= \sum_{d_s\in D} \mathcal{C}_{j\leftarrow\bullet}(u,d_s) \\ \mathcal{C}_{j\rightarrow\bullet}(u) &= \sum_{d_s\in D} \mathcal{C}_{j\rightarrow\bullet}(u,d_s) \end{aligned} \tag{11}$$

where $\mathcal{C}(u)$ are local network connectedness measures aggregated over frequencies.

Proof. See Appendix A. □

In light of the above, all local frequency connectedness measures $\mathcal{C}(u,d)$ for $u = t/T$ approximate well the time-varying frequency connectedness of the process $\boldsymbol{X}_{t,T}$.

2.3 Obtaining Transitory and Persistent Network Measures

In light of the assumptions that underpin our measures, we conjecture that the economy follows a stable time-varying parameter heteroskedastic VAR (TVP-VAR) model as in (1).

We follow Petrova (2019) who establishes a Quasi Bayesian Local-Likelihood approach for inference in the presence of time-varying parameters.

For consistent estimation under the QBLL approach, let $\boldsymbol{X}_{t,T}$ be a time-series we observe with log probability density $l_t\left(\boldsymbol{X}_{t,T}|\boldsymbol{X}_{t-1,T}, \tilde{\boldsymbol{\Phi}}(t/T)\right)$, $\tilde{\boldsymbol{\Phi}}(t/T)$ stacks the time-varying autoregressive coefficient matrices into a finite-dimensional vector that satisfies one of the following conditions.

i) $\tilde{\boldsymbol{\Phi}}_t = \tilde{\boldsymbol{\Phi}}(t/T)$ is a deterministic process where $\boldsymbol{\Phi}(.)$ is a piecewise differentiable function.

ii) $\tilde{\boldsymbol{\Phi}}(t/T)$ is a stochastic process satisfying: $\sup_{j:|j-t|\leq h}||\tilde{\boldsymbol{\Phi}}_t - \tilde{\boldsymbol{\Phi}}_j||^2 = O_p(h/t)$ for $1 \leq h \leq t$, $t \to \infty$.

Both of the above indicate that the parameter sequence drafts gradually over time. The first condition is standard of Dahlhaus (2000) for locally stationary processes which requires the parameter process is a piecewise smooth deterministic function; thus allowing for breaks in parameters. The second condition is a generalization of the first to include stochastic parameter processes exhibiting degrees of persistence necessary for consistent estimation of stochastic driven time-variation. Such condition includes bounded random walk processes and some fractionally integrated processes. The parameters may feature any combination of deterministic trends and/or breaks satisfying conditions i) and ii) above. Our data generating processes (DGPs) in Section 3 are examples of such DGPs.[9]

To obtain the time-varying coefficient estimates at a fixed time point $u = t_0/T$, $\widehat{\boldsymbol{\Phi}}_1(u), ..., \widehat{\boldsymbol{\Phi}}_p(u)$, and the time-varying covariance matrices, $\widehat{\boldsymbol{\Sigma}}(u)$, we follow the QBLL approach of Petrova (2019). Specifically, this approach uses a kernel weighting function that provides larger weights to observations that surround the period whose coefficient and covariance matrices are of interest. Using conjugate priors, the (quasi) posterior distribution of the parameters of the model are analytical. This alleviates the need to use a Markov Chain Monte Carlo (MCMC) simulation algorithm and permits the use of parallel computing. Note also that in using (quasi) Bayesian estimation methods, we obtain a distribution of parameters that we use to construct network measures that provide confidence bands. Details of the model and estimation algorithm are in Appendix B.[10] We provide a computationally efficient package `DynamicNets.jl` in JULIA and `DynamicNets` in MATLAB that allows one to obtain our measures on data the researcher desires.[11]

To estimate the elements of dynamic adjacency matrix, we first need to truncate the infinite VMA(∞) representation of the approximating model with a choice of finite horizon H. Here we note that in the frequency domain quantities, H serves only as an

[9]Figure 1 of Petrova (2019) provides figures of examples and provides further discussion around conditions the parameter sequences must satisfy for consistent estimation.

[10]Unlike traditional TVP VARs time-variation evolves in a non-parametric manner thus making no assumption on the laws of motion within the model. Typically, the model of Primiceri (2005), and many extensions, assume parameters evolve as random walks or autoregressive processes.

[11]The packages are available at https://github.com/barunik/DynamicNets.jl and https://github.com/ellington/DynamicNets

approximation factor, and it has no interpretation as in the time domain. Hence in the applications we advise setting the H sufficiently high to obtain a better approximation, particularly when lower frequencies are of interest. We obtain horizon specific measures using Fourier transforms and set our truncation horizon $H = 100$. Note, we run all results in this paper for $H \in \{50, 100, 200\}$, they are qualitatively similar and available upon request.

Next, estimating dynamic network measures requires the user to choose a kernel and its bandwidth. Typically the larger the bandwidth, the smoother time-evolution of our frequency specific network measures. Therefore, prior to tracking dynamic network connections, it is important the user considers the time-series properties of their data. For example if common peaks (troughs) in the time-series occur frequently and are transient, then a shorter bandwidth may be necessary. Conversely, if tracking network connections among data that evolves gradually over time, like interest rates, a larger bandwidth may be more appropriate. In the context of our study, we use a Normal kernel and explore the implication of bandwidth choice for a variety of data generating processes (DGPs) in Section 3.

It is noteworthy to mention that the choice of a two-sided kernel can come at a cost; especially if one wishes to use network measures for forecasting purposes. In these cases one may wish to: i) estimate the dynamic network recursively throughout time such that the Normal kernel truncates to use only past values at the time T estimate; or ii) choose a one-sided kernel such as those in Hahn et al. (2001); Barigozzi et al. (2020). In practice, we encourage researchers to experiment with a variety of bandwidths to ensure results are not driven by its selection. We also encourage authors to use reasonable bandwidths given the data for their application. For example, if one was using these measures for forecasting daily stock return volatilities, a researcher might consider combination forecasts using Bayesian Model averaging to trade on. Alternatively one might look to use variance minimizing kernels in an attempt to reduce uncertainty around the forecast.[12]

We estimate the j, k element of our dynamic adjacency matrix at time $u = t_0/T$ and horizon $d = (a, b) : a, b \in (-\pi, \pi)$ and $a < b$ such that it corresponds to the transitory (high frequency band) and persistent (low frequency band) element of the adjacency matrix respectively as:

$$\left[\widehat{\boldsymbol{\theta}}(u,d)\right]_{j,k} = \frac{\widehat{\sigma}_{kk}^{-1}\sum_{\omega\in d}\left(\left[\widehat{\boldsymbol{\Psi}}(u,\omega)\widehat{\boldsymbol{\Sigma}}(u)\right]_{j,k}\right)^2}{\sum_{\omega\in(-\pi,\pi)}\left[\widehat{\boldsymbol{\Psi}}(u,\omega)\widehat{\boldsymbol{\Sigma}}(u)\widehat{\boldsymbol{\Psi}}^{\top}(u,\omega)\right]_{j,j}}, \tag{12}$$

where $\widehat{\boldsymbol{\Psi}}(u,\omega) = \sum_{h=0}^{H-1}\sum_h \widehat{\boldsymbol{\Psi}}(u,h)e^{-i\omega h}$ is an estimate of the impulse transfer function from Fourier frequencies $\omega \in \{aH/2\pi, \ldots, bH/2\pi\}$ of impulse response functions that

[12]In the context of our empirical application below where we look at daily realized volatilities of stock returns, we use a bandwidth equal to 8. We also estimate the models using bandwidths of 12, 18, and $\sqrt{T} = \sqrt{3278} \approx 57$. Increasing the bandwidth smooths our network connectedness measures because it assigns larger weights to more distant observations.

cover a specific horizon.[13] From this, estimates of Equations (8)–(10) directly follow. For example if the application uses daily data, one may define transitory (short-term) as horizons corresponding to 1–5 days and persistent (long-term) as horizons corresponding to horizons greater than 5 days. This would require defining the band as $(a, b) = (2\pi/5, 2\pi)$ for the transitory and $(a, b) = (0, 2\pi/5)$ for the persistent networks.

2.4 Testing for Statistical Differences in Connectedness

We now consider how one can determine, from a statistical perspective, differences between connectedness. We discuss in detail here how one can test for differences in connectedness one computes over different frequency bands.

In a Bayesian setting there are three alternatives for hypothesis testing. The first is the Bayes factor, the second uses posterior credible intervals, and the third follows statistical decision theory. We follow the latter and utilize the work of Li et al. (2014), Li et al. (2015) and Liu et al. (2022).[14] These studies focus on developing test statistics of a point null hypothesis using the posterior distribution of parameters from a Bayesian model. The approach requires only the posterior distribution of parameters and has various advantages. First, they overcome the problem of the Jeffreys-Lindely paradox. Second, are not sensitive to the prior and are pivotal quantities. Third, they are easy to compute. Crucially, these statistics directly come from quadratic loss functions, as with classical test statistics, and therefore possess the same distributions as their frequentist counterparts.

Li et al. (2015) develop a Bayesian Lagrange-Multiplier (LM) type test that is asymptotically equivalent to a classical LM test. Using similar assumptions, Liu et al. (2022) develop a Bayesian Wald type test that is asymptotically equivalent to a classical Wald test and requires only the posterior mean and posterior variance of parameters under the null hypothesis.[15] Noting that the VMA(∞) representation of the VAR are nothing more than a transformation of the VAR parameters, we follow the assumptions in Liu et al. (2022) and therefore establish a Bayesian Wald type test for differences between network connectedness across frequency bands. We emphasize our network connectedness measures are manipulations of the VAR parameters themselves and possess a posterior distribution.[16]

[13] Note that $i = \sqrt{-1}$.

[14] Bayes factors involve comparing the marginal likelihoods of two competing models and extensively appear in the literature (see e.g. Koop et al., 2010; Chan, 2020). This is not appropriate in our setting because the network connectedness measures come from a manipulation of a sequence of posterior parameters from the same model; we have no alternative model to specify the marginal likelihood. Using posterior credible intervals is possible and something we consider in the spirit of Cogley et al. (2010). In particular we use the joint posterior distribution of network connectedness measures to compute the probability that connectedness across one frequency band is larger than an analogous measure across another frequency band. These results are in Appendix C.

[15] They also show asymptotic equivalence between their Wald-type test the LM type test in Li et al. (2015).

[16] Lütkepohl (1990) provide the asymptotic distribution for impulse response functions from conventional VAR models one estimates using OLS. Petrova (2019)

We test the null hypothesis, $\mathrm{H}_0 : \mathcal{C}(u, d_a) = \mathcal{C}(u, d_b)$ that network connectedness across frequency band d_a and d_b are equivalent against the alternative, $\mathrm{H}_1 : \mathcal{C}(u, d_a) \neq \mathcal{C}(u, d_b)$. This is equivalent to testing $\mathrm{H}_0 : \mathcal{C}(u, d_a) - \mathcal{C}(u, d_b) = 0$ against $\mathrm{H}_1 : \mathcal{C}(u, d_a) - \mathcal{C}(u, d_b) \neq 0$. The following proposition establishes the test statistic and its asymptotic distribution under the null. We outline the regularity conditions in Appendix A along with a brief discussion on the importance of such conditions.

Proposition 3 (Testing for Heterogeneities in Network Connectedness)**.** *Let* $\bar{\mathcal{D}}(u)$ *and* $\mathbf{V}_{\mathcal{D}}\left(\bar{\mathcal{D}}(u)\right)$ *denote the time* u *posterior mean and variance of the difference between network connectedness across frequency band* d_a *and* d_b*. Then the test statistic under* H_0

$$\begin{aligned} \mathbf{W}\left(\boldsymbol{X}, \mathcal{D}_0(u)\right) &= q_{\mathcal{D}} + \left(\bar{\mathcal{D}}(u) - \mathcal{D}_0(u)\right)^{\top} \left[\mathbf{V}_{\mathcal{D}\mathcal{D}}\left(\bar{\mathcal{D}}(u)\right)\right]^{-1} \left(\bar{\mathcal{D}}(u) - \mathcal{D}_0(u)\right) \quad (13) \\ &= q_{\mathcal{D}} + \mathbf{Wald} \quad (14) \end{aligned}$$

where $\mathbf{Wald} = \left(\bar{\mathcal{D}}(u) - \mathcal{D}_0(u)\right)^{\top} \left[\mathbf{V}_{\mathcal{D}\mathcal{D}}\left(\bar{\mathcal{D}}(u)\right)\right]^{-1} \left(\bar{\mathcal{D}}(u) - \mathcal{D}_0(u)\right)$*,* $q_{\mathcal{D}}$ *is the number of restrictions, and* $\mathcal{D}_0(u) = \mathcal{C}(u, d_a) - \mathcal{C}(u, d_b) = 0$*.*

$$\mathbf{W}\left(\boldsymbol{X}, \mathcal{D}_0(u)\right) - q_{\mathcal{D}} = \mathbf{Wald} + O_p(1) \to^d \chi^2\left(q_{\mathcal{D}}\right)$$

Proof. See Appendix A. □

It is important to note that it is straightforward to generalize Equation (13) to include multiple restrictions. This may be applicable if one requires testing equivalence among more than two frequency bands. We also note that one may also wish to utilize the above to test for differences between directional network connections over frequency bands. In the context of the above one would take the difference between net-directional connections, or pairwise directional connections, over frequency bands and compute the test in an analogous manner to below.

For estimation purposes, the test statistic only requires the posterior mean and the posterior variance of $\mathcal{D}(u)$, $\bar{\mathcal{D}}(u)$ and $\mathbf{V}\left(\bar{\mathcal{D}}\right)$. Let $\left\{\mathcal{D}^{[r]}\right\}_{r=1}^{R}$ denote the posterior draws such that $\mathcal{D}^{[r]} = \widehat{\mathcal{C}}^{[r]}(u, d_a) - \widehat{\mathcal{C}}^{[r]}(u, d_b)$ is the rth posterior draw of the difference between estimates of network connectedness over frequency band d_a and d_b. Then, the estimate of our test statistic for heterogeneities between network connectedness is given by

$$\widehat{\mathbf{W}}\left(\boldsymbol{X}, \mathcal{D}(u) = 0\right) = \frac{\frac{1}{R}\sum_{r=1}^{R}\left(\mathcal{D}^{[r]}(u)\right)^2}{\frac{1}{R}\sum_{r=1}^{R}\left(\mathcal{D}^{[r]}(u) - \bar{\bar{\mathcal{D}}}(u)\right)^2}, \text{ with } \bar{\bar{\mathcal{D}}}(u) = \frac{1}{R}\sum_{r=1}^{R}\mathcal{D}^{[r]}(u) \quad (15)$$

Under the null, we have $\mathbf{W}\left(\boldsymbol{X}, \mathcal{D}_0(u)\right) - q_{\mathcal{D}} \to^d \chi^2\left(q_{\mathcal{D}}\right)$ with $q_{\mathcal{D}} = 1$ in this particular case.[17] Thus, we only need to compare $\widehat{\mathbf{W}}\left(\boldsymbol{X}, \mathcal{D}(u) = 0\right)$ to the critical values of the

[17] In the classical setting one uses Wald tests to check a variety of restrictions, such as a parameter of interest being equal to zero, or equivalence between two parameters of interest under the null hypothesis. When one tests the latter, the resulting Wald test is $\backsim \chi^2(1)$.

$\chi^2(1)$ distribution. Rejecting the null implies that the time u network connectedness over frequency band d_a is statistically different to the corresponding time u network connectedness over frequency band d_b.

It may also be pertinent to test for differences in connectedness over time. Here we fix the frequency band, and now consider differences over time. In this case we test the null hypothesis, $\mathrm{H}_0 : \mathcal{C}(u_1, d) = \mathcal{C}(u_2, d)$ that network connectedness at time u_1 and u_2 across frequency band d are equivalent against the alternative, $\mathrm{H}_1 : \mathcal{C}(u_1, d) \neq \mathcal{C}(u_2, d)$. This is equivalent to testing $\mathrm{H}_0 : \mathcal{C}(u_1, d) - \mathcal{C}(u_2, d) = 0$ against $\mathrm{H}_1 : \mathcal{C}(u_1, d) - \mathcal{C}(u_2, d) \neq 0$. Now letting $\bar{\mathcal{D}}(s)$ and $\mathbf{V}_{\mathcal{DD}}\left(\bar{\mathcal{D}}(s)\right)$ denote the posterior mean and variance of the difference between network connectedness at times u_1 and u_2 and replacing $\bar{\mathcal{D}}(u)$ and $\mathbf{V}_{\mathcal{DD}}\left(\bar{\mathcal{D}}(u)\right)$ in Proposition 3 with these quantities delivers a Wald-type test statistic following a $\chi^2(1)$ distribution. Again, all we need to do is compare the test statistic with critical values of the $\chi^2(1)$ distribution.

3 Monte Carlo Study

In this section, we conduct a Monte Carlo exercise to understand the finite sample properties of our connectedness measures. In order to motivate the need to focus on connections forming conditional on the persistence of shocks, we generate data with different levels of persistence throughout time and also changes in the covariance structure. This will induce differences in network connectedness measures we compute over different frequency bands. Here we concentrate on low and high frequency bands and consider four different data generating processes (DGP) to highlight their uses.

For simplicity, we focus on bivariate VAR(2) models with time-varying parameters and time-varying covariance matrices:

$$\begin{aligned} \boldsymbol{X}_{t,T} &= \boldsymbol{\Phi}_0(u) + \boldsymbol{\Phi}_1(u)\boldsymbol{X}_{t-1,T} + \boldsymbol{\Phi}_2(u)\boldsymbol{X}_{t-2,T} + \boldsymbol{\epsilon}_{t,T}, \\ \boldsymbol{\epsilon}_{t,T} &= \boldsymbol{\Sigma}^{-1/2}(u)\boldsymbol{\eta}_{t,T},\ \boldsymbol{\eta}_{t,T} \backsim (0, \mathbf{I}_2) \end{aligned}$$

where $\boldsymbol{\Phi}_0(u)$ contains the time-varying intercepts and $\boldsymbol{\Phi}_1(u)$ and $\boldsymbol{\Phi}_2(u)$ contain the time-varying autoregressive parameters. The time-varying covariance matrix $\boldsymbol{\Sigma}(u) = \mathbf{A}^{-1}(u)\mathbf{H}(u)\left(\mathbf{A}^{-1}(u)\right)^\top$ with $\mathbf{A}^{-1}(u)$ being a lower triangular matrix with a unit diagonal and $\mathbf{H}(u)$ is a 2×2 diagonal matrix.

DGPI: Our first DGP has residuals such that, $\boldsymbol{\eta}_{t,T} \backsim \mathrm{NID}(0, \mathbf{I}_2)$. The time-varying intercepts follow the process:

$$\left[\boldsymbol{\Phi}_0(u)\right]_j = 0.0025\sin(0.004\pi t) + 0.15\sum_{i=1}^{t}\frac{\nu_i}{\sqrt{t}},\ \nu_i \backsim \mathrm{NID}\left(0, 0.001^2\right),\ j = 1, 2$$

For the time-varying autoregressive parameters, we have

$$\left[\boldsymbol{\Phi}_g(u)\right]_{j,k} = \begin{cases} 0.05\sin(0.002\pi t) + 0.75\sum_{i=1}^{t}\frac{\kappa_i}{\sqrt{t}}, & t \in \{1, ..., 500\}\ \forall g, j, k = 1, 2 \\ 0.45\sin(0.002\pi t) + 0.75\sum_{i=1}^{t}\frac{\kappa_i}{\sqrt{t}}, & t \in \{501, ..., 1000\}\ j = k = 1,\ j = k = 2 \\ 0.05\sin(0.002\pi t) + 0.75\sum_{i=1}^{t}\frac{\kappa_i}{\sqrt{t}}, & t \in \{501, ..., 1000\}\ j = 1, k = 2\ \&\ j = 2, k = 1 \end{cases}$$

with $\kappa_i \backsim \text{NID}\left(0, 0.0001^2\right)$. The (2,1) element of $\mathbf{A}(u)$ have the following dynamics:

$$[\mathbf{A}(u)]_{2,1} = \begin{cases} 0.03\sin\left(0.002\pi t\right) + 0.7\sum_{i=1}^{t}\frac{\upsilon_i}{\sqrt{t}}, & t \in \{1, ..., 500\} \\ 1.5\sin\left(0.002\pi t\right) + 0.7\sum_{i=1}^{t}\frac{\upsilon_i}{\sqrt{t}}, & t \in \{501, ..., 1000\} \end{cases}$$

with $\upsilon_i \backsim \text{NID}\left(0, 0.3^2\right)$. The diagonal elements of $\mathbf{H}(u)$ follow

$$\log\left[\mathbf{H}(u)\right]_{j,j} = \mu_j + \lambda_j\left(\log\left[\mathbf{H}(u-1)\right]_{j,j} - \mu_j\right) + \xi_{j,t}$$

where $\xi_{j,t} \backsim \text{NID}\left(\mu_j, 0.1^2/(1-\lambda_j)\right)$, $\mu_j = 0.01$, $\lambda_j = 0.95$.

This DGP has little to no dependence for the first 500 observations which means connectedness at both high and low frequency bands will be low and close to zero. The latter half of the sample sees the AR coefficients in each equation become persistent as the sin wave becomes negative. Note also that the contemporaneous relationship intensifies. This induces connections at high frequency bands while connections at low frequency bands should be low and close to zero. Note we also allow for non-Gaussian residuals in this DGP such that we draw $\boldsymbol{\eta}_{t,T}$ from a multivariate student-t distribution with 5 degrees of freedom. These results are in Appendix C.

DGPII: For our second DGP, the time-varying intercepts follow the process:

$$\left[\mathbf{\Phi}_0(u)\right]_j = 0.25\sin\left(0.004\pi t\right) + 0.15\sum_{i=1}^{t}\frac{\nu_i}{\sqrt{t}},\ \nu_i \backsim \text{NID}\left(0, 0.1^2\right),\ j = 1, 2$$

and the time-varying autoregressive parameters follow:

$$\left[\mathbf{\Phi}_g(u)\right]_{j,k} = 0.25\sin\left(0.004\pi t\right) + 0.75\sum_{i=1}^{t}\frac{\kappa_i}{\sqrt{t}}, \forall g, j, k = 1, 2$$

with $\kappa_i \backsim \text{NID}\left(0, 0.3^2\right)$. The (2,1) element of $\mathbf{A}(u)$ and the diagonal elements of $\mathbf{H}(u)$ follow the processes:

$$\begin{aligned} [\mathbf{A}(u)]_{2,1} &= 0.3\sin\left(0.008\pi t\right) + 0.7\sum_{i=1}^{t}\frac{\upsilon_i}{\sqrt{t}} \\ \log\left[\mathbf{H}(u)\right]_{j,j} &= \mu_j + \lambda_j\left(\log\left[\mathbf{H}(u-1)\right]_{j,j} - \mu_j\right) + \xi_{j,t}, \end{aligned}$$

where $\upsilon_i \backsim \text{NID}\left(0, 0.3^2\right)$ and $\xi_{j,t} \backsim \text{NID}\left(\mu_j, 0.1^2/(1-\lambda_j)\right)$, $\mu_j = 0.01$, $\lambda_j = 0.95$.

This DGP induces two distinct periods of persistence during observations 100-200 and 600-700 which amplifies connectedness at the low frequency band.

DGPIII: Our third DGP is the same as DGPII but relaxes the assumption that $\boldsymbol{\eta}_{t,T}$ are Gaussian. Instead we assume that the residuals follow a multivariate student-t distribution with 5 degrees of freedom.

DGPIV: Our fourth DGP is the same as DGPII, but increases the periodicity of the sin functions in the time-varying autoregressive matrices from $\sin\left(0.004\pi t\right)$ to

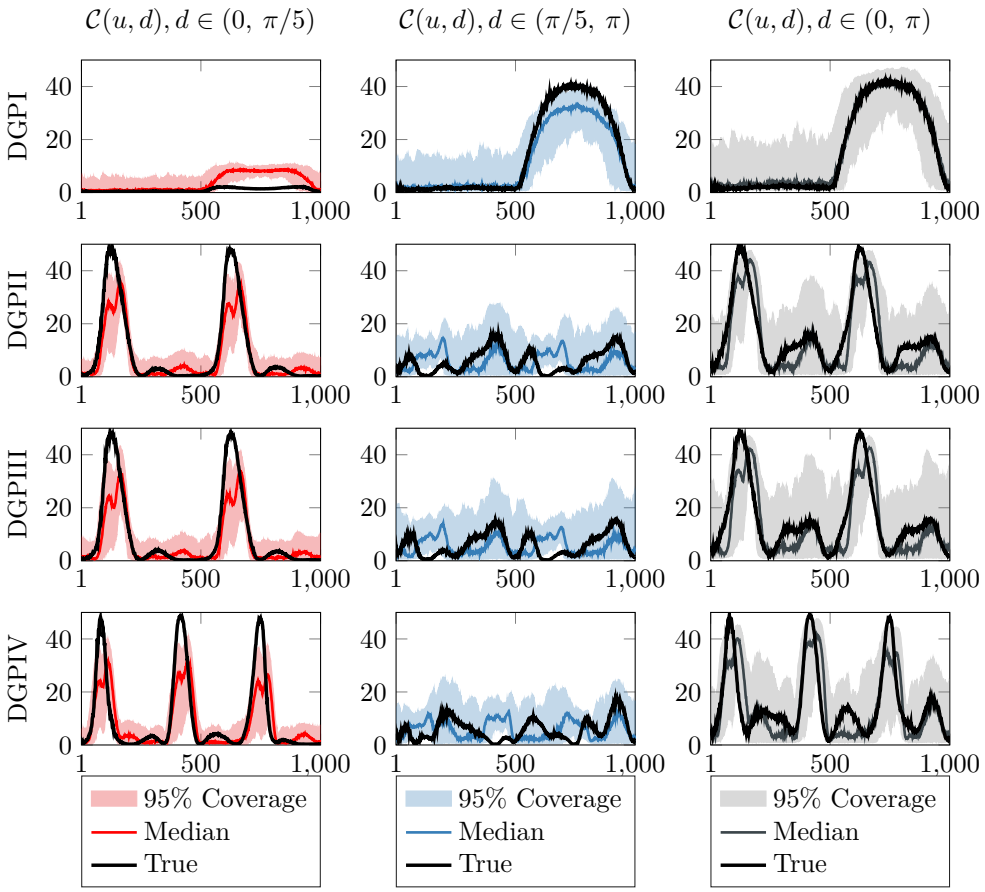

Figure 2: Dynamic network connectedness measures: True and fitted values
Notes: This figure plots the true network connectedness measures for three data generating processes following bi-variate TVP VAR(2) models along with the median and 95% quantiles of estimated network connectedness measures with bandwidth W=8. The left columns report network connectedness on the low-frequency band, $d \in (0,\ \pi/5)$, the middle columns show network connectedness on the high-frequency band, $d \in (\pi/5,\ \pi)$, and the right columns show the aggregate network connectedness such that $d \in (0,\ \pi)$. DGPI (top row) is a TVP VAR(2) model with Gaussian errors, we introduce a break in the time-varying autoregressive matrices and contemporaneous relations from observation 500 that induces large connections across the high frequency band. DGPII (second row) is a TVP VAR(2) where time-varying intercepts and autoregressive matrices following sin wave functions with a stochastic error, time-varying covariance matrix where the off-diagonals follow sin wave functions with a stochastic error, and the diagonal elements follow a stationary AR(1) processes. DGPIII (third row) is a TVP VAR(2) model with student-t errors, time-varying intercepts and autoregressive matrices following sin wave functions with a stochastic error, time-varying covariance matrix where the off-diagonals follow sin wave functions with a stochastic error, and the diagonal elements follow a stationary AR(1) processes. DGPIV (bottom row) is the same as DGPII, but with an increase in the periodicity of the respective sin wave functions the time-varying intercepts and autoregressive matrices follow.

$\sin(0.006\pi t)$. This generates three distinct periods of persistence during observations 50-150, 300-450, and 700-850.

For each of the four DGPs, we generate 100 simulations of length $T = 1000$ and compute the network connectedness measures. We use the median over these simulations as the true network connectedness. Then, for each of the 100 simulations of DGPI–DGPIV, we fit the TVP-VAR model we outline in Section 2.3. In fitting this model we take 1000 draws from the posterior distribution, calculate our network connectedness measures and then save the posterior median. For this exercise, we compute network connectedness on two frequency bands that cover the spectrum. The low-frequency band, which empirically pertains to persistent network connections, is $d \in (0,\ \pi/5)$, and the high-frequency band, pertaining to transitory network connections, is $d \in (\pi/5,\ \pi)$. For completeness, we compute the aggregate connectedness measures that considers the entire spectrum such that $d \in (0,\ \pi)$; this corresponds to a dynamic version of the Diebold and Yilmaz (2014) connectedness measure.

A final noteworthy point is the choice of bandwidth, W, for the kernel weights that induce time-variation into the TVP-VAR model. In this exercise, we consider $W = \{8,\ 12,\ 18\}$. However, for ease of exposition, we only report plots of the network connectedness measure estimates using W=8 in the main text, results of fitted values from $W = \{12,\ 18\}$ are in Appendix C. The larger the bandwidth, the smoother the network measures become. This is because larger bandwidths assign weights to a higher number of observations around the one of interest. In general, we find that larger W results in poorer fit. This highlights the importance of selecting an appropriate bandwidth for the kernel weights relative to the data application.

As we discuss in 2.3, from a practical perspective, we encourage researchers to explore the robustness of their results to different bandwidths. For example, consider a low frequency forecaster looking to predict returns one-month ahead today, using our network connectedness measures, would likely place zero weight on data from the burst of the dot-com bubble and 2008 recession. Likewise, a high frequency investor would likely place little to no weight on data from one-year prior. As we show in Appendix C, lengthening the bandwidth in our Monte Carlo experiment causes the surges in estimates of connectedness to be more gradual.

If the data is low frequency data such as monthly yields, then one could argue to use a wider bandwidth as changes in these data are far smoother than returns or return volatility. If one is looking to describe the nature of connections that form on persistent and transitory components of shocks then we encourage researchers to use multiple bandwidths to check how these dynamics are influenced by such changes. We do not suspect these changes would drastically change the conclusions or results in most applications.

In Appendix C, we conduct further robustness checks for our simulation analysis. In Appendix C.3 we provide analysis on the performance of rolling VAR models for our DGPs. These results show that such connectedness estimates less accurate relative to our approach. In particular, estimates are highly sensitive to the window size and fail to accurately capture the peaks and troughs in connectedness; the latter is prominent as the complexity of the DGP increases (e.g. DGPII-DGPIV). Meanwhile in Appendix

C.4 we conduct a Monte Carlo study on larger scale VAR models under DGPI containing N=10 and N=25 variables respectively. These results show that our approach is robust to increasing the number of variables and tracks connections well within larger-scale models.

Figure 2 reports the true network connectedness measures and the median and 95% quantiles of corresponding estimates from the TVP VAR model using a kernel bandwidth of W=8. We report network connectedness over the low-frequency-band, the high-frequency-band, and aggregate, in the left, middle, and right columns respectively. The top row corresponds to DGPI, and the second, third and fourth rows results from DGPII, DGPIII, and DGPIV respectively. As we can see, the distribution of estimates for each DGP track the true values remarkably well. In almost all cases, the true value lies within the 95% quantiles of the distribution from model estimates. This plot shows that our method provides an accurate representation of horizon specific network connectedness, even when the underlying process has complex dynamics and the true error distribution is non-Gaussian.

Figure 3 further reports the estimates, and numerical standard error bounds of our test for heterogeneities between high frequency band and low frequency band network connectedness measures (Equation 15) for each of the four DGPs we use in the Monte Carlo study. Specifically, for each observation we test the null hypothesis $\mathcal{D}(u) = 0$ where $\mathcal{D}(u) = \widehat{\mathcal{C}}(u, d) - \widehat{\mathcal{C}}(u, c)$ with $d \in (0,\ \pi/5)$ corresponding to the low frequency band and $c \in (\pi/5,\ \pi)$ corresponding to the high frequency band. In each plot, we also report the 5% critical value from the $\chi^2(1)$ distribution of 3.84. Test statistics exceeding this value reject the null in favor of heterogeneities between network connectedness measures across frequency bands.[18] As we can see, our test statistic identifies significant differences between low and high frequency band network connectedness measures that correspond with the peaks we observe in Figure 2 for each DGP. We can see with the non-Gaussian DGP, DGPIII the estimates of the test statistics are smaller relative to the analogous Gaussian DGP, DGPII. However, there are still clear rejections.

We now test for differences in connectedness measures across the same frequency band over time. To do so, we test the first time period against all remaining 999 observations from our DGPs, $u_1 = 1$, $u_2 = \{2, 3, \ldots, 1,000\}$. We do this for both the low-frequency band, $d \in (0,\ \pi/5)$, and the high-frequency band, $d \in (\pi/5,\ \pi)$. Figure 4 reports the estimates of the test statistics, their numerical standard error and the corresponding 95% critical value. First, considering DGPI, it is clear that connectedness across the high frequency band exhibits significant differences between the first observation when we increase connections at observation 500. It is noteworthy to mention that we also see this for the low frequency band. This is expected as we can see that the estimates from our simulations exhibit slight bias here. However, for DGPII–DGPIV we have rejections relative to the first observation at the corresponding periods where we create connectedness across the low frequency band, and no rejections across the high frequency band.

Overall, this shows that our testing procedure indicates rejections of equality in connectedness forming over different frequency bands, and over time for connectedness across

[18] We obtain numerical standard errors in a similar manner to Li et al. (2015).

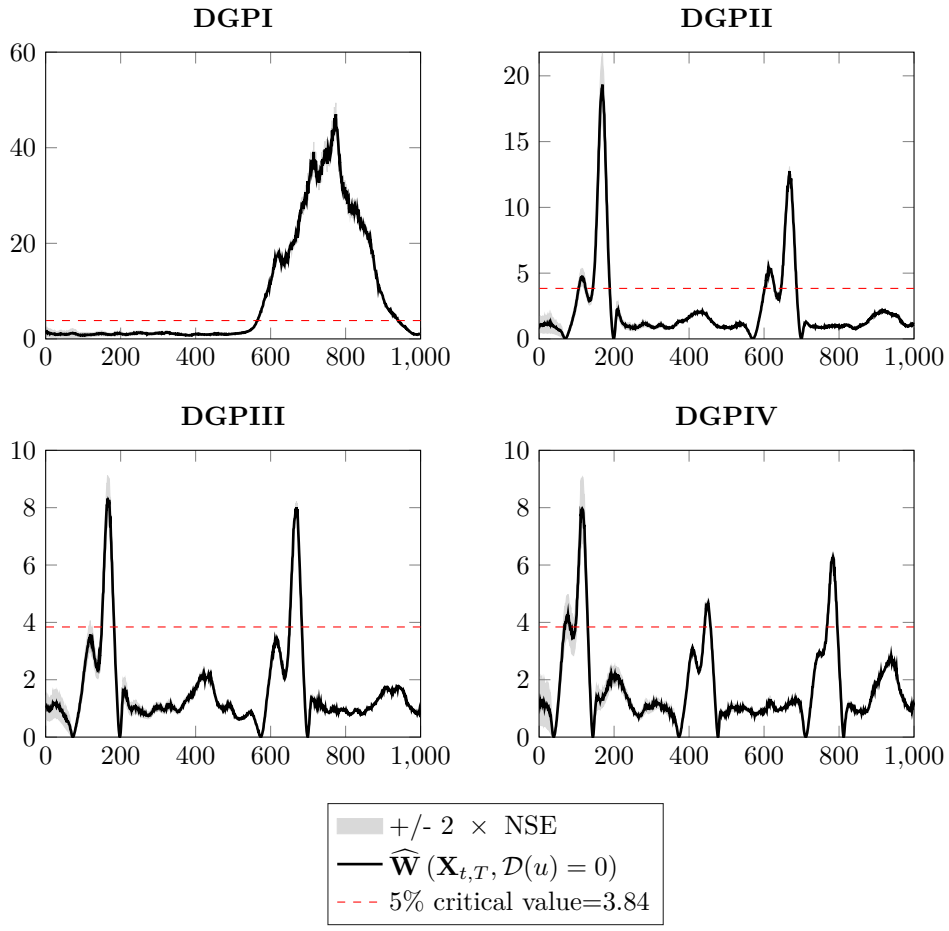

Figure 3: Tests for differences between network connectedness across different frequency bands

Notes: This figure plots the test statistics and $+/-$ 2 $\times$ the numerical standard errors for heterogeneities between network connectedness measures across different frequency bands from four DGPs. The test statistic is $\widehat{\mathbf{W}}(\boldsymbol{X}_{t,T}, \mathcal{D}(u) = 0) - 1$ where $\mathcal{D}(u) = \widehat{\mathcal{C}}(u, d) - \widehat{\mathcal{C}}(u, c)$ with $d \in (0,\ \pi/5)$ corresponding to the low frequency band and $c \in (\pi/5,\ \pi)$ corresponding to the high frequency band. The dashed black line is the 5% critical value from the $\chi^2(1)$ distribution =3.84. Values greater than 3.84 reject the null hypothesis of equivalent network connections across frequency band d and c in favor of differences. DGPI (top left panel) is a TVP VAR(2) model with Gaussian errors, we introduce a break in the time-varying autoregressive matrices and contemporaneous relations from observation 500 that induces large connections across the high frequency band. DGPII (top right panel) is a TVP VAR(2) where time-varying intercepts and autoregressive matrices following sin wave functions with a stochastic error, time-varying covariance matrix where the off-diagonals follow sin wave functions with a stochastic error, and the diagonal elements follow a stationary AR(1) processes. DGPIII (bottom left panel) is a TVP VAR(2) model with student-t errors, time-varying intercepts and autoregressive matrices following sin wave functions with a stochastic error, time-varying covariance matrix where the off-diagonals follow sin wave functions with a stochastic error, and the diagonal elements follow a stationary AR(1) processes. DGPIV (bottom right panel) is the same as DGPII, but with an increase in the periodicity of the respective sin wave functions the time-varying intercepts and autoregressive matrices follow.

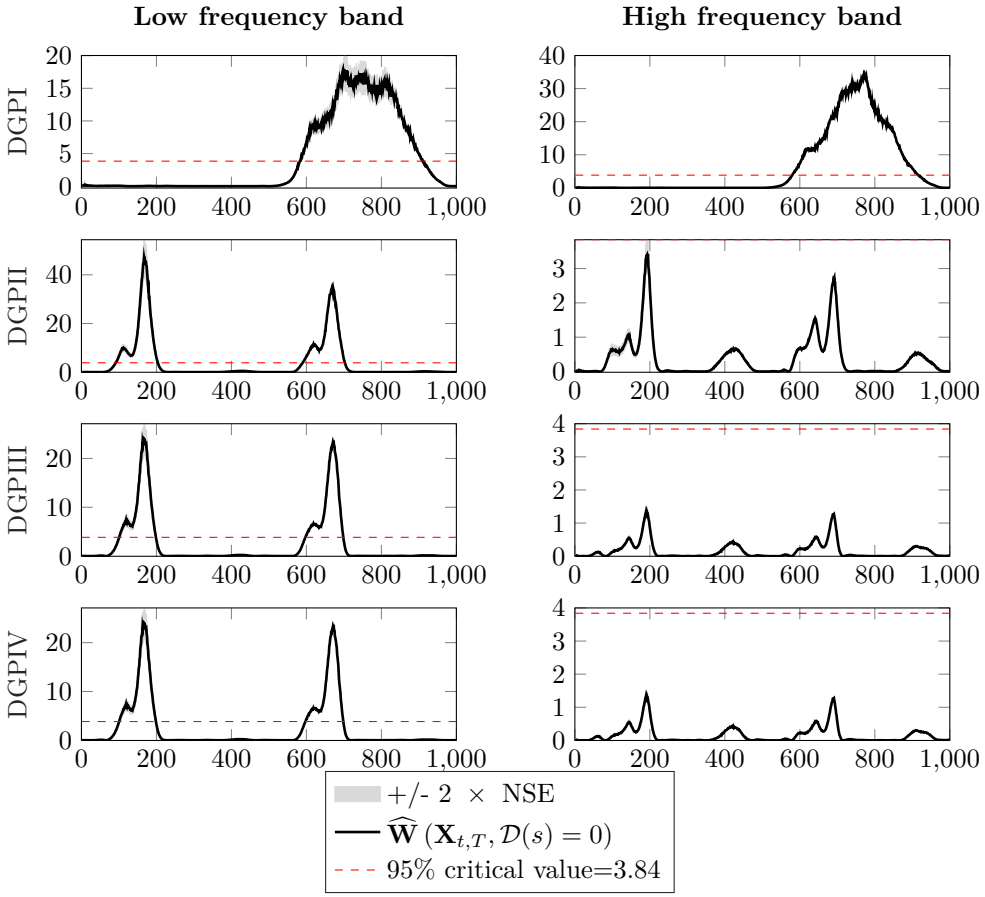

Figure 4: Tests for differences between network connectedness over time
Notes: This figure plots the test statistics and $+/-$ 2 $\times$ the numerical standard errors for heterogeneities between network connectedness measures across different frequency bands from four DGPs. The test statistic is $\widehat{\mathbf{W}}(\boldsymbol{X}_{t,T}, \mathcal{D}(s) = 0) - 1$ where $\mathcal{D}(u) = \widehat{\mathcal{C}}(u_1, d) - \widehat{\mathcal{C}}(u_2, d)$ with $u_1 = 1$, $u_2 = \{2, 3, \ldots, 1,000\}$. Here $d \in (0,\ \pi/5)$ corresponds to the low frequency band (LHS plots) and $d \in (\pi/5,\ \pi)$ corresponds to the high frequency band (RHS plots). The dashed black line is the 5% critical value from the $\chi^2(1)$ distribution =3.84. Values greater than 3.84 reject the null hypothesis of equivalent network connections at time u_1, u_2 across frequency band d. DGPI (top row) is a TVP VAR(2) model with Gaussian errors, we introduce a break in the time-varying autoregressive matrices and contemporaneous relations from observation 500 that induces large connections across the high frequency band. DGPII (second row) is a TVP VAR(2) where time-varying intercepts and autoregressive matrices following sin wave functions with a stochastic error, time-varying covariance matrix where the off-diagonals follow sin wave functions with a stochastic error, and the diagonal elements follow a stationary AR(1) processes. DGPIII (third row) is a TVP VAR(2) model with student-t errors, time-varying intercepts and autoregressive matrices following sin wave functions with a stochastic error, time-varying covariance matrix where the off-diagonals follow sin wave functions with a stochastic error, and the diagonal elements follow a stationary AR(1) processes. DGPIV (fourth row) is the same as DGPII, but with an increase in the periodicity of the respective sin wave functions the time-varying intercepts and autoregressive matrices follow.

the same frequency band, where we should expect to see such differences.

4 Monitoring Persistence in Uncertainty Networks using S&P500 Stocks

Changes in uncertainty can play a key role in driving business cycles and financial turmoil (Bloom et al., 2018). Identifying the sources of such risks is a focus for researchers and practitioners. Some quantify systemic risks emanating from financial markets and sectors (e.g. Billio et al., 2012; Acemoglu et al., 2015), while others examine how sectoral shocks affect aggregate fluctuations (e.g. Gabaix, 2011; Acemoglu et al., 2017; Atalay, 2017). Related to the above, and in response to financial crises, many countries are implementing policies to monitor systemic risk and financial stability. Therefore, we use our framework to identify new measures of transitory and persistent linkages for S&P500 constituents. In section 4.1, we examine sectoral connectedness as well as network structures at a granular level for financial firms.

We use high-frequency tick data for all stocks listed on the S&P500 from 5 July 2005 to 31 August 2018 and compute realised volatility (RV) for all stocks in the sample. To obtain firm-level RVs, we restrict our analysis to five-minute returns during New York Stock Exchange (NYSE) trading hours (i.e. 09:30-16:00). The data are time-synchronised using the same timestamps, eliminating transactions executed on Saturdays and Sundays, US holidays, 24-26 December and 31 December to 2 January due to low activity on these days. This leaves us with 3278 trading days. After cleaning the data, we are left with a cross section of 496 stocks.

To obtain our network connectivity measures, we estimate a TVP VAR model on N=496 stocks with p=2 lags on our T=3278 days of data. We estimate our horizon-specific dynamic network measures on a 48-core server. For each $t \in \{1, 2, \dots, T\}$, we generate 500 simulations of the (quasi) posterior distribution, resulting in a total estimation time of 10 days. We define transitory (short-term) network links as those that form over a 1-5 day horizon, and persistent (long-term) network links at horizons greater than 5 days (i.e. 5 days to the ∞ horizon). Our choice of these horizons stems from the existing literature on volatility modelling using high-frequency data, which shows that daily and weekly fluctuations contain salient information for future volatility (e.g. Corsi and Renò, 2012).[19]

Figure 5 plots measures of transitory and persistent network connectedness from 5 July 2005 to 31 August 2018. Overall, there are significant differences in the level of horizon-specific connectedness across our estimation sample. In general, long-term linkages are muted during periods of economic/financial calm. However, it is clear that long-term connectedness spikes during periods of economic recession or major stock market events. For example, long-term connectedness starts to rise in 2006 and continues to rise during

[19]In this case, due to the size of the system, we diagonalise the covariance matrix of the VAR as an additional precaution to avoid overfitting. For the sectoral network measures in section 4.1, we use a full covariance matrix.

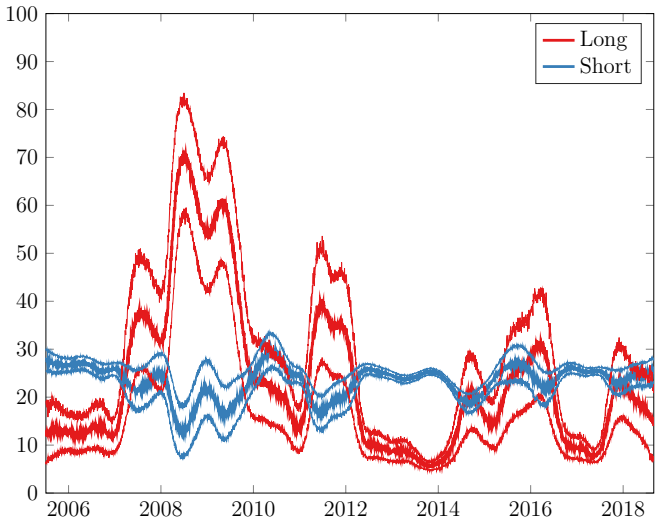

Figure 5: Horizon specific dynamic total network connectedness for S&P500 constituents
This figure plots the posterior median and 1-standard deviation percentiles of horizon specific dynamic total network connectedness, $\mathcal{C}(u,d)$, $d \in \{\mathrm{S},\mathrm{L}\}$ from July 5, 2005 to August 31, 2018. S refers to the short-term, transitory connectedness, which we define as 1 day to 1 week; and L refers to long-term, persistent connectedness, which we define as horizons $>$ 1 week. The spectrum with which the horizons stem from link to the frequency with which we observe the data.

the 2007-2009 recession. In addition, we can see that long-term connectedness increases during 2010-2012. This may be due to fears of contagion from the European sovereign debt crisis, the 2010 flash crash and when the S&P500 entered a bear market in 2011, albeit a short-lived one. We can also see an increase in short- and long-term connectedness in mid-to-late 2015, which is consistent with the stock market sell-off starting in August 2015; this may also be related to fears of contagion from the Chinese stock market crash in late 2015. Overall, we see that long-term network connectedness increases during periods of high systemic risk across our sample.

4.1 Transitory and Persistent Network Connectedness of S&P500 Sectors

Here we focus on the overall network connectedness driven by transitory and persistent shocks to companies in a given sector. We classify stocks into eleven main sectors according to the Global Industry Classification Standard (GICS)[20]. These are: Consumer Discretionary (COND) with 73 stocks; Consumer Staples (CONS) with 34 stocks; Health Care (HLTH) with 53 stocks; Industrials (INDU) with 73 stocks; Information Technology

[20]GICS is an industry taxonomy developed by MSCI and Standard & Poor's for use by the global financial community.

(INFT) with 67 stocks; Materials (MATR) with 33 stocks; Real Estate (REAS) with 29 stocks; Financials (SPF) with 66 stocks; Energy (SPN) with 36 stocks; Communication Services (TELS) with 6 stocks; and Utilities (UTIL) with 26 stocks. Further details, including descriptive statistics (Table D1) for annualised daily RVs, which we compute as $100 \times \sqrt{252 \times RV_t}$ pooling information across companies within each sector over the sample period from 5 July 2005 to 31 August 2018, are reported in Appendix D. The energy sector has the highest mean, while the financial sector has the highest standard deviation, skewness and kurtosis. Overall, we can see that these sectoral RVs show significant differences in terms of the first four moments as well as the minimum and maximum values.

For each of the 11 sectors, we obtain dynamic network measures by estimating a TVP VAR model on all stocks with two lags on our 3278 days of data. On each day, we take 500 draws from the (quasi) posterior distribution.[21]. We define transient and persistent network connectivity in the same way as above.

In Figure 6 we plot the posterior median and 95% confidence bands for transitory and persistent network connectivity for each sector as in Equation 8. These follow directly from the manipulations of the estimates of the dynamic adjacency matrices (see equation 12). Grey bars in these figures represent periods where there are statistically significant differences at the 5% level between transient and persistent network connectivity.[22] Overall, we observe significant differences between these horizon-specific networks for each sector. In general, network connectedness due to the persistent component of shocks exceeds that due to the transitory component of shocks during periods of market turbulence. Then, during periods of calm, the transitory part of the networks becomes more pronounced.

Comparing these measures across sectors, we can see from the real estate and financial sectors that surges in network connectedness driven by persistent shocks drive uncertainty in the sector for much longer periods of time during the Great Recession relative to other sectors, e.g. CONS and HLTH. In addition, the magnitude of persistent network connectedness from the real estate and financial sectors is much greater. Note also that throughout this period we observe much more frequent evidence in favour of statistical differences between networks driven by transitory and persistent shocks for these sectors; particularly relative to COND, CONS, HLTH and MATR. Although the other sectors show spikes in persistent network connectedness, these do not occur until around February 2008. This highlights how long-term systemic risks within the real estate and financial sectors intensify during this period and are stronger relative to other sectors.

We also see clear spikes in persistent sectoral network connectedness in May-October 2011 and again in 2015-2016. The former coincides with the S&P500 entering a bear market and the latter with declines in major stock markets around the world. In 2015-2016, we see much lower long-term connectedness of consumer staples, utilities, real estate and telecoms relative to other sectors. We expect the long-term linkages of industrials,

[21] We estimate our horizon-specific dynamic network measures on a 64-core server, resulting in a total estimation time of about 4–5 hours to obtain network estimates from the 11 sectors.

[22] We plot the values of these test statistics in Figure D1 in Appendix D.

materials, energy, information technology and financials to be strong during this period, as the decline in global equity markets is linked to falling commodity prices and the depreciation of Asian currencies against the US dollar.

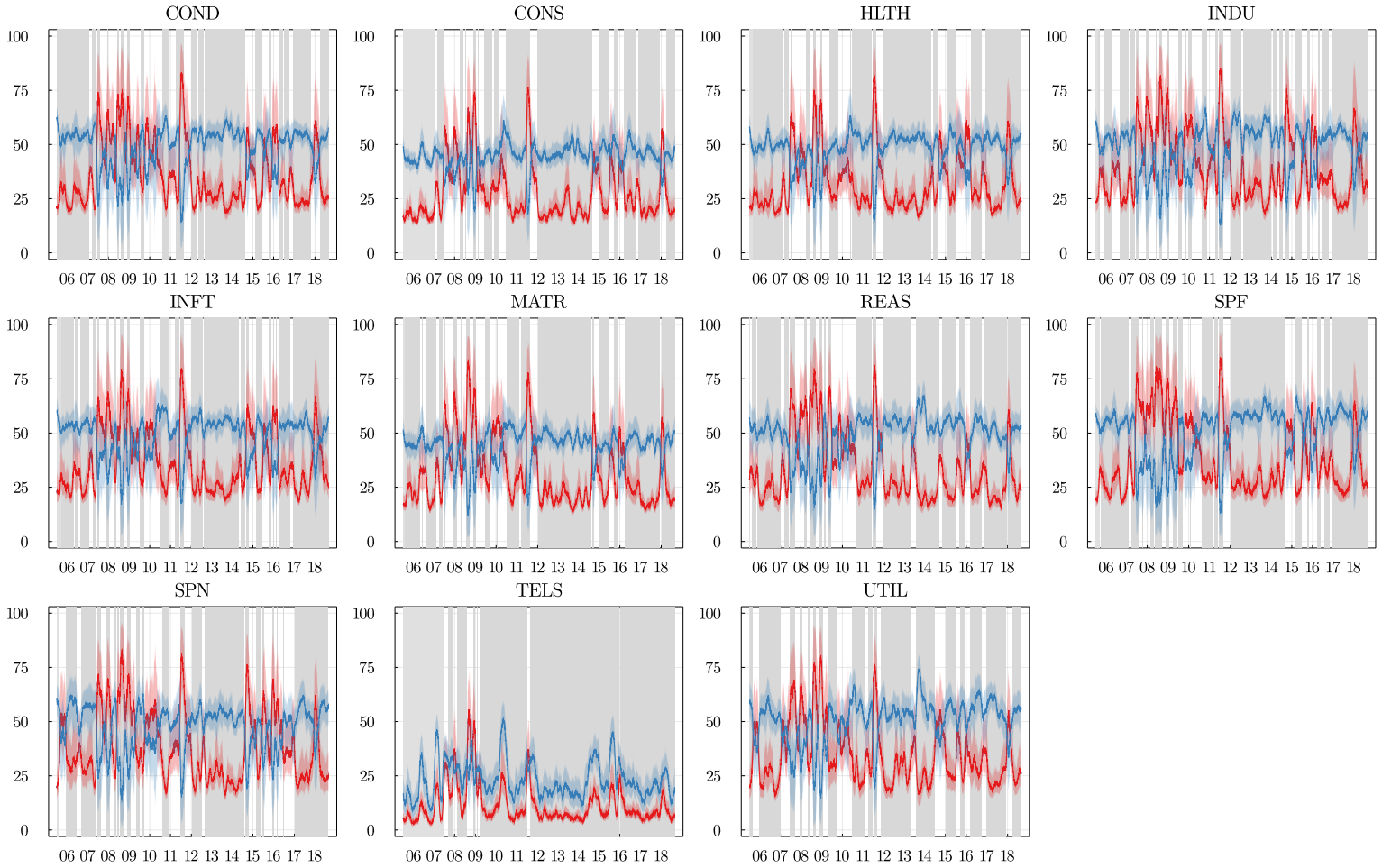

Figure 6: Dynamics and Persistence in the U.S. Uncertainty Networks
This figure plots the (quasi) posterior median and 95% confidence bands of network connectedness specific to the transitory, or short-term (in blue) and persistent, or long-term (in red) shocks to realized volatility of the S&P500 sectors: Consumers Discretionary (COND), Consumer Staples (CONS), Health Care (HLTH), Industrials (INDU), Information Technology (INFT), Materials (MATR), Real Estate (REAS), Financials (SPF), Energy (SPN), Comunication Services (TELS), and Utilities (UTIL) from July 8, 2005 to August 31, 2018. We define transitory (short-term), as connections made over the 1 day to 1 week horizon; we characterize connections greater than 1 week as persistent (long-term). Grey bars indicate periods with significant heterogeneity in persistence structures.

In terms of sectoral network linkages due to transitory shocks, the main differences we observe are in the size of the linkages. Note, however, that there are subtle differences in the time profiles. For each sector, we see that transitory linkages are much stronger during periods of uncertainty at the firm level than persistent linkages during periods of calm.

Digging deeper into our investigation of whether there are significant differences between transitory and persistent network connectedness for each sector, two main points emerge. First, we document significant heterogeneity in persistence over substantial periods of time. Figure 6 shows that during periods of tranquillity, transitory network linkages between sectors are stronger relative to linkages from persistent shocks. During periods of turbulence, however, our results show that persistent linkages intensify in all sectors, but their magnitude differs considerably.[23] Second, statistical differences between transitory and persistent connectedness in one sector do not necessarily imply differences in other sectors. For example, over the 2014-2017 period, we see clear differences in the rejections that transitory and persistent connectedness are equivalent in REAS, SPF, SPN and UTIL.

In general, the temporal nature of our results aligns well with Bianchi et al. (2019), which documents a regime dependent impact of systemic risk on financial markets. They also provide empirical support for Gabaix (2011) and Acemoglu et al. (2012), and uncover new measures of sectoral uncertainties (or sector-wide risks). Testing for statistical differences between transitory and persistent sectoral network connectedness adds further substance to our suggestion that one should consider dynamic network structures that form across frequency bands.

For researchers, our measures of network connectedness may contain useful information for real economy or forecasting purposes; there is already evidence that network connectedness contains predictive content for the real economy (e.g. Baruník et al., 2020). For practitioners, tracking persistence in sectoral networks can be useful for informing macroprudential policy. This is because one can use these measures as online monitoring tools to study the evolution and persistence of sectoral network connectedness.

Although we provide evidence of substantial heterogeneity in the persistence of network structures, there are commonalities in the time profiles. We attribute this to the high degree of correlation between the data used to proxy uncertainty in our investigation. Herskovic et al. (2016) exploits the correlation structure of idiosyncratic return volatilities and shows that a common factor among the drivers of firm-level volatilities has pricing implications. Our network connectedness measures, by definition, refer to this correlation structure and provide an aggregate description of the network at each point in time. However, our network measures are more informative. We are able to obtain measures that contain information about the overall network structure. These may relate to directional connections (in and out degrees) or concentration (i.e. a high influence of a small number of firms/nodes on the overall network); both of which have been shown to contain information with economic implications (see e.g. Herskovic, 2018; Herskovic

[23] From Figure D1 in Appendix D we also document that the magnitude of the differences evolves substantially over time.

et al., 2020).

4.2 Network Connections at a Granular Level

By focusing on shocks to a single financial institution that affect the wider system, our research contributes to the large literature on measuring systemic risk. Many understand systemic risk as many market participants realizing severe losses as a result of propagation throughout the (financial) system.[24] During periods of financial turbulence, uncertainty shocks, the drying up of liquidity, and insolvencies have the ability to spread rapidly affecting many institutions across the market. Lessons from recent financial crises spur the demand for financial regulations in order to mitigate firm behaviors consistent with increasing systemic risk.

From a prudential perspective, safeguarding against this type of risk requires quantifying systemic risk. The existing literature offers many measures of such types of risk, these include: the expected shortfall measure of Acharya et al. (2017); Co-Value-at-Risk (CoVaR) (Adrian and Brunnermeier, 2016); and network connectedness measures (see e.g. Demirer et al., 2018). The former measures relate to specific risk channels and as such can aid in calibrating regulatory tools. The latter approach permits one to quantify the overall influence of individual institutions to overall systemic risk, and hence identify systemically important financial institutions (SIFIs).

Our approach relates closely with the above. However, under our framework one can characterise SIFIs, or important variables for different applications, throughout time as well as understanding whether the influence is persistent or transitory in nature. The benefit of this is twofold. First, enhancing prudential authorities' understanding of whether SIFIs influence are transient or long lasting can help refine policies in order to mitigate adverse firm behavior. For example, one could tailor policies by increasing the capital requirements of SIFIs for those who transmit persistent shocks that contribute significantly to systemic risk. Second, since systemic risk threatens the stability of the entire financial sector, knowing the frequency-specific sources of instability facilitates monitoring of changes to the such risks.

To illustrate how policy makers might use our approach, we examine the network structures of the financial sector at the granular level of 65 firms. Since our application uses daily data we have transitory and persistent network structures at every observation in our sample. We therefore focus on two different dates. Figure 7 shows the network structures driven by transitory and persistent shocks for the SPF sector on October 24, 2008. This date corresponds to the start of the global financial crisis. Figure 8 reports the corresponding network structures one year later on October 24, 2009. For each plot, arrows indicate the direction and strength of the connections, while a transparent (full colour) vertex indicates a stock that sends (receives) more shocks than it receives (sends). The size of the vertices indicates the net direction of the connections.

We can see that on 24 October 2008, the persistent links are larger relative to the transitory ones, suggesting that shocks within the financial sector create links that re-

[24] For a comprehensive review of the literature on systemic risk, see Benoit et al. (2017).

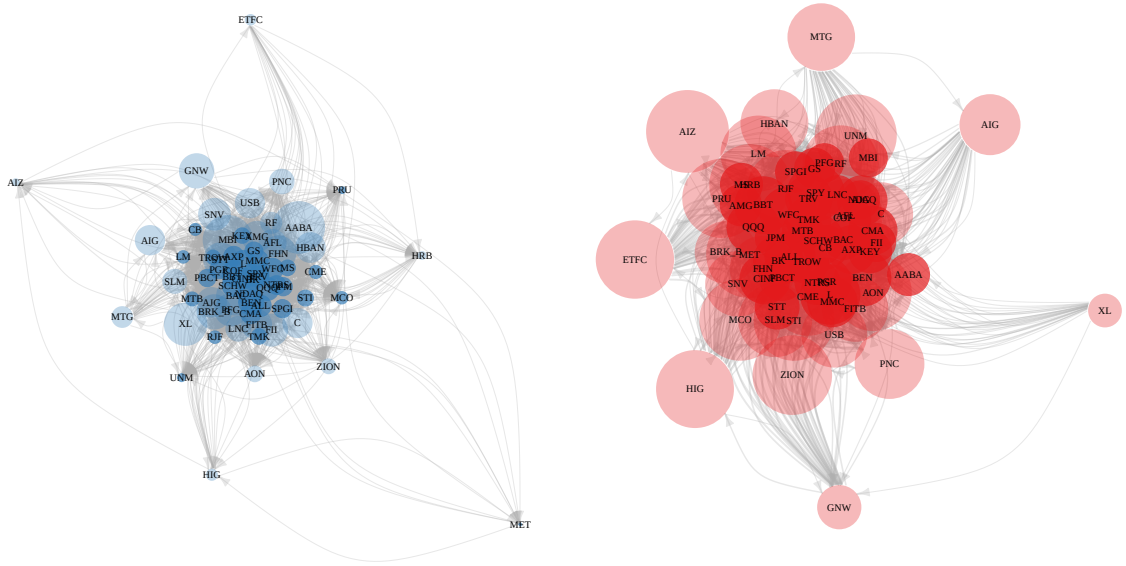

Figure 7: Transitory and persistent networks of finance: 24 October 2008
The left (right) figure shows the network connections between the assets comprising the SPF sector driven by transitory (persistent) shocks on 24 October 2008, corresponding to the day when the VIX peaked. Arrows indicate the direction of the connections and the strength of the lines indicates the strength of the connections. Grey (black) vertices indicate firms that receive (send) more shocks than they send (receive). The size of the vertices indicates the net amount of shocks.

late to the long-term. This suggests that systemic risk within the systems stems from persistent network structures. Now looking one year later, it is clear that connections are far weaker across both transitory and persistent network structures and systemic risk is relatively lower. The main takeaway from these plots is the strong differences in the overall structure of the horizon-specific networks.

In Appendix D, we plot heatmaps showing the strength of financial institutions connections across persistent and transitory network structures for these same two dates in Figures D2–D3. Persistent shocks tend to drive the links with greater strength, and so we focus our discussion here. Zooming in to examine the contribution of specific firms, we can see that Truist Financial Corp (BBT), Franklin Resources (BEN), Loews Corporation (L), SPY, or Wells Fargo & Co (WFC) transmitted persistent shocks to the financial sector and thus are identified as SIFIs that affected the system with persistent shocks. We can follow the contributions from the columns of the heatmap in Figure D2. The impact of a bank increases with the number of rows containing a stronger and warmer red colour. As we can see, those banks we name above affect many other financial institutions at the start of the financial crisis, and such impacts are long-lasting.

Conversely, Metlife (MET), Moody's (MCO), Unum (UNM), H&R Block (HRB) and Assurant (AIZ) receiv the most shocks on 24 October 2008. One year later, we can see that the structure changes dramatically. While WFC, SPY and L seem to be strong SIFIs, BEN is a nearly non-contributing bank. This highlights how our approach tracks dynamics of key financial institutions within the system and their influence across persistent and transitory network structures. For completeness, we rank all institutions in the financial sector according to the strength of transitory and persistent shocks they transmit/receive during the same two dates in Table D2 in Appendix D.

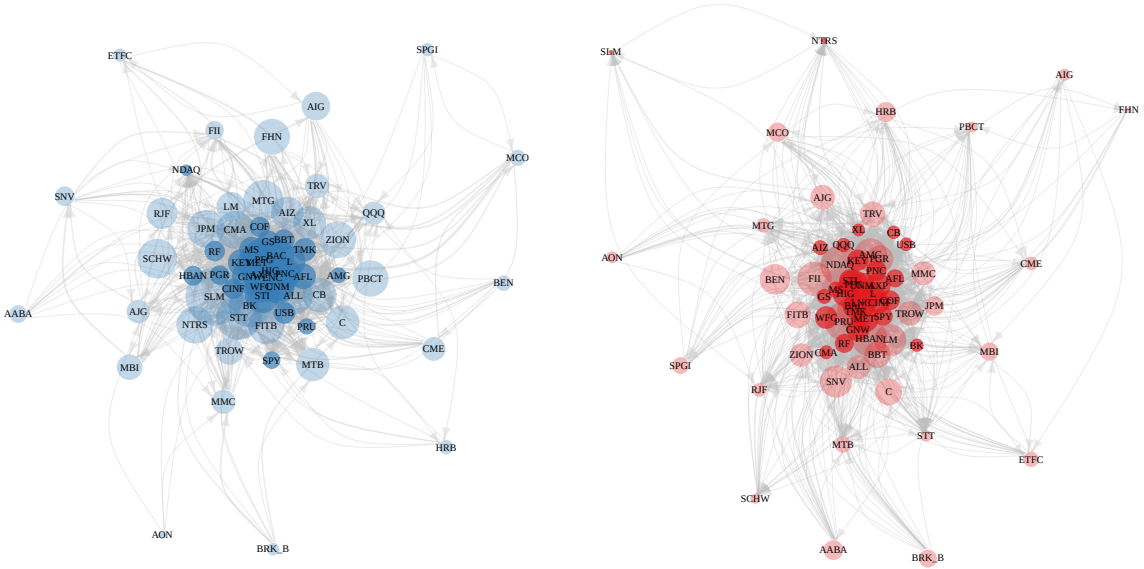

Figure 8: Transitory and Persistent Network of Financials: 24 October 2009
The left (right) figure shows the network connections between the assets that make up the SPF sector by transitory (persistent) connectedness. Arrows indicate the direction of the connections and the strength of the lines indicates the strength of the connections. Grey (black) vertices denote firms that receive (send) more shocks than they send (receive). The size of the vertices indicates the net amount of shocks.

Overall, these network structures show how the role of a firm can change not only over time, but also in terms of persistence. Researchers may wish to use these network structures to assess the pricing implications of such risks. For example, users could assess the role of directional linkages in an empirical asset pricing application that builds on the theoretical work of Branger et al. (2020). Our framework also allows one to track dynamic network structures that would complement studies such as Herskovic (2018) and Gofman et al. (2020). The advantage of our approach is that one does not have to rely on monthly or annual data to capture such networks. Finally, looking at dynamic adjacency matrices can help economists understand how shocks dynamically determine

network structures in models of monetary policy (e.g. Pasten et al., 2020) or the Phillips curve (e.g. Rubbo, 2020).

5 Conclusion

This paper proposes a novel framework for measuring dynamic network connections in a multivariate data system. We track dynamic connections driven by different degrees of persistence using a spectral decomposition of time-varying variance decomposition matrices. Our approach properly accounts for the characteristics of the shocks that create such links. We outline a procedure that allows one to test for statistical differences in connectedness over time and frequency. We provide Monte Carlo evidence that our measures are able to reliably track connectedness and correctly identify statistical differences from different data generating processes.

Empirically, we show that transitory and persistent measures of network connectedness improve our understanding of systemic risks arising from uncertainty networks. This is because our approach allows one to track connectedness across the transitory and persistent components of shocks. This is particularly useful during periods of heightened uncertainty, as our measures indicate whether systemic risks from network connections are transitory or persistent in nature. Ultimately, this could lead to better decision-making by macroprudential supervisors and investment decisions by market participants.

References

Acemoglu, D., V. M. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi (2012). The network origins of aggregate fluctuations. *Econometrica 80*(5), 1977–2016.

Acemoglu, D., A. Ozdaglar, and A. Tahbaz-Salehi (2015). Systemic risk and stability in financial networks. *American Economic Review 105*(2), 564–608.

Acemoglu, D., A. Ozdaglar, and A. Tahbaz-Salehi (2017). Microeconomic origins of macroeconomic tail risks. *American Economic Review 107*(1), 54–108.

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

Acharya, V. V., L. H. Pedersen, T. Philippon, and M. Richardson (2017). Measuring systemic risk. *The review of financial studies 30*(1), 2–47.

Adrian, T. and M. K. Brunnermeier (2016). Covar. *American Economic Review 106*(7), 1705–41.

Atalay, E. (2017). How important are sectoral shocks? *American Economic Journal: Macroeconomics 9*(4), 254–80.

Balke, N. S. and M. E. Wohar (2002). Low-frequency movements in stock prices: A state-space decomposition. *Review of Economics and Statistics 84*(4), 649–667.

Bandi, F. and A. Tamoni (2017). Business-cycle consumption risk and asset prices. *Available at SSRN 2337973*.

Bandi, F. M., S. Chaudhuri, A. W. Lo, and A. Tamoni (2021). Spectral factor models. *Journal of Financial Economics forthcoming*.

Bansal, R., D. Kiku, and A. Yaron (2010). Long run risks, the macroeconomy, and asset prices. *American Economic Review 100*(2), 542–46.

Bansal, R. and A. Yaron (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance 59*(4), 1481–1509.

Barigozzi, M., M. Hallin, S. Soccorsi, and R. von Sachs (2020). Time-varying general dynamic factor models and the measurement of financial connectedness. *Journal of Econometrics*.

Baruník, J., M. Bevilacqua, and R. Tunaru (2020). Asymmetric network connectedness of fears. *Review of Economics and Statistics*, 1–41.

Baruník, J. and T. Křehlík (2018). Measuring the frequency dynamics of financial connectedness and systemic risk. *Journal of Financial Econometrics 16*(2), 271–296.

Barunik, J., T. Krehlik, and L. Vacha (2016). Modeling and forecasting exchange rate volatility in time-frequency domain. *European Journal of Operational Research 251*(1), 329–340.

Benoit, S., J.-E. Colliard, C. Hurlin, and C. Pérignon (2017). Where the risks lie: A survey on systemic risk. *Review of Finance 21*(1), 109–152.

Bianchi, D., M. Billio, R. Casarin, and M. Guidolin (2019). Modeling systemic risk with Markov switching graphical SUR models. *Journal of Econometrics 210*(1), 58–74.

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*(3), 535–559.

Bloom, N., M. Floetotto, N. Jaimovich, I. Saporta-Eksten, and S. J. Terry (2018). Really uncertain business cycles. *Econometrica 86*(3), 1031–1065.

Branger, N., P. Konermann, C. Meinerding, and C. Schlag (2020). Equilibrium asset pricing in directed networks. *Review of Finance Forthcoming*.

Brennan, M. J. and Y. Zhang (2018). Capital asset pricing with a stochastic horizon. *Journal of Financial and Quantitative Analysis*, 1–45.

Calabrese, R. and S. A. Osmetti (2019). A new approach to measure systemic risk: A bivariate copula model for dependent censored data. *European Journal of Operational Research 279*(3), 1053–1064.

Chan, J. C. (2020). Large Bayesian VARs: A flexible Kronecker error covariance structure. *Journal of Business & Economic Statistics 38*(1), 68–79.

Chaudhuri, S. E. and A. W. Lo (2019). Dynamic alpha: A spectral decomposition of investment performance across time horizons. *Management Science 65*(9), 4440–4450.

Cogley, T., G. E. Primiceri, and T. J. Sargent (2010). Inflation-gap persistence in the US. *American Economic Journal: Macroeconomics 2*(1), 43–69.

Corsi, F. and R. Renò (2012). Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling. *Journal of Business & Economic Statistics 30*(3), 368–380.

Dahlhaus, R. (1996). On the kullback-leibler information divergence of locally stationary processes. *Stochastic processes and their applications 62*(1), 139–168.

Dahlhaus, R. (2000). A likelihood approximation for locally stationary processes. *Annals of Statistics 28*(6), 1762–1794.

Dahlhaus, R., W. Polonik, et al. (2009). Empirical spectral processes for locally stationary time series. *Bernoulli 15*(1), 1–39.

Demirer, M., F. X. Diebold, L. Liu, and K. Yilmaz (2018). Estimating global bank network connectedness. *Journal of Applied Econometrics 33*(1), 1–15.

Dew-Becker, I. and S. Giglio (2016). Asset pricing in the frequency domain: theory and empirics. *Review of Financial Studies 29*(8), 2029–2068.

Diebold, F. X. and K. Yilmaz (2014). On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics 182*(1), 119–134.

Drechsler, I. and A. Yaron (2011). What's vol got to do with it. *Review of Financial Studies 24*(1), 1–45.

Engle, R. and B. Kelly (2012). Dynamic equicorrelation. *Journal of Business & Economic Statistics 30*(2), 212–228.

Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica 79*(3), 733–772.

Gallant, A. R. and H. White (1988). A unified theory of estimation and inference for nonlinear dynamic models. *Blackwell*.

Garvey, M. D., S. Carnovale, and S. Yeniyurt (2015). An analytical framework for supply network risk propagation: A Bayesian network approach. *European Journal of Operational Research 243*(2), 618–627.

Geraci, M. V. and J.-Y. Gnabo (2018). Measuring interconnectedness between financial institutions with Bayesian time-varying vector autoregressions. *Journal of Financial and Quantitative Analysis 53*(3), 1371–1390.

Gerrard, R., I. Kyriakou, J. P. Nielsen, and P. Vodička (2022). On optimal constrained investment strategies for long-term savers in stochastic environments and probability hedging. *European Journal of Operational Research*.

Giglio, S., M. Maggiori, and J. Stroebel (2015). Very long-run discount rates. *The Quarterly Journal of Economics 130*(1), 1–53.

Gofman, M., G. Segal, and Y. Wu (2020). Production networks and stock returns: The role of vertical creative destruction. *The Review of Financial Studies 33*(12), 5856–5905.

Hahn, J., P. Todd, and W. Van der Klaauw (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica 69*(1), 201–209.

Haven, E., X. Liu, and L. Shen (2012). De-noising option prices with the wavelet method. *European Journal of Operational Research 222*(1), 104–112.

Herskovic, B. (2018). Networks in production: Asset pricing implications. *Journal of Finance 73*(4), 1785–1818.

Herskovic, B., B. Kelly, H. Lustig, and S. Van Nieuwerburgh (2016). The common factor in idiosyncratic volatility: Quantitative asset pricing implications. *Journal of Financial Economics 119*(2), 249–283.

Herskovic, B., B. Kelly, H. Lustig, and S. Van Nieuwerburgh (2020). Firm volatility in granular networks. *Journal of Political Economy 128*(11), 4097–4162.

Kadiyala, K. R. and S. Karlsson (1997). Numerical methods for estimation and inference in Bayesian VAR-models. *Journal of Applied Econometrics 12*(2), 99–132.

Koop, G., R. Leon-Gonzalez, and R. W. Strachan (2010). Dynamic probabilities of restrictions in state space models: an application to the Phillips curve. *Journal of Business & Economic Statistics 28*(3), 370–379.

Li, Y., X.-B. Liu, and J. Yu (2015). A Bayesian chi-squared test for hypothesis testing. *Journal of Econometrics 189*(1), 54–69.

Li, Y., J. Yu, and T. Zeng (2020). Deviance information criterion for latent variable models and misspecified models. *Journal of Econometrics 216*(2), 450–493.

Li, Y., T. Zeng, and J. Yu (2014). A new approach to Bayesian hypothesis testing. *Journal of Econometrics 178*, 602–612.

Liu, X., Y. Li, J. Yu, and T. Zeng (2022). Posterior-based Wald-type statistics for hypothesis testing. *Journal of Econometrics 230*(1), 83–113.

Lütkepohl, H. (1990). Asymptotic distributions of impulse response functions and forecast error variance decompositions of vector autoregressive models. *Review of Economics and Statistics*, 116–125.

Pasten, E., R. Schoenle, and M. Weber (2020). The propagation of monetary policy shocks in a heterogeneous production economy. *Journal of Monetary Economics 116*, 1–22.

Pesaran, H. H. and Y. Shin (1998). Generalized impulse response analysis in linear multivariate models. *Economics letters 58*(1), 17–29.

Petrova, K. (2019). A quasi-Bayesian local likelihood approach to time varying parameter VAR models. *Journal of Econometrics*.

Primiceri, G. E. (2005). Time varying structural vector autoregressions and monetary policy. *Review of Economic Studies 72*(3), 821–852.

Richmond, R. J. (2019). Trade network centrality and currency risk premia. *Journal of Finance 74*(3), 1315–1361.

Roueff, F. and A. Sanchez-Perez (2016). Prediction of weakly locally stationary processes by auto-regression. *arXiv preprint arXiv:1602.01942*.

Rubbo, E. (2020). Networks, Phillips curves and monetary policy. *Available https://economics.yale.edu/sites/default/files/rubbo_jmp.pdf*.

Sévi, B. (2014). Forecasting the volatility of crude oil futures using intraday data. *European Journal of Operational Research 235*(3), 643–659.

Sun, E. W. and T. Meinl (2012). A new wavelet-based denoising algorithm for high-frequency financial data mining. *European Journal of Operational Research 217*(3), 589–599.

White, H. (1996). *Estimation, inference and specification analysis*. Number 22. Cambridge university press.

Yang, Z. and Y. Zhou (2017). Quantitative easing and volatility spillovers across countries and asset classes. *Management Science 63*(2), 333–354.

For Online Publication: Technical Appendix

A Proofs

Proposition 1. Let us have the VMA(∞) representation of the locally stationary TVP VAR model (Dahlhaus et al., 2009; Roueff and Sanchez-Perez, 2016)

$$\boldsymbol{X}_{t,T} = \sum_{h=-\infty}^{\infty} \boldsymbol{\Psi}_{t,T}(h)\boldsymbol{\epsilon}_{t-h} \tag{A.1}$$

$\boldsymbol{\Psi}_{t,T}(h) \approx \boldsymbol{\Psi}(t/T,h)$ is a stochastic process satisfying $\sup_\ell ||\boldsymbol{\Psi}_t - \boldsymbol{\Psi}_\ell||^2 = O_p(h/t)$ for $1 \leq h \leq t$ as $t \to \infty$, hence in a neighborhood of a fixed time point $u = t/T$ the process $\boldsymbol{X}_{t,T}$ can be approximated by a stationary process $\widetilde{\boldsymbol{X}}_t(u)$

$$\widetilde{\boldsymbol{X}}_t(u) = \sum_{h=-\infty}^{\infty} \boldsymbol{\Psi}(u,h)\boldsymbol{\epsilon}_{t-h} \tag{A.2}$$

with $\boldsymbol{\epsilon}$ being *iid* process with $\mathbb{E}[\boldsymbol{\epsilon}_t] = 0$, $\mathbb{E}[\boldsymbol{\epsilon}_s\boldsymbol{\epsilon}_t] = 0$ for all $s \neq t$, and the local covariance matrix of the errors $\boldsymbol{\Sigma}(u)$. Under suitable regularity conditions $|\boldsymbol{X}_{t,T} - \widetilde{\boldsymbol{X}}_t(u)| = O_p(|t/T - u| + 1/T)$.

Since the errors are assumed to be serially uncorrelated, the total local covariance matrix of the forecast error conditional on the information at time $t-1$ is given by

$$\boldsymbol{\Omega}(u,H) = \sum_{h=0}^{H} \boldsymbol{\Psi}(u,h)\boldsymbol{\Sigma}(u)\boldsymbol{\Psi}^\top(u,h). \tag{A.3}$$

Next, we consider the local covariance matrix of the forecast error conditional on knowledge of today's shock and future expected shocks to k-th variable. Starting from the conditional forecasting error,

$$\boldsymbol{\xi}^k(u,H) = \sum_{h=0}^{H} \boldsymbol{\Psi}(u,h)\Big[\boldsymbol{\epsilon}_{t+H-h} - \mathbb{E}(\boldsymbol{\epsilon}_{t+H-h}|\boldsymbol{\epsilon}_{k,t+H-h})\Big], \tag{A.4}$$

assuming normal distribution of $\boldsymbol{\epsilon}_t \sim N(0,\boldsymbol{\Sigma})$, we obtain[25]

$$\mathbb{E}(\boldsymbol{\epsilon}_{t+H-h}|\boldsymbol{\epsilon}_{k,t+H-h}) = \sigma_{kk}^{-1}\Big[\boldsymbol{\Sigma}(u)\Big]_{.k}\boldsymbol{\epsilon}_{k,t+H-h} \tag{A.5}$$

and substituting (A.5) to (A.4), we obtain

$$\boldsymbol{\xi}^k(u,H) = \sum_{h=0}^{H} \boldsymbol{\Psi}(u,h)\Big[\boldsymbol{\epsilon}_{t+H-h} - \sigma_{kk}^{-1}\Big[\boldsymbol{\Sigma}(u)\Big]_{.k}\boldsymbol{\epsilon}_{k,t+H-h}\Big]. \tag{A.6}$$

[25]Note to notation: $[\boldsymbol{A}]_{j,k}$ denotes the jth row and kth column of matrix $\boldsymbol{A}$ denoted in bold. $[\boldsymbol{A}]_{j,\cdot}$ denotes the full jth row; this is similar for the columns. A $\sum A$, where A is a matrix that denotes the sum of all elements of the matrix A.

Finally, the local forecast error covariance matrix is

$$\boldsymbol{\Omega}^k(u,H) = \sum_{h=0}^{H} \boldsymbol{\Psi}(u,h)\boldsymbol{\Sigma}(u)\boldsymbol{\Psi}^\top(u,h) - \sigma_{kk}^{-1}\sum_{h=0}^{H} \boldsymbol{\Psi}(u,h)\Big[\boldsymbol{\Sigma}(u)\Big]_{.k}\Big[\boldsymbol{\Sigma}(u)\Big]_{.k}^\top\boldsymbol{\Psi}^\top(u,h). \quad \text{(A.7)}$$

Then

$$\Big[\boldsymbol{\Delta}(u,H)\Big]_{(j)k} = \Big[\boldsymbol{\Omega}(u,H) - \boldsymbol{\Omega}^k(u,H)\Big]_{j,j} = \sigma_{kk}^{-1}\sum_{h=0}^{H}\left(\Big[\boldsymbol{\Psi}(u,h)\boldsymbol{\Sigma}(u)\Big]_{j,k}\right)^2 \quad \text{(A.8)}$$

is the unscaled local H-step ahead forecast error variance of the j-th component with respect to the innovation in the k-th component. Scaling the equation with H-step ahead forecast error variance with respect to the jth variable yields the desired time varying generalized forecast error variance decompositions (TVP GFEVD)

$$\Big[\boldsymbol{\theta}(u,H)\Big]_{j,k} = \frac{\sigma_{kk}^{-1}\sum_{h=0}^{H}\left(\Big[\boldsymbol{\Psi}(u,h)\boldsymbol{\Sigma}(u)\Big]_{j,k}\right)^2}{\sum_{h=0}^{H}\Big[\boldsymbol{\Psi}(u,h)\boldsymbol{\Sigma}(u)\boldsymbol{\Psi}^\top(u,h)\Big]_{j,j}} \quad \text{(A.9)}$$

Next, we derive the frequency representation of the quantity in (A.9) using the fact that unique time varying spectral density of $\boldsymbol{X}_{t,T}$ at frequency ω which is locally the same as the spectral density of $\widetilde{\boldsymbol{X}}_t(u)$ at $u = t/T$ can be defined as a Fourier transform of VMA(∞) filtered series over frequencies $\omega \in (-\pi,\pi)$ as

$$\boldsymbol{S}_{\boldsymbol{X}}(u,\omega) = \sum_{h=-\infty}^{\infty}\mathbb{E}\Big[\boldsymbol{X}_{t+h}(u)\boldsymbol{X}_t^\top(u)\Big]e^{-i\omega h} = \Big\{\boldsymbol{\Psi}(u,e^{-i\omega})\Big\}\boldsymbol{\Sigma}(u)\Big\{\boldsymbol{\Psi}(u,e^{+i\omega})\Big\}^\top, \quad \text{(A.10)}$$

where we consider a time varying frequency response function $\boldsymbol{\Psi}(u,e^{-i\omega}) = \sum_h e^{-i\omega h}\boldsymbol{\Psi}(u,h)$ which can be obtained as a Fourier transform of the coefficients with $i = \sqrt{-1}$.

Letting $H \to \infty$, we have time varying generalized forecast error variance decompositions

$$\Big[\boldsymbol{\theta}(u)\Big]_{j,k} = \frac{\sigma_{kk}^{-1}\sum_{h=0}^{\infty}\left(\Big[\boldsymbol{\Psi}(u,h)\boldsymbol{\Sigma}(u)\Big]_{j,k}\right)^2}{\sum_{h=0}^{\infty}\Big[\boldsymbol{\Psi}(u,h)\boldsymbol{\Sigma}(u)\boldsymbol{\Psi}^\top(u,h)\Big]_{j,j}} = \frac{\mathcal{A}}{\mathcal{B}}. \quad \text{(A.11)}$$

Starting with frequency domain counterpart of the nominator $\mathcal{A}$, we will use the standard integral

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i\omega(r-v)}d\omega = \begin{cases} 1 & \text{for } r = v \\ 0 & \text{for } r \neq v. \end{cases} \quad \text{(A.12)}$$

Using the fact that $\sum_{h=0}^{\infty}\phi(h)\psi(h) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{v=0}^{\infty}\sum_{r=0}^{\infty}\phi(r)\psi(v)e^{i\omega(r-v)}d\omega$, we can rewrite (A.11) as

$$
\begin{aligned}
\sigma_{kk}^{-1}\sum_{h=0}^{\infty}\left(\Big[\boldsymbol{\Psi}(u,h)\boldsymbol{\Sigma}(u)\Big]_{j,k}\right)^2 &= \sigma_{kk}^{-1}\sum_{h=0}^{\infty}\left(\sum_{z=1}^{n}\Big[\boldsymbol{\Psi}(u,h)\Big]_{j,z}\Big[\boldsymbol{\Sigma}(u)\Big]_{z,k}\right)^2 \\
&= \sigma_{kk}^{-1}\frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{r=0}^{\infty}\sum_{v=0}^{\infty}\left(\sum_{x=1}^{n}\Big[\boldsymbol{\Psi}(u,r)\Big]_{j,x}\Big[\boldsymbol{\Sigma}(u)\Big]_{x,k}\right)\left(\sum_{y=1}^{n}\Big[\boldsymbol{\Psi}(u,v)\Big]_{j,y}\Big[\boldsymbol{\Sigma}(u)\Big]_{y,k}\right)e^{i\omega(r-v)}d\omega \\
&= \sigma_{kk}^{-1}\frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{r=0}^{\infty}\sum_{v=0}^{\infty}\left(\sum_{x=1}^{n}\Big[\boldsymbol{\Psi}(u,r,e^{i\omega r})\Big]_{j,x}\Big[\boldsymbol{\Sigma}(u)\Big]_{x,k}\right)\left(\sum_{y=1}^{n}\Big[\boldsymbol{\Psi}(u,v,e^{-i\omega v})\Big]_{j,y}\Big[\boldsymbol{\Sigma}(u)\Big]_{y,k}\right)d\omega \\
&= \sigma_{kk}^{-1}\frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\sum_{r=0}^{\infty}\sum_{x=1}^{n}\Big[\boldsymbol{\Psi}(u,r,e^{i\omega r})\Big]_{j,x}\Big[\boldsymbol{\Sigma}(u)\Big]_{x,k}\right)\left(\sum_{v=0}^{\infty}\sum_{y=1}^{n}\Big[\boldsymbol{\Psi}(u,v,e^{-i\omega v})\Big]_{j,y}\Big[\boldsymbol{\Sigma}(u)\Big]_{y,k}\right)d\omega \\
&= \sigma_{kk}^{-1}\frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\sum_{x=1}^{n}\Big[\boldsymbol{\Psi}(u,e^{i\omega})\Big]_{j,x}\Big[\boldsymbol{\Sigma}(u)\Big]_{x,k}\right)\left(\sum_{y=1}^{n}\Big[\boldsymbol{\Psi}(u,e^{-i\omega})\Big]_{j,y}\Big[\boldsymbol{\Sigma}(u)\Big]_{y,k}\right)d\omega \\
&= \sigma_{kk}^{-1}\frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\Big[\boldsymbol{\Psi}(u,e^{-i\omega})\boldsymbol{\Sigma}(u)\Big]_{j,k}\right)\left(\Big[\boldsymbol{\Psi}(u,e^{i\omega})\boldsymbol{\Sigma}(u)\Big]_{j,k}\right)d\omega \\
&= \sigma_{kk}^{-1}\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\Big[\boldsymbol{\Psi}(u,e^{-i\omega})\boldsymbol{\Sigma}(u)\Big]_{j,k}\right|^2 d\omega
\end{aligned}
\tag{A.13}
$$

Hence we have established that

$$
\mathcal{A} = \sigma_{kk}^{-1}\sum_{h=0}^{\infty}\left(\Big[\boldsymbol{\Psi}(u,h)\boldsymbol{\Sigma}(u)\Big]_{j,k}\right)^2 = \sigma_{kk}^{-1}\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\Big[\boldsymbol{\Psi}(u,e^{-i\omega})\boldsymbol{\Sigma}(u)\Big]_{j,k}\right|^2 d\omega \tag{A.14}
$$

from (A.11), we use the local spectral representation of the VMA coefficients in the second step. The rest is a manipulation with the last step invoking the definition of modulus squared of a complex number to be defined as $|z|^2 = zz^*$. Note that we can use this simplification without loss of generality, because the $VMA(\infty)$ representation that is described by the coefficients $\boldsymbol{\Psi}(u,h)$ has a spectrum that is always symmetric.

Next, we concentrate on $\mathcal{B}$ from (A.11). Using similar steps and the positive semidefiniteness of the matrix $\boldsymbol{\Sigma}(u)$ that ascertains that there exists $\boldsymbol{P}(u)$ such that $\boldsymbol{\Sigma}(u) = \boldsymbol{P}(u)\boldsymbol{P}^{\top}(u)$.

$$\begin{aligned}
\sum_{h=0}^{\infty}\Big[\boldsymbol{\Psi}(u,h)\boldsymbol{\Sigma}(u)\boldsymbol{\Psi}^{\top}(u,h)\Big] &= \sum_{h=0}^{\infty}\Big[\boldsymbol{\Psi}(u,h)\boldsymbol{P}(u)\Big]\Big[\boldsymbol{\Psi}(u,h)\boldsymbol{P}(u)\Big]^{\top} \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{r=0}^{\infty}\sum_{v=0}^{\infty}\Big[\boldsymbol{\Psi}(u,r,e^{i\omega r})\boldsymbol{P}(u)\Big]\Big[\boldsymbol{\Psi}(u,v,e^{-i\omega v})\boldsymbol{P}(u)\Big]^{\top}d\omega \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{r=0}^{\infty}\Big[\boldsymbol{\Psi}(u,r,e^{i\omega r})\boldsymbol{P}(u)\Big]\sum_{v=0}^{\infty}\Big[\boldsymbol{\Psi}(u,v,e^{-i\omega v})\boldsymbol{P}(u)\Big]^{\top}d\omega \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\Big[\boldsymbol{\Psi}(u,e^{i\omega})\boldsymbol{P}(u)\Big]\Big[\boldsymbol{\Psi}(u,e^{-i\omega})\boldsymbol{P}(u)\Big]^{\top}d\omega \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\Big[\Big\{\boldsymbol{\Psi}(u,e^{i\omega})\Big\}\boldsymbol{\Sigma}(u)\Big\{\boldsymbol{\Psi}(u,e^{-i\omega})\Big\}^{\top}\Big]d\omega
\end{aligned} \tag{A.15}$$

That establishes the fact that

$$\mathcal{B} = \sum_{h=0}^{\infty}\Big[\boldsymbol{\Psi}(u,h)\boldsymbol{\Sigma}(u)\boldsymbol{\Psi}^{\top}(u,h)\Big]_{j,j} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\Big[\Big\{\boldsymbol{\Psi}(u,e^{i\omega})\Big\}\boldsymbol{\Sigma}(u)\Big\{\boldsymbol{\Psi}(u,e^{-i\omega})\Big\}^{\top}\Big]_{j,j}d\omega \tag{A.16}$$

from (A.11), and we have shown that

$$\Big[\boldsymbol{\theta}(u)\Big]_{j,k} = \frac{\sigma_{kk}^{-1}\sum_{h=0}^{\infty}\Big(\Big[\boldsymbol{\Psi}(u,h)\boldsymbol{\Sigma}(u)\Big]_{j,k}\Big)^2}{\sum_{h=0}^{\infty}\Big[\boldsymbol{\Psi}(u,h)\boldsymbol{\Sigma}(u)\boldsymbol{\Psi}^{\top}(u,h)\Big]_{j,j}} = \frac{\sigma_{kk}^{-1}\int_{-\pi}^{\pi}\Big|\Big[\boldsymbol{\Psi}(u,e^{-i\omega})\boldsymbol{\Sigma}(u)\Big]_{j,k}\Big|^2 d\omega}{\int_{-\pi}^{\pi}\Big[\Big\{\boldsymbol{\Psi}(u,e^{i\omega})\Big\}\boldsymbol{\Sigma}(u)\Big\{\boldsymbol{\Psi}(u,e^{-i\omega})\Big\}^{\top}\Big]_{j,j}d\omega} \tag{A.17}$$

Finally, focusing on a frequency band $d = (a,b) : a,b \in (-\pi,\pi), a < b$, we have

$$\Big[\boldsymbol{\theta}(u,d)\Big]_{j,k} = \frac{\sigma_{kk}^{-1}\int_{a}^{b}\Big|\Big[\boldsymbol{\Psi}(u,e^{-i\omega})\boldsymbol{\Sigma}(u)\Big]_{j,k}\Big|^2 d\omega}{\int_{-\pi}^{\pi}\Big[\Big\{\boldsymbol{\Psi}(u,e^{i\omega})\Big\}\boldsymbol{\Sigma}(u)\Big\{\boldsymbol{\Psi}(u,e^{-i\omega})\Big\}^{\top}\Big]_{j,j}d\omega} \tag{A.18}$$

This completes the proof. □

Proposition 2. Using the Remark 1 and appropriate substitutions, it immediately follows

that

$$\begin{aligned}\sum_{d_s \in D} \mathcal{C}(u, d_s) &= \sum_{d_z \in D} \left(\sum_{\substack{j,k=1 \\ j \neq k}}^{N} \left[\widetilde{\boldsymbol{\theta}}(u, d_s)\right]_{j,k} \Big/ \sum_{j,k=1}^{N} \left[\widetilde{\boldsymbol{\theta}}(u, \infty)\right]_{j,k} \right) \\ &= \left(\sum_{d_z \in D} \sum_{\substack{j,k=1 \\ j \neq k}}^{N} \left[\widetilde{\boldsymbol{\theta}}(u, d_s)\right]_{j,k} \right) \Big/ \sum_{j,k=1}^{N} \left[\widetilde{\boldsymbol{\theta}}(u, \infty)\right]_{j,k} \\ &= \sum_{\substack{j,k=1 \\ j \neq k}}^{N} \left[\widetilde{\boldsymbol{\theta}}(u, \infty)\right]_{j,k} \Big/ \sum_{j,k=1}^{N} \left[\widetilde{\boldsymbol{\theta}}(u, \infty)\right]_{j,k} \\ &= \mathcal{C}(u)\end{aligned} \tag{A.19}$$

Similarly, quantities $\mathcal{C}_{j \leftarrow \bullet}(u)$ and $\mathcal{C}_{j \rightarrow \bullet}(u)$ will sum over frequency bands. This completes the proof. □

Proposition 3. Before proving proposition 3, we first outline the regularity conditions that resemble those in Liu et al. (2022). Note that $\mathcal{D}(u)$ are transformations of the parameters of the candidate model that possess a posterior distribution from which we are able to compute posterior means and variances. Petrova (2019) establishes asymptotic normality of the VAR parameters using the QBLL approach under a linear Gaussian local likelihood function we use in this paper; we refer the interested reader to Petrova (2019)'s for details.[26]

Assumptions

1. $\mathcal{D}(u) \in \boldsymbol{\Theta}$, where $\boldsymbol{\Theta}$, the parameter space, is a compact subset of $\mathbb{R}^q$ where $q = q_{\mathcal{D}}$
2. $\{X_t\}_{t=1}^{\infty}$ satisfies the α-mixing condition with the coefficient $\alpha(m) = O\left(m^{\frac{-2r}{r-2}-\epsilon}\right)$ where $\epsilon > 0$, $r > 2$.
3. $\forall t$, $l_t\left(\mathcal{D}(u)\right)$ is three-times differentiable on $\boldsymbol{\Theta}$ almost surely.
4. For any $\mathcal{D}(u)$, $\mathcal{D}(u)^{\top} \in \boldsymbol{\Theta}$, $||l_t^{(j)}(\mathcal{D}(u)) - l_t^{(j)}(\mathcal{D}(u))^{\top}|| \leq c_t^j\left(\boldsymbol{X}^t\right) ||\mathcal{D}(u) - \mathcal{D}(u)^{\top}||$ in probability, where $c_t^j\left(\boldsymbol{X}^t\right) > 0$, $\sup_t \mathbf{E}||c_t^j\left(\boldsymbol{X}^t\right)|| < \infty$, and $n^{-1}\sum_{t=1}^{n}\left(c_t^j\left(\boldsymbol{X}^t\right) - \mathbf{E}\left(c_t^j\left(\boldsymbol{X}^t\right)\right)\right) \rightarrow^p 0$, $j = \{0, 1, 2\}$.
5. $\forall \mathcal{D}(u) \in \boldsymbol{\Theta}$, $\exists M_t(\boldsymbol{X}^t) > 0$, such that $l_t^{(j)}(\mathcal{D}(u))$ exists, $\sup_{\mathcal{D}(u) \in \boldsymbol{\Theta}} ||l_t^{(j)}(\mathcal{D}(u))|| \leq M_t(\boldsymbol{X}^t)$, and $\sup_t \mathbf{E}||M_t(\boldsymbol{X}^t)||^{r+\delta} \leq M$ for some $\delta > 0$, $M < \infty$, where r is the same as in Assumption 2, and $j = \{0, 1, 2\}$.
6. $\left\{l_t^{(j)}(\mathcal{D}(u))\right\}$ is L_2-near epoch dependent of size - 1 for $0 \leq j \leq 1$ and $-\frac{1}{2}$ for j=2 uniformly on $\boldsymbol{\Theta}$.

[26] Lütkepohl (1990) provides the asymptotic distribution for impulse response functions from VAR models.

7. Let $\mathcal{D}(u)_n^0$ be the minimizer of the KL-divergence between the DGP $g(\boldsymbol{X})$ and the candidate model $p(\boldsymbol{X}|\mathcal{D}(u))$, that is,

$$\mathcal{D}(u)_n^0 = \arg\min_{\mathcal{D}(u)\in\boldsymbol{\Theta}} \frac{1}{n}\int \log\frac{g(\boldsymbol{X})}{p(\boldsymbol{X}|\mathcal{D}(u)}g(\boldsymbol{X})d\boldsymbol{X},$$

where $\{\mathcal{D}_n^0\}$ is the sequence of minimizers in $\boldsymbol{\Theta}$. For any $\epsilon > 0$,

$$\lim_{n\to\infty}\sup \sup_{\mathcal{D}(u)\in\boldsymbol{\Theta}\setminus N(\mathcal{D}(u),\epsilon)} \frac{1}{n}\sum_{t=1}^{n}\left\{\mathbf{E}[l_t(\mathcal{D}(u))] - \mathbf{E}[l_t(\mathcal{D}(u)_n^0)]\right\} < 0$$

where $N(\mathcal{D}(u),\epsilon)$ is the open ball of radius ϵ around $\mathcal{D}(u)$.

8. $\{-\mathbf{H}_n(\mathcal{D}(u)_n^0), \mathbf{B}_n(\mathcal{D}(u)_n^0)\}$ are positive definite uniformly on n and $\mathbf{H}_n(\mathcal{D}(u)_n^0) + \mathbf{B}_n(\mathcal{D}(u)_n^0) = 0$.

9. The prior density $p(\mathcal{D}(u))$ is thrice continuously differentiable and $0 < p(\mathcal{D}(u)) < \infty$ uniformly on n. There exists an n^* such that for any $n > n^*$, the posterior distribution $p(\mathcal{D}(u)|\boldsymbol{X})$ is proper, and $\int||\mathcal{D}(u)||^2 p(\mathcal{D}(u)|\boldsymbol{X})d\mathcal{D}(u) < \infty$.

Remarks

1. Assumption 1 is the compactness condition. Assumptions 2 and 6 imply weak dependence in X_t and l_t. Assumption 3 are continuity conditions. Assumption 4 is the Lipshitz condition for l_t to develop the uniform of large numbers for dependent and heterogeneous stochastic processes. Assumption 5 contains the dominance condition for l_t. Assumption 7 is the identification condition in Gallant and White (1988). Such assumptions are well-known to develop Maximum Likelihood theory; namely consistency and asymptotic normality for dependent and heterogeneous data.

2. In Assumption 2, a mixing process is only a mixing process if the function depends on a finite number of lagged values in the mixing process. In most latent variable models however, the likelihood function depend on distant past or future of the process. Assumption 6 controls the dependence of the function Gallant and White (1988). The differentiability in Assumption 3 and the domination condition in Assumption 5 are important to develop a high order Laplace expansion. Together with assumption 9, the concentration condition indicates that we can apply the stochastic Laplace expansion to the posterior distribution and establish the asymptotic normality of the posterior distribution.

3. Assumption 8 gives the requirement for a good model, and assumption 9 ensures the second moment of the prior is finite which leads to a finite second moment of the posterior whilst also implying the the prior is negligible asymptotically.

We note that Li et al. (2020) further outlines the importance of these assumptions. Now, using Lemma A.1 in Liu et al. (2022), we have

$$\begin{aligned}
\mathbf{E}\left[\left(\mathcal{D}(u)-\widehat{\mathcal{D}}(u)\right)|\boldsymbol{X}\right] &= o_p\left(n^{-1/2}\right) \\
\mathbf{V}\left(\widehat{\mathcal{D}}(u)\right) &= \mathbf{E}\left[\left(\mathcal{D}(u)-\widehat{\mathcal{D}}(u)\right)\left(\mathcal{D}(u)-\widehat{\mathcal{D}}(u)\right)^{\top}|\boldsymbol{X}\right] \\
&= -\frac{1}{n}\bar{\mathbf{H}}_n^{-1}\left(\widehat{\mathcal{D}}(u)\right)+o_p\left(n^{-1}\right)=O_p(n^{-1})
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
\mathbf{V}\left(\bar{\mathcal{D}}(u)\right) &= \mathbf{E}\left[\left(\mathcal{D}(u)-\widehat{\mathcal{D}}(u)\right)\left(\mathcal{D}(u)-\widehat{\mathcal{D}}(u)\right)^{\top}|\boldsymbol{X}\right] \\
&= \mathbf{E}\left[\left(\mathcal{D}(u)-\widehat{\mathcal{D}}(u)+\widehat{\mathcal{D}}(u)-\bar{\mathcal{D}}(u)\right)\left(\mathcal{D}(u)-\widehat{\mathcal{D}}(u)+\widehat{\mathcal{D}}(u)-\bar{\mathcal{D}}(u)\right)^{\top}|\boldsymbol{X}\right] \\
&= \mathbf{V}\left(\widehat{\mathcal{D}}(u)\right)-\mathbf{E}\left[\left(\widehat{\mathcal{D}}(u)-\bar{\mathcal{D}}(u)\right)\left(\left(\widehat{\mathcal{D}}(u)-\bar{\mathcal{D}}(u)\right)\right)^{\top}|\boldsymbol{X}\right] \\
&= \mathbf{V}\left(\widehat{\mathcal{D}}(u)\right)+o_p\left(n^{-1/2}\right)o_p\left(n^{-1/2}\right) \\
&= -\frac{1}{n}\bar{\mathbf{H}}_n^{-1}\left(\widehat{\mathcal{D}}(u)\right)+o_p\left(n^{-1}\right)=O_p(n^{-1})
\end{aligned}$$

Under classical asymptotic theory for maximum likelihood estimation White (1996), $\widehat{\mathcal{D}}(u)-\mathcal{D}_0(u)=O_p(n^{-1/2})$ under the null, H_0. Therefore:

$$\begin{aligned}
&\left(\widehat{\mathcal{D}}(u)-\mathcal{D}_0(u)\right)^{\top}\left[\mathbf{V}_{\mathcal{D}\mathcal{D}}\left(\bar{\mathcal{D}}(u)\right)\right]^{-1}\left(\widehat{\mathcal{D}}(u)-\mathcal{D}_0(u)\right) \\
&=\left(\widehat{\mathcal{D}}(u)-\mathcal{D}_0(u)\right)^{\top}\left[-n^{-1}\bar{\mathbf{H}}_{n,\mathcal{D}\mathcal{D}}^{-1}\left(\widehat{\mathcal{D}}(u)\right)+o_p\left(n^{-1}\right)\right]^{-1}\left(\widehat{\mathcal{D}}(u)-\mathcal{D}_0(u)\right)
\end{aligned}$$

$$\begin{aligned}
&= \sqrt{n}\left(\widehat{\mathcal{D}}(u)-\mathcal{D}_0(u)\right)^{\top}\left[-\bar{\mathbf{H}}_{n,\mathcal{D}\mathcal{D}}^{-1}\left(\widehat{\mathcal{D}}(u)\right)+o_p(1)\right]^{-1}\sqrt{n}\left(\widehat{\mathcal{D}}(u)-\mathcal{D}_0(u)\right) \\
&= \sqrt{n}\left(\widehat{\mathcal{D}}(u)-\mathcal{D}_0(u)\right)^{\top}\left[-\bar{\mathbf{H}}_{n,\mathcal{D}\mathcal{D}}^{-1}\left(\widehat{\mathcal{D}}(u)\right)\right]^{-1}\sqrt{n}\left(\widehat{\mathcal{D}}(u)-\mathcal{D}_0(u)\right) \\
&+ o_p(1)\sqrt{n}\left(\widehat{\mathcal{D}}(u)-\mathcal{D}_0(u)\right)^{\top}\sqrt{n}\left(\widehat{\mathcal{D}}(u)-\mathcal{D}_0(u)\right) \\
&= \sqrt{n}\left(\widehat{\mathcal{D}}(u)-\mathcal{D}_0(u)\right)^{\top}\left[-\bar{\mathbf{H}}_{n,\mathcal{D}\mathcal{D}}^{-1}\left(\widehat{\mathcal{D}}(u)\right)\right]^{-1}\sqrt{n}\left(\widehat{\mathcal{D}}(u)-\mathcal{D}_0(u)\right) \\
&+ o_p(1)\sqrt{n}O_p\left(n^{-1/2}\right)\sqrt{n}O_p\left((n^{-1/2}\right) \\
&= \sqrt{n}\left(\widehat{\mathcal{D}}(u)-\mathcal{D}_0(u)\right)^{\top}\left[-\bar{\mathbf{H}}_{n,\mathcal{D}\mathcal{D}}^{-1}\left(\widehat{\mathcal{D}}(u)\right)\right]^{-1}+o_p(1)
\end{aligned}$$

$$= \mathbf{Wald}+o_p(1) \tag{A.20}$$

Under the null, H_0, we can show that

$$\begin{aligned}
&\mathbf{W}\left(\boldsymbol{X}, \mathcal{D}_0(u)\right) = \\
&q_{\mathcal{D}} + \left(\bar{\mathcal{D}}(u) - \widehat{\mathcal{D}}(u) + \widehat{\mathcal{D}}(u) - \mathcal{D}_0(u)\right)^{\top} \left[\mathbf{V}_{\mathcal{D}\mathcal{D}}\left(\bar{\mathcal{D}}(u)\right)\right]^{-1} \left(\bar{\mathcal{D}}(u) - \widehat{\mathcal{D}}(u) + \widehat{\mathcal{D}}(u) - \mathcal{D}_0(u)\right)
\end{aligned}$$

$$\begin{aligned}
&= q_{\mathcal{D}} + \left(\widehat{\mathcal{D}}(u) - \mathcal{D}_0(u)\right)^{\top} \left[\mathbf{V}_{\mathcal{D}\mathcal{D}}\left(\bar{\mathcal{D}}(u)\right)\right]^{-1} \left(\widehat{\mathcal{D}}(u) - \mathcal{D}_0(u)\right) \\
&+ 2\left(\bar{\mathcal{D}}(u) - \widehat{\mathcal{D}}(u)\right)^{\top} \left[\mathbf{V}_{\mathcal{D}\mathcal{D}}\left(\bar{\mathcal{D}}(u)\right)\right]^{-1} \left(\widehat{\mathcal{D}}(u) - \mathcal{D}_0(u)\right) \\
&+ \left(\bar{\mathcal{D}}(u) - \widehat{\mathcal{D}}(u)\right)^{\top} \left[\mathbf{V}_{\mathcal{D}\mathcal{D}}\left(\bar{\mathcal{D}}(u)\right)\right]^{-1} \left(\bar{\mathcal{D}}(u) - \widehat{\mathcal{D}}(u)\right) \\
&= q_{\mathcal{D}} + \left(\widehat{\mathcal{D}}(u) - \mathcal{D}_0(u)\right)^{\top} \left[\mathbf{V}_{\mathcal{D}\mathcal{D}}\left(\bar{\mathcal{D}}(u)\right)\right]^{-1} \left(\widehat{\mathcal{D}}(u) - \mathcal{D}_0(u)\right) \\
&+ 2o_p\left(n^{-1/2}\right) O_p(n) O_p\left(n^{-1/2}\right) + o_p\left(n^{-1/2}\right) O_p(n) o_p\left(n^{-1/2}\right)
\end{aligned}$$

$$= q_{\mathcal{D}} + \left(\widehat{\mathcal{D}}(u) - \mathcal{D}_0(u)\right)^{\top} \left[\mathbf{V}_{\mathcal{D}\mathcal{D}}\left(\bar{\mathcal{D}}(u)\right)\right]^{-1} \left(\widehat{\mathcal{D}}(u) - \mathcal{D}_0(u)\right) + o_p(1) \tag{A.21}$$

Which if assumptions 1–9 hold, from equation (A) and (A.21) under H_0 we have

$$\mathbf{W}\left(\boldsymbol{X}, \mathcal{D}_0(u)\right) = \mathbf{Wald} + q_{\mathcal{D}} + o_p(1) \rightarrow^{d} \chi^2\left(q_{\mathcal{D}}\right) + q_{\mathcal{D}}$$

□

B Estimation Details

To estimate our high dimensional systems, we follow the Quasi-Bayesian Local-Liklihood (QBLL) approach of Petrova (2019). let $\boldsymbol{X}_t$ be an $N \times 1$ vector generated by a stable time-varying parameter (TVP) heteroskedastic VAR model with p lags:

$$\boldsymbol{X}_{t,T} = \boldsymbol{\Phi}_1(t/T)\boldsymbol{X}_{t-1,T} + \ldots + \boldsymbol{\Phi}_p(t/T)\boldsymbol{X}_{t-p,T} + \boldsymbol{\epsilon}_{t,T}, \tag{B.1}$$

where $\boldsymbol{\epsilon}_{t,T} = \boldsymbol{\Sigma}^{-1/2}(t/T)\boldsymbol{\eta}_{t,T}$ with $\boldsymbol{\eta}_{t,T} \sim NID(0, \boldsymbol{I}_M)$ and $\boldsymbol{\Phi}(t/T) = (\boldsymbol{\Phi}_1(t/T), \ldots, \boldsymbol{\Phi}_p(t/T))^{\top}$ are the time varying autoregressive coefficients. Note that all roots of the locally stationary VAR polynomial, lie outside the unit circle, and $\boldsymbol{\Sigma}_t^{-1}$ is a positive definite time-varying covariance matrix. Stacking the time-varying intercepts and autoregressive matrices in the vector $\phi(t/T)$ with $\bar{\boldsymbol{X}}'_{t,T} = (\mathbf{I}_N \otimes x_{t,T})$, $x_{t,T} = \left(1, x'_{t-1,T}, \ldots, x'_{t-p,T}\right)$ and $\otimes$ denotes the Kronecker product, the model can be written as:

$$\boldsymbol{X}_{t,T} = \bar{\boldsymbol{X}}'_{t,T}\phi(t/T) + \boldsymbol{\Sigma}^{-\frac{1}{2}}(t/T)\boldsymbol{\eta}_{t,T} \tag{B.2}$$

We obtain the time-varying parameters of the model by employing Quasi-Bayesian Local Likelihood (QBLL) methods. Estimation of (B.1) requires re-weighting the likelihood function. Essentially, the weighting function gives higher proportions to observations surrounding the time period whose parameter values are of interest. The local likelihood

function at discrete time period s, where we drop the double time index for notational convenience, is given by:

$$\mathrm{L}_s\left(\boldsymbol{X}_s|\phi_s,\boldsymbol{\Sigma}_s,\bar{\boldsymbol{X}}_s\right)\propto$$
$$|\boldsymbol{\Sigma}_s|^{\mathrm{trace}(\mathbf{D}_s)/2}\exp\left\{-\frac{1}{2}(\boldsymbol{X}_s-\bar{\boldsymbol{X}}_s'\phi_s)'\left(\boldsymbol{\Sigma}_s\otimes\mathbf{D}_s\right)(\boldsymbol{X}_s-\bar{\boldsymbol{X}}_s'\phi_s)\right\} \quad \text{(B.3)}$$

The $\mathbf{D}_s$ is a diagonal matrix whose elements hold the weights:

$$\begin{aligned}
\mathbf{D}_s &= \mathrm{diag}(\varrho_{s1},\ldots,\varrho_{sT}) && \text{(B.4)}\\
\varrho_{st} &= \zeta_{T,s}w_{st}/\sum_{t=1}^{T}w_{st} && \text{(B.5)}\\
w_{st} &= (1/\sqrt{2\pi})\exp((-1/2)((k-t)/W)^2), \quad \text{for } s,t\in\{1,\ldots,T\} && \text{(B.6)}\\
\zeta_{Ts} &= \left(\left(\sum_{t=1}^{T}w_{st}\right)^2\right)^{-1} && \text{(B.7)}
\end{aligned}$$

where ϱ_{st} is a normalised kernel function. w_{st} uses a Normal kernel weighting function. ζ_{Ts} gives the rate of convergence and behaves like the bandwidth parameter W in (B.6). it is the kernel function that provides greater weight to observations surrounding the parameter estimates at time s relative to more distant observations.

Using a Normal-Wishart prior distribution for $\phi_s|\,\boldsymbol{\Sigma}_s$ for $s\in\{1,\ldots,T\}$:

$$\begin{aligned}
\phi_s|\boldsymbol{\Sigma}_s &\backsim \mathcal{N}\left(\phi_{0s},(\boldsymbol{\Sigma}_s\otimes\boldsymbol{\Xi}_{0s})^{-1}\right) && \text{(B.8)}\\
\boldsymbol{\Sigma}_s &\backsim \mathcal{W}\left(\alpha_{0s},\boldsymbol{\Gamma}_{0s}\right) && \text{(B.9)}
\end{aligned}$$

where ϕ_{0s} is a vector of prior means, $\boldsymbol{\Xi}_{0s}$ is a positive definite matrix, α_{0s} is a scale parameter of the Wishart distribution ($\mathcal{W}$), and $\boldsymbol{\Gamma}_{0s}$ is a positive definite matrix.

The prior and weighted likelihood function implies a Normal-Wishart quasi posterior distribution for $\phi_s|\,\boldsymbol{\Sigma}_s$ for $s=\{1,\ldots,T\}$. Formally let $\mathbf{A}=(\bar{x}_1',\ldots,\bar{x}_T')'$ and $\mathbf{Y}=(x_1,\ldots,x_T)'$ then:

$$\begin{aligned}
\phi_s|\boldsymbol{\Sigma}_s,\mathbf{A},\mathbf{Y} &\backsim \mathcal{N}\left(\tilde{\phi}_s,\left(\boldsymbol{\Sigma}_s\otimes\tilde{\boldsymbol{\Xi}}_s\right)^{-1}\right) && \text{(B.10)}\\
\boldsymbol{\Sigma}_s &\backsim \mathcal{W}\left(\tilde{\alpha}_s,\tilde{\boldsymbol{\Gamma}}_s^{-1}\right) && \text{(B.11)}
\end{aligned}$$

with quasi posterior parameters

$$\begin{aligned}
\tilde{\phi}_s &= \left(\mathbf{I}_N\otimes\tilde{\boldsymbol{\Xi}}_s^{-1}\right)\left[\left(\mathbf{I}_N\otimes\mathbf{A}'\mathbf{D}_s\mathbf{A}\right)\hat{\phi}_s+\left(\mathbf{I}_N\otimes\boldsymbol{\Xi}_{0s}\right)\phi_{0s}\right] && \text{(B.12)}\\
\tilde{\boldsymbol{\Xi}}_s &= \tilde{\boldsymbol{\Xi}}_{0s}+\mathbf{A}'\mathbf{D}_s\mathbf{A} && \text{(B.13)}\\
\tilde{\alpha}_s &= \alpha_{0s}+\sum_{t=1}^{T}\varrho_{st} && \text{(B.14)}\\
\tilde{\boldsymbol{\Gamma}}_s &= \boldsymbol{\Gamma}_{0s}+\mathbf{Y}'\mathbf{D}_s\mathbf{Y}+\boldsymbol{\Phi}_{0s}\boldsymbol{\Gamma}_{0s}\boldsymbol{\Phi}_{0s}'-\tilde{\boldsymbol{\Phi}}_s\tilde{\boldsymbol{\Gamma}}_s\tilde{\boldsymbol{\Phi}}_s' && \text{(B.15)}
\end{aligned}$$

where $\hat{\phi}_s = (\mathbf{I}_N \otimes \mathbf{A}'\mathbf{D}_s\mathbf{A})^{-1}(\mathbf{I}_N \otimes \mathbf{A}'\mathbf{D}_s)\, y$ is the local likelihood estimator for ϕ_s. The matrices $\mathbf{\Phi}_{0s}$, $\tilde{\mathbf{\Phi}}_s$ are conformable matrices from the vector of prior means, ϕ_{0s}, and a draw from the quasi posterior distribution, $\tilde{\phi}_s$, respectively.

The motivation for employing these methods are threefold. First, we are able to estimate large systems that conventional Bayesian estimation methods do not permit. This is typically because the state-space representation of an N-dimensional TVP VAR (p) requires an additional $N(3/2 + N(p + 1/2))$ state equations for every additional variable. Conventional Markov Chain Monte Carlo (MCMC) methods fail to estimate larger models, which in general confine one to (usually) fewer than 6 variables in the system. Second, the standard approach is fully parametric and requires a law of motion. This can distort inference if the true law of motion is misspecified. Third, the methods used here permit direct estimation of the VAR's time-varying covariance matrix, which has an inverse-Wishart density and is symmetric positive definite at every point in time.

In estimating the model, we use p=2 and a Minnesota Normal-Wishart prior with a shrinkage value $\varphi = 0.05$ and centre the coefficient on the first lag of each variable to 0.1 in each respective equation. The prior for the Wishart parameters are set following Kadiyala and Karlsson (1997). For each point in time, we run 500 simulations of the model to generate the (quasi) posterior distribution of parameter estimates. Note we experiment with various lag lengths, $p = \{2, 3, 4, 5\}$; shrinkage values, $\varphi = \{0.01, 0.25, 0.5\}$; and values to centre the coefficient on the first lag of each variable, $\{0, 0.05, 0.2, 0.5\}$. Network measures from these experiments are qualitatively similar. Notably, adding lags to the VAR and increasing the persistence in the prior value of the first lagged dependent variable in each equation increases computation time.

C Monte Carlo Study: Additional Results

Figures C1 and C2 report dynamic network connectedness measures under kernel bandwidths W=$\{12, 18\}$. As we can see, the larger is W, the smoother the measures become.

C.1 DGPI with fat tails

Here we report the connectedness measures for DGPI in the main text with residuals that follow a multivariate student-t distribution. In Figure C3 we report the true and estimated connectedness measures for this alternative DGP at bandwidths W=$\{8, 12, 18\}$. Overall, these plots correspond to the top row in Figure 2 within the main text.

C.2 Posterior probabilities for differences in network connectedness measures

Here we examine whether the network connectedness measures one computes over frequency bands are different from one another. We use the time u joint posterior distribution of horizon specific network connectedness to compute the probability of heterogeneities. Consider the time u network connectedness measure on frequency band d, $\mathcal{C}(u,d)$, and the time u network connectedness measure on frequency band c, $\mathcal{C}(u,c)$ where $d < c$. Then, the probability that the time u network connectedness on frequency band d is greater than the time u network connectedness measure on frequency band c, $\Pr\left(\mathcal{C}(u,d) > \mathcal{C}(u,c)\right)$, is given by

$$\Pr\left(\mathcal{C}(u,d) > \mathcal{C}(u,c)\right) = \sum \left(\mathcal{C}^r(u,d) > \mathcal{C}^r(u,c)\right) / \mathrm{R}$$

where $\mathcal{C}^r(u,d)$, $\mathcal{C}^r(u,c)$ denote the rth draw from the posterior distribution of each network connectedness measure, and R denotes the total number of draws taken from the posterior distribution.

If $\Pr\left(\mathcal{C}(u,d) > \mathcal{C}(u,c)\right)$ >0.95 this tells us with 95% confidence that the network connectedness measure across frequency band d is greater than the corresponding measure across frequency band c. Similarly, if $\Pr\left(\mathcal{C}(u,d) > \mathcal{C}(u,c)\right)$ <0.05 this tells us with 95% confidence that the network connectedness measure across frequency band c is greater than the corresponding measure across frequency band d. Note here that $\Pr\left(\mathcal{C}(u,d) > \mathcal{C}(u,c)\right) = 1 - \Pr\left(\mathcal{C}(u,c) > \mathcal{C}(u,d)\right)$. These probabilities provide statistical evidence of whether heterogeneities exit between network connectedness measures over different frequency bands.

We report the probabilities for differences between network connectedness in Table C1 at three specific dates, u=$\{400,\ 650,\ 1000\}$. First, we report the true probability that the low-frequency-band network connectedness measure is greater than the high-frequency-band network connectedness measure, $\Pr\left(\mathcal{C}(u,d) > \mathcal{C}(u,c)\right)$. Then, we provide the estimates of sample average probabilities from TVP VAR models with kernel bandwidths W=$\{8,\ 12,\ 18\}$. Again, this indicates that our method works well at specific

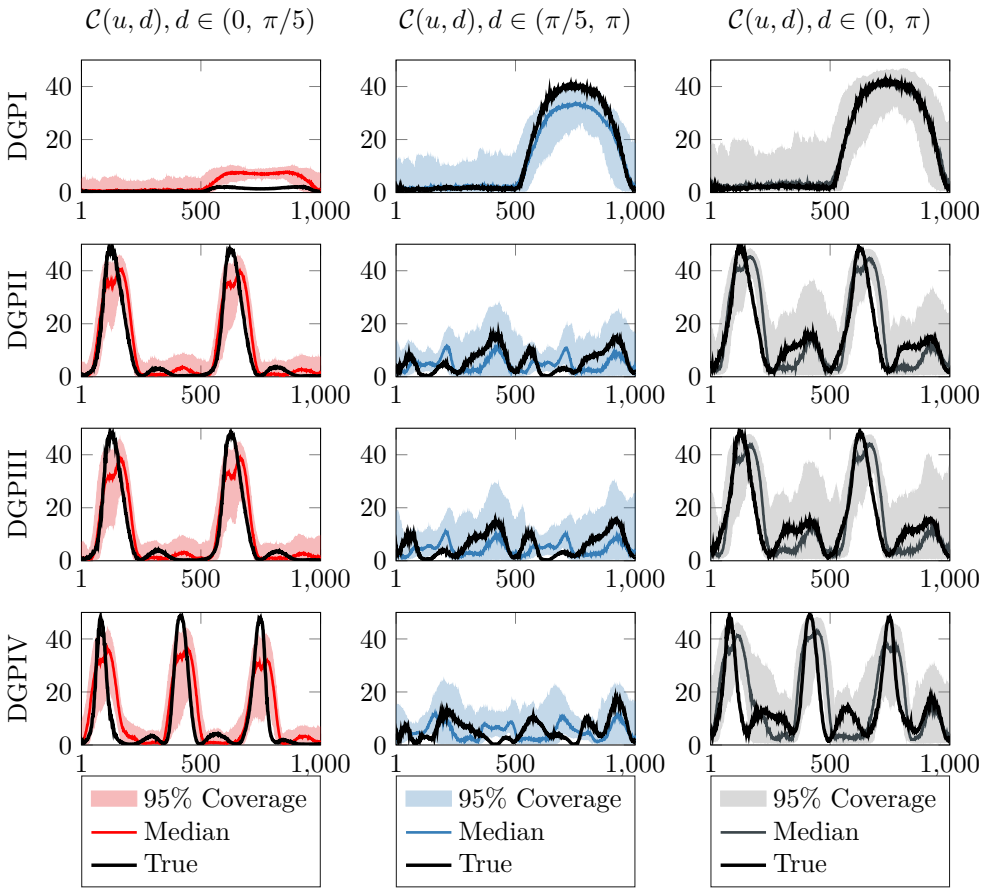

Figure C1: Dynamic network connectedness measures: True and fitted values
Notes: This figure plots the true network connectedness measures for three data generating processes following bi-variate TVP VAR(2) models along with the median and 95% quantiles of estimated network connectedness measures with bandwidth W=12. The left columns report network connectedness on the low-frequency band, $d \in (0,\ \pi/5)$, the middle columns show network connectedness on the high-frequency band, $d \in (\pi/5,\ \pi)$, and the right columns show the aggregate network connectedness such that $d \in (0,\ \pi)$. DGPI (top row) is a TVP VAR(2) model with Gaussian errors, we introduce a break in the time-varying autoregressive matrices and contemporaneous relations from observation 500 that induces large connections across the high frequency band. DGPII (second row) is a TVP VAR(2) where time-varying intercepts and autoregressive matrices following sin wave functions with a stochastic error, time-varying covariance matrix where the off-diagonals follow sin wave functions with a stochastic error, and the diagonal elements follow a stationary AR(1) processes. DGPIII (third row) is a TVP VAR(2) model with student-t errors, time-varying intercepts and autoregressive matrices following sin wave functions with a stochastic error, time-varying covariance matrix where the off-diagonals follow sin wave functions with a stochastic error, and the diagonal elements follow a stationary AR(1) processes. DGPIV (bottom row) is the same as DGPII, but with an increase in the periodicity of the respective sin wave functions the time-varying intercepts and autoregressive matrices follow.

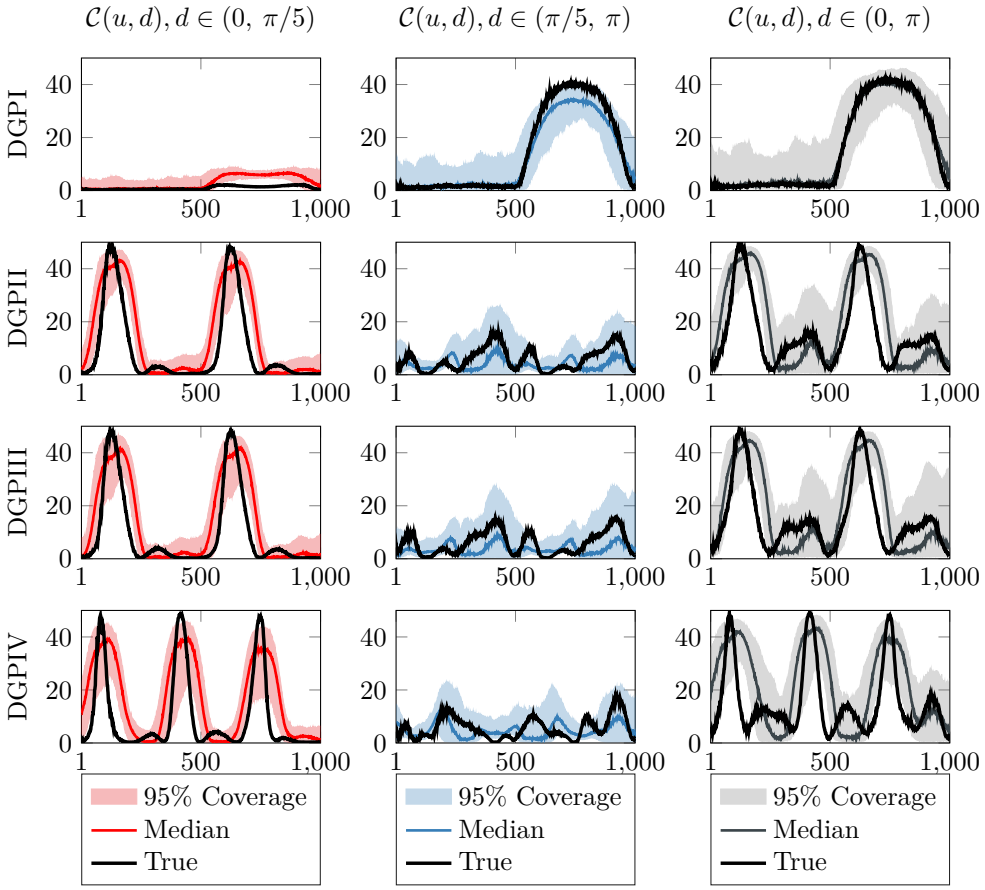

Figure C2: Dynamic network connectedness measures: True and fitted values
Notes: This figure plots the true network connectedness measures for three data generating processes following bi-variate TVP VAR(2) models along with the median and 95% quantiles of estimated network connectedness measures with bandwidth W=18. The left columns report network connectedness on the low-frequency band, $d \in (0,\ \pi/5)$, the middle columns show network connectedness on the high-frequency band, $d \in (\pi/5,\ \pi)$, and the right columns show the aggregate network connectedness such that $d \in (0,\ \pi)$. DGPI (top row) is a TVP VAR(2) model with Gaussian errors, we introduce a break in the time-varying autoregressive matrices and contemporaneous relations from observation 500 that induces large connections across the high frequency band. DGPII (second row) is a TVP VAR(2) where time-varying intercepts and autoregressive matrices following sin wave functions with a stochastic error, time-varying covariance matrix where the off-diagonals follow sin wave functions with a stochastic error, and the diagonal elements follow a stationary AR(1) processes. DGPIII (third row) is a TVP VAR(2) model with student-t errors, time-varying intercepts and autoregressive matrices following sin wave functions with a stochastic error, time-varying covariance matrix where the off-diagonals follow sin wave functions with a stochastic error, and the diagonal elements follow a stationary AR(1) processes. DGPIV (bottom row) is the same as DGPII, but with an increase in the periodicity of the respective sin wave functions the time-varying intercepts and autoregressive matrices follow.

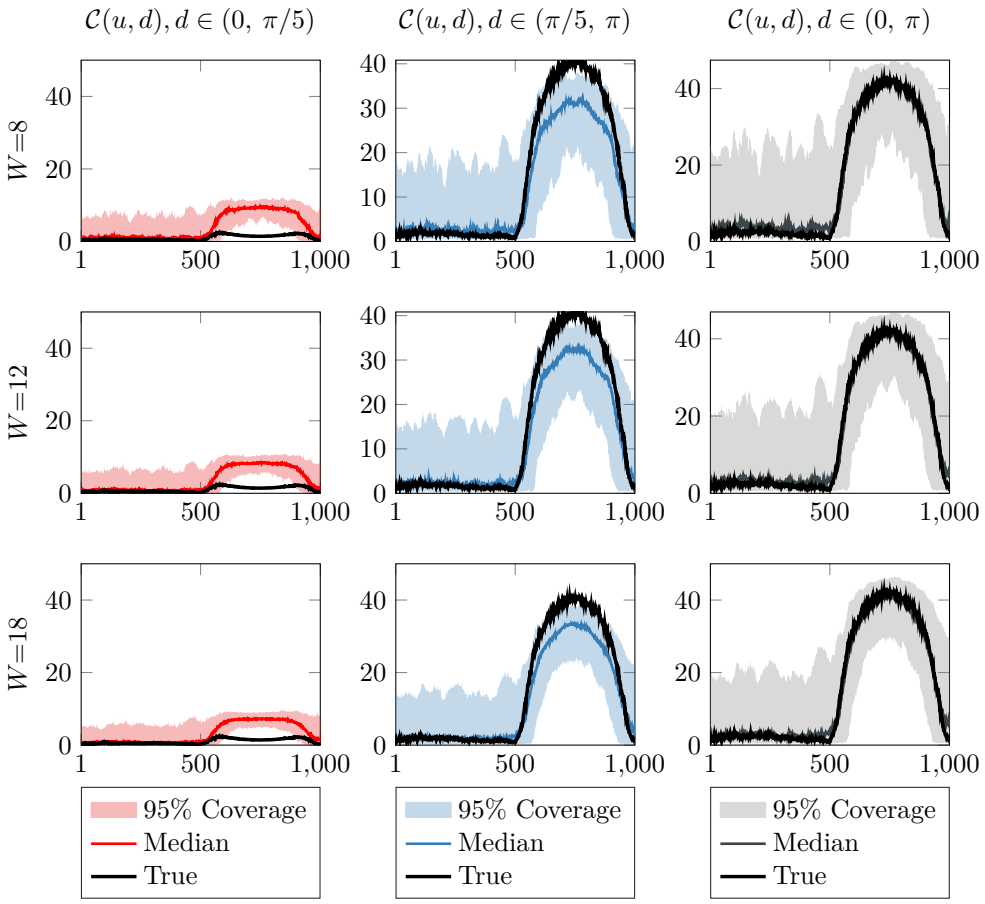

Figure C3: Dynamic network connectedness measures: True and fitted values DGPI with fat tails

Notes: This figure plots the true network connectedness measures for DGPI with residuals following a multivariate student-t distribution. Along with this the figures report the median and 95% quantiles of estimated network connectedness measures with bandwidth W=8, W=12, and W=18 in the top, middle and bottom rows respectively. The left columns report network connectedness on the low-frequency band, $d \in (0,\ \pi/5)$, the middle columns show network connectedness on the high-frequency band, $d \in (\pi/5,\ \pi)$, and the right columns show the aggregate network connectedness such that $d \in (0,\ \pi)$.

observations throughout each DGP. It is noteworthy to mention here that for bandwidth W={18} the time profiles relative to true probabilities become less accurate as can be seen from DGPIII in Panel B.

Table C1: Probability of statistical differences between low-frequency-band and high-frequency-band network connectedness: Sample average.

Notes: This table reports the probability that the low-frequency-band network connectedness measure is greater than the high-frequency-band network connectedness measure, $\Pr\left(\mathcal{C}(u,d) > \mathcal{C}(u,c)\right)$. Panel A shows results for observation u=400, and Panels B and C show results for observation u=650, and u=1000 respectively. We report the true probabilities for the three DGPs in our Monte Carlo study, followed by the fitted probabilities from TVP VARs with different kernel bandwidths, $W = \{8, 12, 18\}$. Probabilities greater (less) than 0.95 (0.05) indicate statistical differences at each time period. DGPI (left column) is a TVP VAR(2) model with Gaussian errors, time-varying intercepts and autoregressive matrices following sin wave functions with a stochastic error, time-varying covariance matrix where the off-diagonals follow sin wave functions with a stochastic error, and the diagonal elements follow a stationary AR(1) processes. DGPII (middle column) is a TVP VAR(2) model with student-t errors, time-varying intercepts and autoregressive matrices following sin wave functions with a stochastic error, time-varying covariance matrix where the off-diagonals follow sin wave functions with a stochastic error, and the diagonal elements follow a stationary AR(1) processes. DGPIII (right column) is the same as DGPI, but with an increase in the periodicity of the respective sin wave functions the time-varying intercepts and autoregressive matrices follow.

A: u=400	**DGPII**	**DGPIII**	**DGPIV**
True Probability	0.00	0.00	0.99
Fitted Probability, W=8	0.00	0.00	0.97
Fitted Probability, W=12	0.00	0.00	1.00
Fitted Probability, W=18	0.00	0.00	1.00
B: u=650	**DGPII**	**DGPIII**	**DGPIV**
True Probability	1.00	1.00	0.00
Fitted Probability, W=8	1.00	0.90	0.00
Fitted Probability, W=12	1.00	0.98	0.00
Fitted Probability, W=18	1.00	1.00	0.21
C: u=1000	**DGPII**	**DGPIII**	**DGPIV**
True Probability	0.00	0.00	0.00
Fitted Probability, W=8	0.00	0.00	0.00
Fitted Probability, W=12	0.00	0.00	0.00
Fitted Probability, W=18	0.00	0.00	0.00

C.3 An Insight into how Rolling Window VAR models track connectedness from our DGPs

Here, we examine the performance of Rolling VAR models in tracking the true connectedness measures from the DGPs we outline in Section 3. We consider window sizes of 60, 120, and 240 periods. For each of the 100 DGPs we generate, we estimate connectedness from a VAR model that uses a fixed window that rolls through the sample. Therefore, we have 940 estimates when the window size is 60, 880 when the window size is 120, and 760 when the window size is 240.

Figures C4, C5, and C6 show the results for window sizes of 60, 120 and 240 respectively. First, looking at Figure C4 we can see that the rolling VAR tracks connectedness relatively well. In all DGPs, the rolling VAR estimation picks up the surges in true connectedness with reasonable accuracy. However, when true connectedness falls, the rolling VAR responds with a lag as one would expect. When we examine Figures C5 and C6 the issue becomes more prominent. The rolling VAR still does reasonably well for DGPI, a relatively simple DGP, however, it fails to pick up when true connectedness falls as we increase the window size.

These results show that as the dynamics among variables increases with complexity, the rolling VAR performs less well in tracking true connectedness relative to our approach. In particular, our approach picks up the surges and falls in connectedness with a higher degree of accuracy.

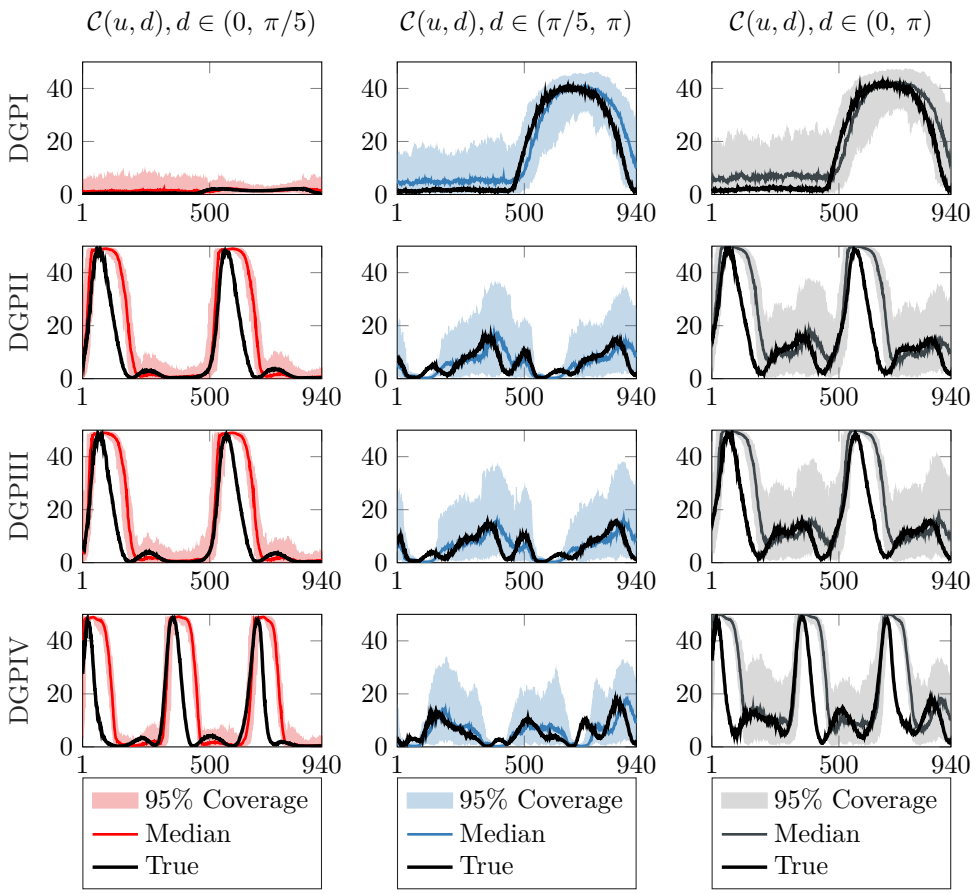

Figure C4: Dynamic network connectedness measures: True and fitted values
Notes: This figure plots the true network connectedness measures for three data generating processes following bi-variate TVP VAR(2) models along with the median and 95% quantiles of estimated network connectedness measures from a rolling VAR model we estimate using OLS and a 60 period window rolling through the sample. The left columns report network connectedness on the low-frequency band, $d \in (0,\ \pi/5)$, the middle columns show network connectedness on the high-frequency band, $d \in (\pi/5,\ \pi)$, and the right columns show the aggregate network connectedness such that $d \in (0,\ \pi)$. DGPI (top row) is a TVP VAR(2) model with Gaussian errors, we introduce a break in the time-varying autoregressive matrices and contemporaneous relations from observation 500 that induces large connections across the high frequency band. DGPII (second row) is a TVP VAR(2) where time-varying intercepts and autoregressive matrices following sin wave functions with a stochastic error, time-varying covariance matrix where the off-diagonals follow sin wave functions with a stochastic error, and the diagonal elements follow a stationary AR(1) processes. DGPIII (third row) is a TVP VAR(2) model with student-t errors, time-varying intercepts and autoregressive matrices following sin wave functions with a stochastic error, time-varying covariance matrix where the off-diagonals follow sin wave functions with a stochastic error, and the diagonal elements follow a stationary AR(1) processes. DGPIV (bottom row) is the same as DGPII, but with an increase in the periodicity of the respective sin wave functions the time-varying intercepts and autoregressive matrices follow.

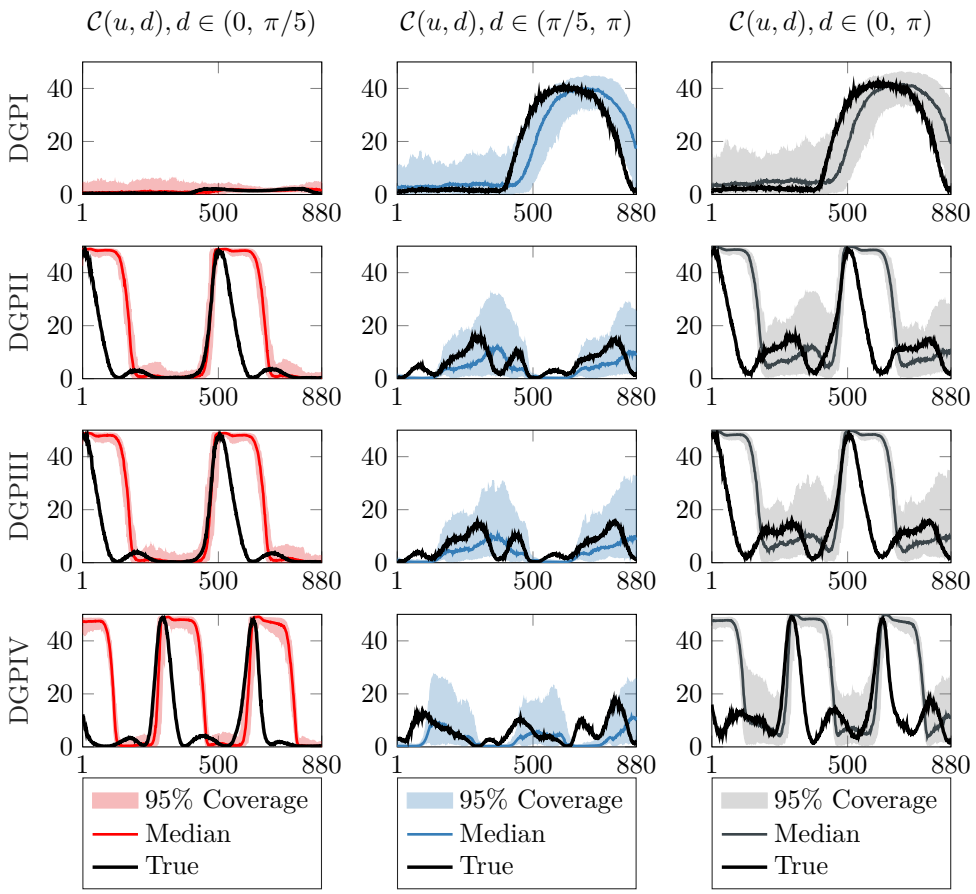

Figure C5: Dynamic network connectedness measures: True and fitted values
Notes: This figure plots the true network connectedness measures for three data generating processes following bi-variate TVP VAR(2) models along with the median and 95% quantiles of estimated network connectedness measures from a rolling VAR model we estimate using OLS and a 120 period window rolling through the sample. The left columns report network connectedness on the low-frequency band, $d \in (0,\ \pi/5)$, the middle columns show network connectedness on the high-frequency band, $d \in (\pi/5,\ \pi)$, and the right columns show the aggregate network connectedness such that $d \in (0,\ \pi)$. DGPI (top row) is a TVP VAR(2) model with Gaussian errors, we introduce a break in the time-varying autoregressive matrices and contemporaneous relations from observation 500 that induces large connections across the high frequency band. DGPII (second row) is a TVP VAR(2) where time-varying intercepts and autoregressive matrices following sin wave functions with a stochastic error, time-varying covariance matrix where the off-diagonals follow sin wave functions with a stochastic error, and the diagonal elements follow a stationary AR(1) processes. DGPIII (third row) is a TVP VAR(2) model with student-t errors, time-varying intercepts and autoregressive matrices following sin wave functions with a stochastic error, time-varying covariance matrix where the off-diagonals follow sin wave functions with a stochastic error, and the diagonal elements follow a stationary AR(1) processes. DGPIV (bottom row) is the same as DGPII, but with an increase in the periodicity of the respective sin wave functions the time-varying intercepts and autoregressive matrices follow.

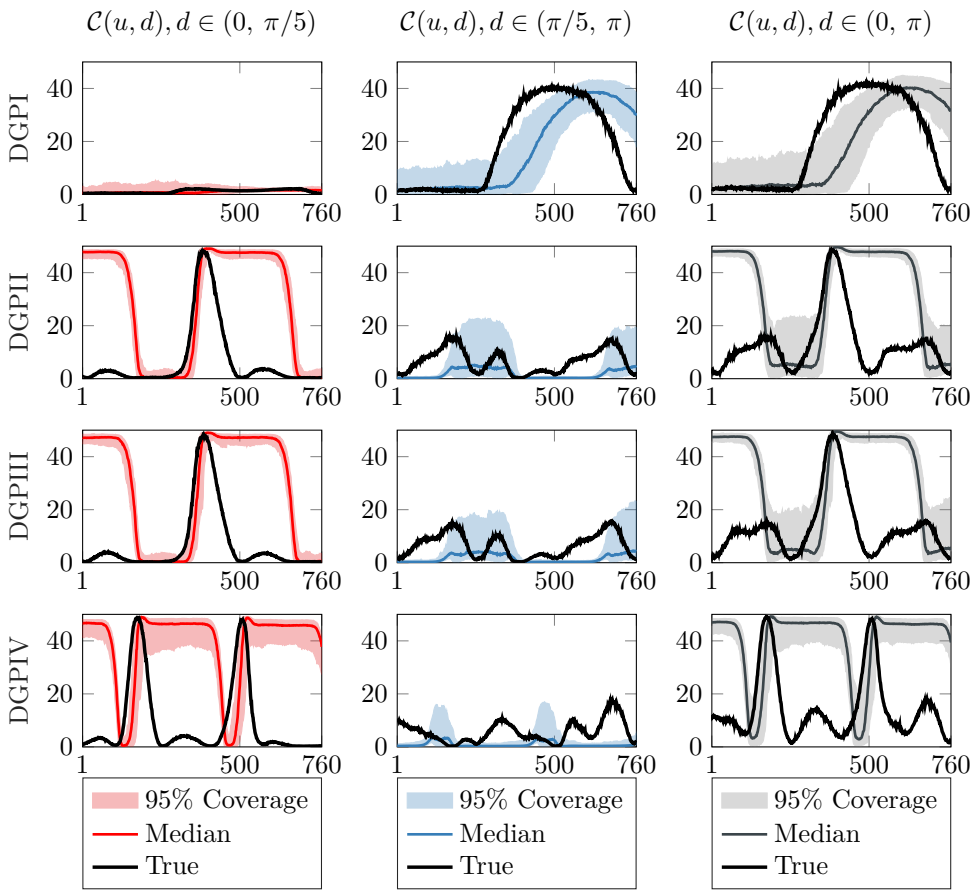

Figure C6: Dynamic network connectedness measures: True and fitted values
Notes: This figure plots the true network connectedness measures for three data generating processes following bi-variate TVP VAR(2) models along with the median and 95% quantiles of estimated network connectedness measures from a rolling VAR model we estimate using OLS and a 240 period window rolling through the sample. The left columns report network connectedness on the low-frequency band, $d \in (0,\ \pi/5)$, the middle columns show network connectedness on the high-frequency band, $d \in (\pi/5,\ \pi)$, and the right columns show the aggregate network connectedness such that $d \in (0,\ \pi)$. DGPI (top row) is a TVP VAR(2) model with Gaussian errors, we introduce a break in the time-varying autoregressive matrices and contemporaneous relations from observation 500 that induces large connections across the high frequency band. DGPII (second row) is a TVP VAR(2) where time-varying intercepts and autoregressive matrices following sin wave functions with a stochastic error, time-varying covariance matrix where the off-diagonals follow sin wave functions with a stochastic error, and the diagonal elements follow a stationary AR(1) processes. DGPIII (third row) is a TVP VAR(2) model with student-t errors, time-varying intercepts and autoregressive matrices following sin wave functions with a stochastic error, time-varying covariance matrix where the off-diagonals follow sin wave functions with a stochastic error, and the diagonal elements follow a stationary AR(1) processes. DGPIV (bottom row) is the same as DGPII, but with an increase in the periodicity of the respective sin wave functions the time-varying intercepts and autoregressive matrices follow.

C.4 Larger scale DGPs

In this section, we study how our approach performs using larger scale DGPs. For the sake of brevity and simplicity we refine our focus on DGPI, that turns on connectedness across higher frequencies at observation 500. We look at simulating N=10 and N=25 TVP-VAR(1) processes in the following manner:

$$\begin{aligned} \boldsymbol{X}_{t,T} &= \boldsymbol{\Phi}_0(u) + \boldsymbol{\Phi}_1(u)\boldsymbol{X}_{t-1,T} + \boldsymbol{\epsilon}_{t,T}, \\ \boldsymbol{\epsilon}_{t,T} &= \boldsymbol{\Sigma}^{-1/2}(u)\boldsymbol{\eta}_{t,T},\ \boldsymbol{\eta}_{t,T} \backsim (0, \mathbf{I}_N) \end{aligned}$$

where $\boldsymbol{\Phi}_0(u)$ contains the time-varying intercepts and $\boldsymbol{\Phi}_1(u)$ contains the time-varying autoregressive parameters. The time-varying covariance matrix $\boldsymbol{\Sigma}(u) = \mathbf{A}^{-1}(u)\mathbf{H}(u)\left(\mathbf{A}^{-1}(u)\right)^{\top}$ with $\mathbf{A}^{-1}(u)$ being a lower triangular matrix with a unit diagonal and $\mathbf{H}(u)$ is a $N \times N$ diagonal matrix.

Large Scale DGPI: The DGPs have residuals such that, $\boldsymbol{\eta}_{t,T} \backsim \text{NID}\left(0, \mathbf{I}_N\right)$. The time-varying intercepts follow the process:

$$\left[\boldsymbol{\Phi}_0(u)\right]_j = 0.0025\sin\left(0.004\pi t\right) + 0.15\sum_{i=1}^{t}\frac{\nu_i}{\sqrt{t}},\ \nu_i \backsim \text{NID}\left(0, 0.001^2\right),\ j = \{1, \ldots, N\}$$

For the time-varying autoregressive parameters for $t \in \{1, ..., 500\}$ the j, k elements of lie between

$$\left[\boldsymbol{\Phi}_1(u)\right]_{j,k} = \begin{cases} 0.005\sin\left(0.002\pi t\right) + 0.75\sum_{i=1}^{t}\frac{\kappa_i}{\sqrt{t}}, & \forall j, k \\ 0.0051\sin\left(0.002\pi t\right) + 0.75\sum_{i=1}^{t}\frac{\kappa_i}{\sqrt{t}}, & \forall j, k \end{cases}$$

Then from $t \in \{1, ..., 500\}$ the diagonal elements of $\boldsymbol{\Phi}_1(u)$ lie between

$$\left[\boldsymbol{\Phi}_1(u)\right]_{j,k} = \begin{cases} 0.5\sin\left(0.002\pi t\right) + 0.75\sum_{i=1}^{t}\frac{\kappa_i}{\sqrt{t}}, & j = k \\ 0.7\sin\left(0.002\pi t\right) + 0.75\sum_{i=1}^{t}\frac{\kappa_i}{\sqrt{t}}, & j = k \end{cases}$$

with the off-diagonal elements of $\boldsymbol{\Phi}_1(u)$ lying between

$$\left[\boldsymbol{\Phi}_1(u)\right]_{j,k} = \begin{cases} 0.005\sin\left(0.002\pi t\right) + 0.75\sum_{i=1}^{t}\frac{\kappa_i}{\sqrt{t}}, & \forall j \neq k \\ 0.0051\sin\left(0.002\pi t\right) + 0.75\sum_{i=1}^{t}\frac{\kappa_i}{\sqrt{t}}, & \forall j \neq k \end{cases}$$

with $\kappa_i \backsim \text{NID}\left(0, 0.0001^2\right)$. We obtain values by simulating random uniform numbers between the values we specify in the above.

The lower off-diagonal elements of $\mathbf{A}(u)$ for $t \in \{1, ..., 500\}$ lie between the following values and possess the following dynamics:

$$\left[\mathbf{A}(u)\right]_{j,k} = \begin{cases} 0.00\sin\left(0.002\pi t\right) + 0.7\sum_{i=1}^{t}\frac{\upsilon_i}{\sqrt{t}} \\ 0.03\sin\left(0.002\pi t\right) + 0.7\sum_{i=1}^{t}\frac{\upsilon_i}{\sqrt{t}} \end{cases}$$

Then, from $t \in \{501, ..., 1000\}$ they lie between the following values and have the following dynamics:

$$[\mathbf{A}(u)]_{j,k} = \begin{cases} 0.40 \sin(0.002\pi t) + 0.7 \sum_{i=1}^{t} \frac{\upsilon_i}{\sqrt{t}} \\ 0.60 \sin(0.002\pi t) + 0.7 \sum_{i=1}^{t} \frac{\upsilon_i}{\sqrt{t}} \end{cases}$$

with $\upsilon_i \backsim \text{NID}(0, 0.3^2)$. The diagonal elements of $\mathbf{H}(u)$ follow

$$\log[\mathbf{H}(u)]_{j,j} = \mu_j + \lambda_j \left(\log[\mathbf{H}(u-1)]_{j,j} - \mu_j \right) + \xi_{j,t}$$

where $\xi_{j,t} \backsim \text{NID}(\mu_j, 0.1^2/(1-\lambda_j))$, $\mu_j = 0.01$, $\lambda_j = 0.95$.

As before, this DGP has little to no dependence for the first 500 observations which means connectedness at both high and low frequency bands will be low and close to zero. The latter half of the sample sees the AR coefficients in each equation become persistent as the sin wave becomes negative. Note also that the contemporaneous relationship intensifies. This induces connections at high frequency bands while connections at low frequency bands should be low and close to zero.

For the N=10 and N=25 DGPs, we generate 100 simulations of length $T = 1000$ and compute the network connectedness measures. We use the median over these simulations as the true network connectedness. Then, for each of the 100 simulations, we fit the TVP-VAR model, with a bandwidth of W=8, we outline in Section 2.3. In fitting this model we take 1000 draws from the posterior distribution, calculate our network connectedness measures and then save the posterior median. For this exercise, we compute network connectedness on two frequency bands that cover the spectrum. The low-frequency band, which empirically pertains to persistent network connections, is $d \in (0, \pi/5)$, and the high-frequency band, pertaining to transitory network connections, is $d \in (\pi/5, \pi)$. For completeness, we compute the aggregate connectedness measures that considers the entire spectrum such that $d \in (0, \pi)$; this corresponds to a dynamic version of the Diebold and Yilmaz (2014) connectedness measure.

Figure C7 reports the true network connectedness measures and the median and 95% quantiles of corresponding estimates from the TVP-VAR model. We report network connectedness over the low-frequency-band, the high-frequency-band, and aggregate, in the left, middle, and right columns respectively. The top row corresponds to the N=10 DGP, and the second row corresponds to the N=25 DGP. As we can see, the distribution of estimates for each DGP track the true values well, when we expect to see surges in connectedness at the high frequency band from observation 500. In all cases, the true value lies within the 95% quantiles of the distribution from model estimates. This plot shows that our method is robust to increasing the number of variables within the VAR model and still provides an accurate representation of horizon specific network connectedness. Notably, as the as the number of variables increases, there is some bias present in the estimation; particularly for the N=25 models. However, such bias is worse when looking at rolling windows. These results are available upon request.

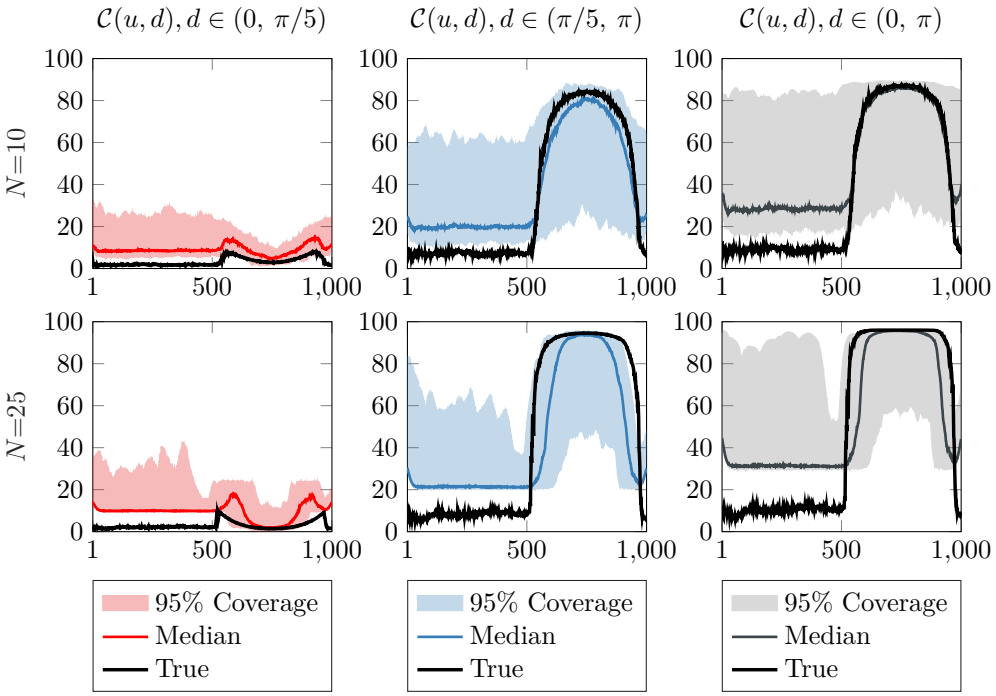

Figure C7: Dynamic network connectedness measures: True and fitted values for Larger Scale DGPs

Notes: The top row of this figure plots the true network connectedness measures for data generating processes following 10-variable TVP VAR(1) process along with the median and 95% quantiles of estimated network connectedness measures from a TVP-VAR model we estimate using the methods in Petrova (2019). The left columns report network connectedness on the low-frequency band, $d \in (0,\ \pi/5)$, the middle columns show network connectedness on the high-frequency band, $d \in (\pi/5,\ \pi)$, and the right columns show the aggregate network connectedness such that $d \in (0,\ \pi)$. The DGPs are N=10 (top row) and N=25 (bottom row) TVP VAR(1) models with Gaussian errors, we introduce a break in the time-varying autoregressive matrices and contemporaneous relations from observation 500 that induces large connections across the high frequency band.

D S&P 500 Sectors Breakdown and Additional Empirical Results

In this short section we present a U.S. stock market sectors breakdown and description where the S&P 500 index is used as a proxy for the stock market. The information in this section are reported as of January 25, 2019. For more details and updated information, see also https://us.spindices.com/indices/equity/sp-500.

- Consumer Discretionary: The consumer discretionary sector consists of businesses that have demand that rises and falls based on general economic conditions such as washers and dryers, sporting goods, new cars, and diamond engagement rings. At present, the consumer discretionary sector contains 11 sub-industries: Automobile Components Industry, Automobiles Industry, Distributors Industry, Diversified Consumer Services Industry, Hotels, Restaurants & Leisure Industry, Household Durables Industry, Leisure Products Industry, Multiline Retail Industry, Specialty Retail Industry, Textile, Apparel & Luxury Goods Industry, Internet & Direct Marketing. The total value of all consumer discretionary stocks in the United States came to $4.54 trillion, or about 10.11% of the market. Examples of consumer discretionary stocks include Amazon and Starbucks.

- Communication Services: From telephone access to high-speed internet, the communication services sector of the economy keeps us all connected. At present, the communication services sector is made up of five industries: Diversified Telecommunication Services, Wireless Telecommunication Services, Entertainment Media, Interactive Media and Services. the total value of all communication services stocks in the United States came to $4.42 trillion, or 10.33% of the market. The communications industry includes stocks such as AT&T and Verizon, but also the giants Alphabet Inc A and Facebook from 2004 and 2012, respectively.

- Consumer Staples: The consumer staples sector consists of businesses that sell the necessities of life, ranging from bleach and laundry detergent to toothpaste and packaged food. At present, the consumer staples sector contains six industries: Beverages Industry, Food & Staples Retailing Industry, Food Products Industry, Household Products Industry, Personal Products Industry, Tobacco Industry. The total value of all consumer staples stocks in the United States came to $2.95 trillion, or about 7.18% of the market and includes companies such as Procter & Gamble.

- Energy: The energy sector consists of businesses that source, drill, extract, and refine the raw commodities we need to keep the country going, such as oil and gas. At present, the energy sector contains two industries: Energy Equipment & Services Industry, and Oil, Gas & Consumable Fuels Industry. The total value of all energy stocks in the United States came to $3.36 trillion, or about 5.51% of the market. Major energy stocks include Exxon Mobil and Chevron.

- Financial: The financial sector consists of banks, insurance companies, real estate investment trusts, credit card issuers. At present, the financial sector contains seven industries: Banking Industry, Capital Markets Industry, Consumer Finance Industry, Diversified Financial Services Industry, Insurance Industry, Mortgage Real Estate Investment Trusts (REITs) Industry, Thrifts & Mortgage Finance Industry. The total value of all financial stocks in the United States came to $6.89 trillion, or about 13.63% of the market. JPMorganChase, GoldmanSachs, and Bank of America are examples of financial stocks.

- Health Care: The health care sector consists of drug companies, medical supply companies, and other scientific-based operations that are concerned with improving and healing human life. At present, the health care sector contains six industries: Biotechnology Industry, Health Care Equipment & Supplies Industry, Health Care Providers & Services Industry, Health Care Technology Industry, Life Sciences Tools & Services Industry, Pharmaceuticals Industry. The total value of all health care stocks in the United States came to $5.25 trillion, or about 15.21% of the market. Examples of health care stocks include Johnson & Johnson, and Pfizer.

- Industrials: The industrial sector comprises railroads and airlines to military weapons and industrial conglomerates. At present, the industrial sector contains fourteen industries: Aerospace & Defense Industry, Air Freight & Logistics Industry, Airlines Industry, Building Products Industry, Commercial Services & Supplies Industry, Construction & Engineering Industry, Electrical Equipment Industry, Industrial Conglomerates Industry, Machinery Industry, Marine Industry, Professional Services Industry, Road & Rail Industry, Trading Companies & Distributors Industry, Transportation Infrastructure Industry. The total value of all industrial stocks in the United States came to $3.80 trillion, or about 9.33% of the market.

- Information Technology: the information technology (IT) sector is home to the hardware, software, computer equipment, and IT services operations that make it possible for you to be reading this right now. At present, the information technology sector contains six industries: Communications Equipment Industry, Electronic Equipment, Instruments & Components Industry, IT Services Industry, Semiconductors & Semiconductor Equipment Industry, Software Industry, Technology Hardware, Storage & Peripherals Industry. The total value of all information technology stocks in the United States came to $7.10 trillion, or about 19.85% of the market. It is the largest sector in the S&P 500. Top IT stocks include Microsoft and Apple.

- Materials: The building blocks that supply the other sectors with the raw materials it needs to conduct business, the material sector manufacturers, logs, and mines everything from precious metals, paper, and chemicals to shipping containers, wood pulp, and industrial ore. At present, the material sector contains five industries: Chemicals Industry, Construction Materials Industry, Containers & Packaging Industry, Metals & Mining Industry, Paper & Forest Products Industry. The total

value of all materials stocks in the United States came to $1.77 trillion, or about 2.71% of the market. Major materials stocks include Dupont.

- Real Estate: The real estate sector includes all Real Estate Investment Trusts (REITs) with the exception of Mortgage REITs, which is housed under the financial sector. The sector also includes companies that manage and develop properties. At present, the Real Estate sector is made up of two industries: Equity Real Estate Investment Trusts, Real Estate Management & Development. The total value of all real estate stocks in the United States came to $1.17 trillion, or 2.96% of the market. The real estate industry includes stocks such as Simon Property Group and Prologis.

- Utilities: The utilities sector of the economy is home to the firms that make our lights work when we flip the switch, let our stoves erupt in flame when we want to cook food, make water come out of the tap when we are thirsty, and more. At present, the utilities sector is made up of five industries: Electric Utilities Industry, Gas Utilities Industry, Independent Power and Renewable Electricity Producers Industry, Multi-Utilities Industry, Water Utilities Industry. The total value of all utilities stocks in the United States came to $1.27 trillion, or about 3.18% of the market. Utilities stocks include many local electricity and water companies including Dominion Resources.

Table D1: Descriptive Statistics

Notes: This table reports descriptive statistics of (annualized) realized volatility estimates for all firms in the Consumers Discretionary (COND), Consumer Staples (CONS), Health Care (HLTH), Industrials (INDU), Information Technology (INFT), Materials (MATR), Real Estate (REAS), Financials (SPF), Energy (SPN), Comunication Services (TELS), and Utilities (UTIL) sectors. All descriptive statistics pool information across all stocks within the sector from July 5, 2005 to August 31, 2018. Mean is the sample average, Std is the sample standard deviation, Min and Max are the minimum and maximum values, Skew and Kurt are the sample skewness and kurtosis respectively.

	COND	CONS	HLTH	INDU	INFT	MATR	REAS	SPF	SPN	TELS	UTIL
Mean	28.024	19.129	23.492	25.934	25.618	28.263	25.233	26.445	31.459	27.911	18.399
Std	19.487	13.016	16.822	18.611	16.242	18.921	20.579	26.289	19.71	20.261	11.371
Min	4.219	2.975	0.000	3.062	1.862	4.509	3.776	1.859	4.721	5.237	4.314
Max	652.061	508.833	598.948	842.509	460.001	413.752	656.948	1526.688	752.732	294.655	302.308
Skew	4.071	4.869	5.402	4.451	3.968	3.200	4.237	6.733	3.448	2.832	4.666
Kurt	38.155	60.767	67.314	51.261	36.928	19.088	35.824	133.412	33.112	15.423	41.019

Test statistics for differences between sectoral short-term and long-term network connectedness

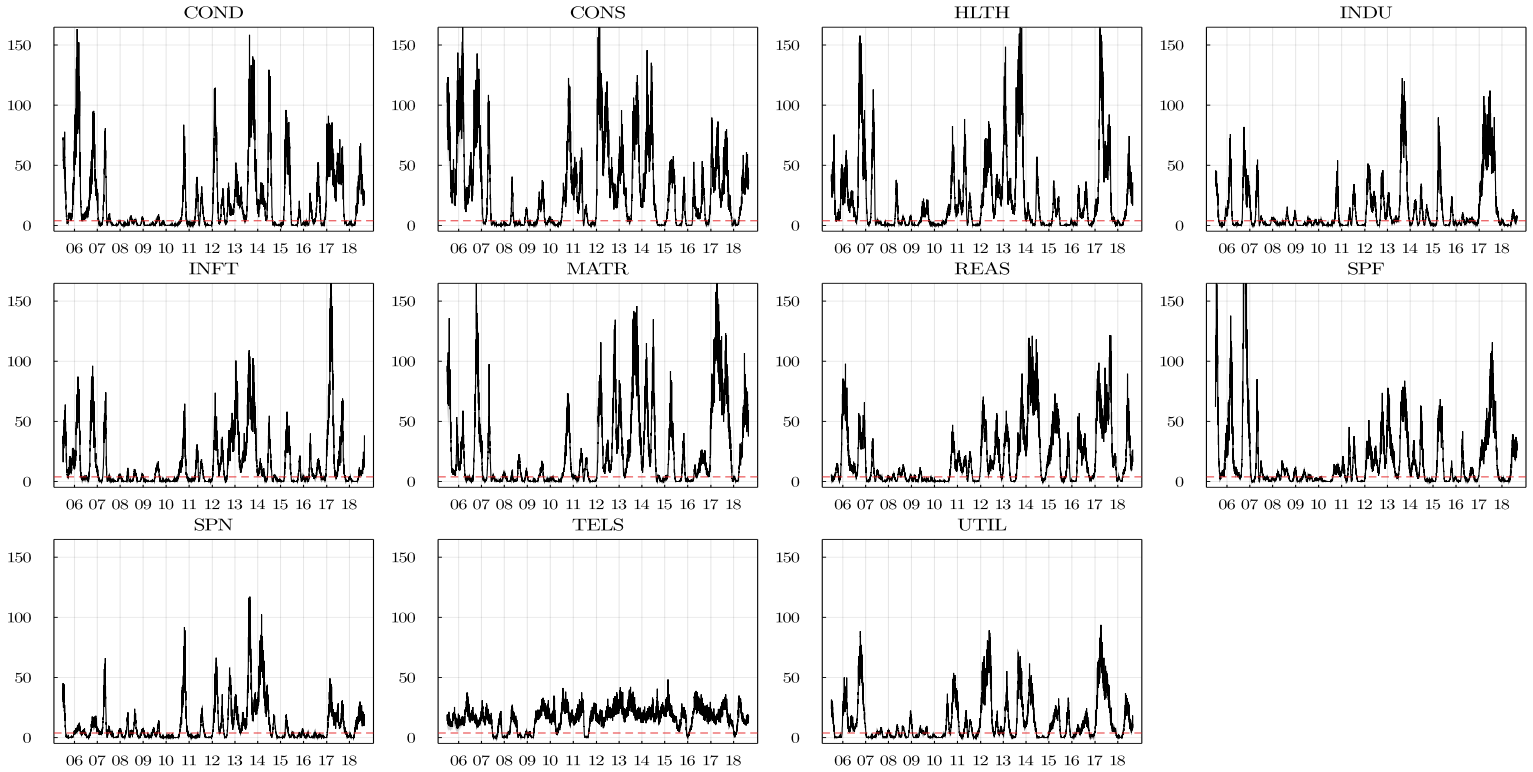

Figure D1: Tests for Heterogeneity of Short-term and Long-term of S&P500 Realized Volatility Networks Connectedness

This figure plots the test statistics and numerical standard errors for heterogeneities between short-term and long-term network connectedness measures that we compute on the realized volatilities firms within each S&P500 sectors: Consumers Discretionary (COND), Consumer Staples (CONS), Health Care (HLTH), Industrials (INDU), Information Technology (INFT), Materials (MATR), Real Estate (REAS), Financials (SPF), Energy (SPN), Comunication Services (TELS), and Utilities (UTIL) from July 8, 2005 to August 31, 2018.

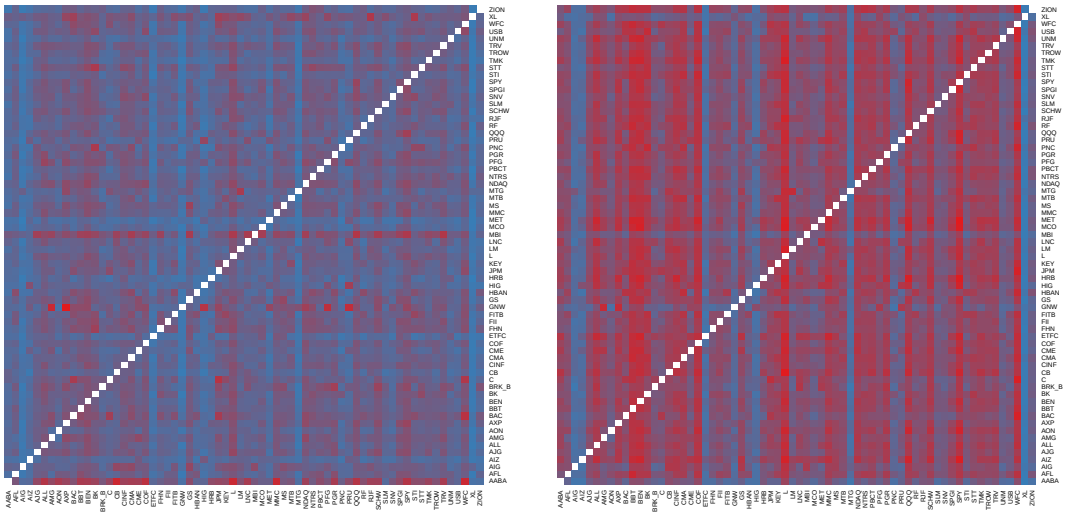

Figure D2: Transitory and Persistent Network of Financials: October 24, 2008
The left (right) figure depicts network connections among assets constituting the SPF sector driven by transitory (persistent) shocks during October 24, 2008 corresponding to the day when VIX reached highest value.

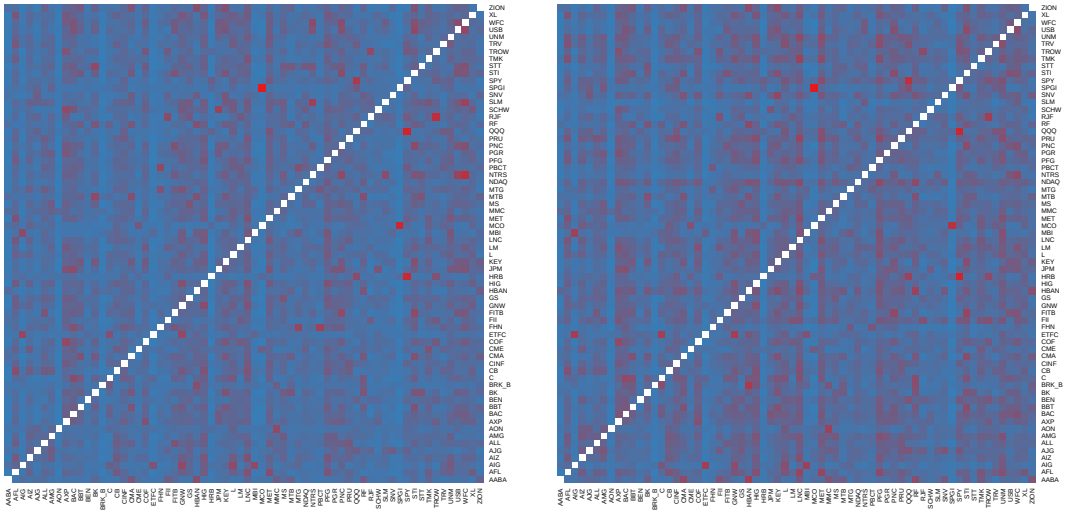

Figure D3: Transitory and Persistent network of Financials: October 24, 2009
The left (right) figure depicts network connections among assets constituting the SPF sector by transitory (persistent) connectedness. Arrows denote the direction of connections and the strength of lines denotes the strength of connections.

Table D2: Ranking of Institutions in the Financial Sector

Notes: This table reports ranking of institutions in financial sector according to contribution of shocks. The table ranks institutions at the two different periods according to being transmitters or receivers of transitory or persistent shocks from those transmitting or receiving the lowest to the ones that contribute the most.

	Transmitting Shocks				Receiving Shocks			
	Oct 24, 2008		Oct 24, 2009		Oct 24, 2008		Oct 24, 2009	
	transitory	persistent	transitory	persistent	transitory	persistent	transitory	persistent
AABA	9	41	6	5	**66**	1	6	35
AFL	26	27	53	48	32	35	25	57
AIG	2	3	9	7	49	6	21	8
AIZ	8	6	35	39	5	**64**	47	28
AJG	32	30	13	14	51	19	14	49
ALL	58	58	36	31	25	44	44	32
AMG	40	38	34	43	54	15	19	59
AON	23	18	1	1	18	51	4	10
AXP	55	51	61	61	53	16	48	42
BAC	44	35	58	57	**62**	5	24	58
BBT	57	**65**	44	40	37	30	32	43
BEN	**65**	**64**	11	11	26	43	11	52
BK	59	57	37	38	44	25	56	18
BRK_B	22	20	2	2	60	7	1	21
C	21	19	20	23	52	14	20	45
CB	31	46	33	33	8	59	40	30
CINF	52	56	51	52	38	32	37	41
CMA	56	53	42	45	20	48	45	31
CME	34	25	10	12	13	55	31	9
COF	51	61	48	47	30	40	49	33
ETFC	5	4	4	6	3	60	2	16
FHN	28	22	14	9	59	11	**62**	2
FII	16	32	24	26	56	12	13	61
FITB	24	21	39	28	48	17	36	39
GNW	3	7	59	58	**63**	4	55	29
GS	46	36	49	42	47	23	54	22
HBAN	15	11	38	49	42	24	10	**64**
HIG	6	5	**63**	**62**	12	53	60	26
HRB	12	26	5	3	4	**62**	5	25
JPM	47	40	32	30	21	47	53	20
KEY	41	44	54	55	31	38	57	23
L	**66**	**66**	55	59	36	34	29	55
LM	29	16	29	29	16	49	26	54
LNC	27	34	**66**	**66**	23	37	35	51
MBI	20	29	8	10	**65**	3	15	15
MCO	30	14	7	8	2	**65**	7	17
MET	7	23	**64**	**64**	1	**66**	43	44
MMC	54	55	23	27	41	28	27	40
MS	39	42	52	44	46	22	46	37
MTB	35	37	18	16	17	50	39	14
MTG	4	2	21	19	29	20	52	11
NDAQ	**63**	54	30	34	61	8	9	**66**
NTRS	49	47	26	20	34	36	**64**	3
PBCT	45	50	15	13	14	54	51	6
PFG	42	33	**65**	**65**	57	10	33	53
PGR	48	52	41	35	40	29	30	47
PNC	10	8	57	53	33	27	59	27
PRU	17	12	40	46	11	56	12	**65**
QQQ	**62**	60	28	36	35	33	23	48
RF	25	28	46	54	19	46	41	34
RJF	33	31	16	18	15	52	38	12
SCHW	60	59	17	15	43	26	61	4
SLM	18	13	19	17	24	42	**66**	1
SNV	14	17	12	22	55	13	8	56
SPGI	50	39	3	4	27	41	3	13
SPY	**64**	**63**	45	51	39	31	17	60
STI	43	24	**62**	60	9	58	50	38
STT	38	45	27	21	58	9	**63**	5
TMK	37	48	50	50	7	61	16	**63**
TROW	36	43	31	37	22	45	28	46
TRV	53	49	25	25	50	21	22	50
UNM	19	10	56	**63**	6	**63**	18	**62**
USB	13	15	47	41	45	18	58	19
WFC	61	**62**	60	56	28	39	**65**	7
XL	1	1	43	32	**64**	2	42	36
ZION	11	9	22	24	10	57	34	24