

MIT Open Access Articles

Multilayer network analysis of oil linkages

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Casarin, Roberto et al. "Multilayer network analysis of oil linkages." *Econometrics Journal* 23, 2 (January 2020): 269–296. © 2020 Royal Economic Society

As Published: <http://dx.doi.org/10.1093/ectj/utaa003>

Publisher: Oxford University Press (OUP)

Persistent URL: <https://hdl.handle.net/1721.1/130503>

Version: Original manuscript: author's manuscript prior to formal peer review

Terms of use: Creative Commons Attribution-Noncommercial-Share Alike



MULTILAYER NETWORK ANALYSIS OF OIL LINKAGES

ROBERTO CASARIN ENRIQUE TER HORST GERMAN MOLINA

RAMON ESPINASA CARLOS SUCRE ROBERTO RIGOBON

SUMMARY

This manuscript proposes a new approach for unveiling existing linkages within the international oil market across multiple driving factors beyond production. A multi-layer, multi-country network is extracted through a novel Bayesian graphical vector autoregressive model, which allows for a more comprehensive, dynamic representation of the network linkages than traditional or static pairwise Granger-causal inference approaches. Building on the complementary strengths in Espinasa et al. (2017) and Rousan et al. (2018), the layers of the network include country- and region-specific oil production levels and rigs, both through simultaneous and lagged temporal dependences among key factors, while controlling for oil prices and a world economic activity index. The proposed approach extracts relationships across all variables through a dynamic, cross-regional network. This approach is highly scalable, and adjusts for time-evolving linkages. The model outcome is a set of time-varying graphical networks which unveil both static representations of world oil linkages and variations in micro-economic relationships both within and between oil producers. An example is provided, illustrating the evolution of intra- and inter-regional relationships for two major inter-connected oil producers: the United States, with a regional decomposition of its production and rig deployment, and Arabian Peninsula and key middle east producers, with a country-based decomposition of production and rig deployment, while controlling for oil prices and global economic indices. Production is less affected to concurrent changes in oil prices and the overall economy than rigs. However, production is a lagged driver for prices, rather than rigs, which indicates that the linkage between rigs and production may not be fully accounted for in the markets.

Keywords: Bayesian Graphical Models, Dynamic Multilayer Network analysis, Rigs, Production, Granger Causality, Oil Linkages.

1. INTRODUCTION

With an increase in globalization, the concept of networks has become an integral part of human activities, including information and product exchanges at the individual level (Dunbar et al., 2015), at the corporate level (Bernard and Moxnes, 2018), and at the global level (Maluck and Donner, 2015). Actions of economic agents in local economies can now have larger impact on a global scale due to strong connections among individuals, institutions, and markets. These connections play a fundamental role in the spread of information and risk (Broner and Ventura, 2016). Actions of major financial institutions have significant

consequences on the stability or public confidence in the entire system (Billio et al., 2012, Battiston et al., 2012, Brunnermeier and Pedersen, 2009).

The academic literature has attempted to mimic this societal evolution, and network models have become a center of attraction for statistical/econometric modeling of dependencies in real world phenomena, due to their intuitive approach to visualization and interpretation of complex relationships (Goldenberg et al., 2010).

Network analysis and graph theory (Bollobas, 2001, 2002) has been shown to be a promising tool with the potential to help in monitoring the both interconnectedness and centrality (Bonacich, 2007) of financial institutions and markets, and assess general system vulnerabilities (Bianchi et al., 2018). This has led to considerable increase in research on the statistical properties of network measures for systemic risk analysis (Battiston et al., 2012, Billio et al., 2012, Diebold and Yilmaz, 2014, Hautsch et al., 2014, Puliga et al., 2016).

Oil markets provide a natural application of network models due to their highly interconnected nature (Economou et al., 2017). Multiple approaches have been proposed to the use of network models in oil markets, either isolating the production-only relationships and providing a dynamic representation of production-driven linkages (Rousan et al., 2018), or providing a fuller picture of the micro-level relationships, but offering only a static approach to their linkages (Espinasa et al., 2017).

As discussed in Espinasa et al. (2017), production of oil consists of several steps: exploration, drilling, extraction, and commercialization. When demand for oil increases and prices rise, the supply chain responds by increasing exploration efforts and drilling activity, which is followed by increased production levels. This does not imply that production is the result of current, or even recent, drilling activity. Before oil production can expand, some physical prerequisites must be fulfilled. The most important one is that new oil wells have to be developed, which requires drilling activity. Current oil production cannot respond immediately to simultaneous changes in the oil price. There is a lead time to increase capacity and bring production in line, always assuming that there is no spare production capacity.

The approach proposed in this paper builds on Rousan et al. (2018) (multi-country, single-layer) and Espinasa et al. (2017) (multi-layer, single country), yet it differs from those two constrained perspectives of the linkage exploration problem by providing a dynamic,

yet complete network representation of the linkages across (and while controlling for) a wider set of relevant factors identified at the micro level. This study also contributes to the recent and expanding stream of literature on applied network econometrics (Diebold and Yilmaz, 2014), by adapting Bayesian graphical model for SVAR processes Ahelegbey et al. (2016a) to the expanded oil network structure studied in (Espinasa et al., 2017), leading to jointly estimating the shape of the world oil network. Also the findings in this paper are not feasible with the aforementioned approaches, as some dynamic linkages occur at the inter-factor level.

This manuscript elicitates a novel method for identification of current state and evolution, within- and cross-country, of the micro-economic network linkages, accounting for relationships between factors affecting both supply and demand, including locally-based factors (rigs activity and oil production per regional area or country), global oil prices, and the state of the global economy. The main novelty of the proposed approach is its multi-dimensional nature, in contrast to previous cross-sectional proposed analyses. To our knowledge, this is the first model in the oil literature to combine through a holistic single system: (1) multiple factors at the micro level per oil-producing region/country; (2) inter-temporal associations (both contemporaneous and lagged) between factors; and (3) dynamic linkages between factors, both within- and between-region/country. This comprehensive world network model offers the additional advantage of allowing for scenario analyses, which in turn allow for anticipation of (co-)movements. The proposed methodology allows economic actors affected by oil dynamics to react or hedge to particular geographical or geopolitical risks by focusing on the (evolving) core linkages most affecting their own production or supply.

Ahelegbey et al. (2016a) show that estimating networks by using conditional Granger-causality or pairwise Granger-causality (Billio et al., 2012) can lead to spurious linkages and to overestimation of the level of connectedness. In order to avoid these issues we employ graphical models for multivariate systems (Whittaker, 1990, Dawid and Lauritzen, 1993, Lauritzen, 1996, Carvalho and West, 2007, Wang and West, 2009, Wang, 2010, Rodriguez et al., 2011, Wang et al., 2011). More specifically, this paper proposes a Bayesian Graphical VAR model to retrieve the world oil network (Corander and Villani, 2006, Ahelegbey et al., 2016a,b), with the added benefit from the Bayesian framework of a lesser reliance on data

length, which is especially relevant in more recent oil producing areas, or those for which limited data is available.

The paper is organized as follows: Section 2 describes the data used in the study and provides a short description of the model used to extract the world network. Sections 3 and 4 describe the results, decomposing them into simultaneous and lagged networks. Section 5 provides a summary of the proposed approach and conclusions.

2. DATA AND METHODOLOGY

To illustrate the proposed methodology, a group of intra- and inter-connected subsets of oil producing countries and regions is selected. This is not meant to be a comprehensive list of oil producers, but an illustrating subset that is sufficiently large to: (1) demonstrate the scalability of the methodology; (2) unveil intra-regional linkages (including inter-temporal and cross-factor linkages); and (3) unveil inter-regional linkages (again, also inter-temporal and cross-factor).

The regions and countries included in the analysis are the following (sorted alphabetically by their short codes in parenthesis, which will be used for descriptive and graphical representation throughout the manuscript): Alaska (A), Golf of Mexico (GM), Irak (IK), Kuwait (KW), Lower48¹ (LO), Oman (OM), Qatar (QA), Saudi Arabia (SA), United Arab Emirates (UAE), USA Bakken (USB), USA Eagle Ford (USAE), and USA Permian (USAP).

For the purpose of non-graphical representations, the variable short code will be used, while for graphical purposes, the variables will be color-coded. Following Espinasa et al. (2017), two region- and country-specific variables are considered in this manuscript: (1) Production, which will have a short code of Y and a blue color code; and (2) Rigs, which will have a short code of R , and a red color code. The approach is clearly scalable, and more region-specific factors of interest could easily be introduced. Common factors affecting all regions and countries (and other factors), both lagged and synchronously, are denoted as PT and DT , representing the oil price and economic index variables, respectively.

¹Lower48 includes the United States without Alaska, Golf of Mexico, Bakken, Eagle Ford and Permian.

2.1 Data description

Building on data used in Espinasa et al. (2017), the number of rigs per country at each point in time is taken from Baker Hughes Inc.², an oil services company that provides the count of active oil rigs for the 65 most important oil producing countries (with the exception of Russia and China) on a monthly basis.

Also following Espinasa et al. (2017), we use the classification that distinguishes between those rigs used for crude oil and natural gas activities. Baker Hughes began publishing data with this distinction in January 1995. Baker Hughes registers rotary rigs³ ; while they are not the only type of rigs available in the market, they are the most used worldwide.

For the purpose of this study, prices were taken from the most widely used crude price index: West Texas Intermediate (WTI), settled in Cushing Oklahoma on a monthly basis. Price is expressed in real terms using the US CPI provided by the Bureau of Labor Statistics, and log-returns are used throughout the manuscript for this variable. The sensitivity of the analysis to this choice of WTI was tested by comparing the results with Brent-price driven results, which rendered low differences (this was expected due to the large correlation between the two indices and the level-agnostic approach proposed in this paper). Finally, average production per country was used, provided on a monthly basis by the International Energy Agency (IEA), for the relevant period. The monthly crude supply can be found in the IEA MODS database (Monthly Oil Data Service).

The data used in this study has been detrended, so that the focus remains exclusively in the relationships among all factors, rather than spurious common-trend-driven relationships. It corresponds to 18 full years of monthly observations, going from January 2000 to December 2017.

2.2 Bayesian Graphical Model Approach

In this section a primer on Bayesian Graphical Models is provided in the context of the oil network application. Graphical models are stochastic models that summarize the depen-

²http://investor.shareholder.com/bhi/rig_counts/rc_index.cfm

³According to Baker Hughes Inc., a rotary rig rotates the drill pipe from surface to drill a new well (or sidetracking an existing one) to explore for, developing and producing oil or natural gas.

2.2 Bayesian Graphical Model Approach

dence structure among random variables by means of graphs (Brillinger, 1996). Specifically, a graph is characterized by nodes and edges, where the nodes represent variables and the edges depict the nature of the interaction among variables. For instance, the relationship $P \rightarrow Q$ means the variable P causes the variable Q . The node P from which a directed edge originates is the parent (explanatory variable), and its end Q is the child (response variable).

One of the appealing features of the graphical approach to multivariate time series analysis is the possibility of giving a graphical representation of the logical implications of models as well as the conditional independence relationships. As an example, assume $P \rightarrow Q \rightarrow R$, then P and R would be probabilistically dependent in the absence of Q ; but conditional on Q , they would be independent. In oil network terms, as can be seen in Fig. 1, Lower48 rigs affect Saudi Arabia production, and Saudi Arabia production affects UAE production, but there is no direct link between Lower48 rigs and UAE production without the presence of Saudi Arabia as a relevant intermediate actor from the perspective of UAE. This aligns with the regional and world-wide weight of each of the three producers, and their relative influence on each other.

A graphical approach provides a coherent framework to represent and assess these relationships. In this paper, both undirected and directed graphs are employed. In the directed graphs, the edge orientation presents a direction of Granger-like causation among the variables. Edges linking two nodes through a clock-wise oriented curve represent a Granger-causal relationships between the two connected variables (the one originating the edge Granger-causes the one receiving the edge, where origination and reception are based on the clock-wise orientation of the connecting edge). Using such class of models provides information on the structural dynamics among the variables by means of the directed edges.

Let $Y_t = (Y_{1,t}, Y_{2,t}, \dots, Y_{n,t})$, where $Y_{i,t}$ is a realization of the i -th variable at time t . Equation (1) summarizes a graphical representation with a one-to-one correspondence between the coefficient matrices and directed acyclic graphs (DAG):

$$Y_{i,t-s} \rightarrow Y_{j,t} \iff \Psi_{s,j,i}^* \neq 0 \quad 0 \leq s \leq p \quad (1)$$

where $\Psi_{s,i,j}^*$ contains the coefficient representing the directed linkage between variables j and i occurring between the concurrent values of the latter and those s periods lagged for the former.

By considering the structural dynamics as a causal dependence among variables, the relationship in (1) for $1 \leq s \leq p$ can be referred to as lagged (temporal) dependence, and the relationship for $s = 0$ can be referred as the contemporaneous dependence. Temporal dependence is based on the time flow and relies on the assumption that causes precede effects in time. The temporal dependence structure is represented naturally through a directed graph, with edge orientation indicating the direction of the causation. Contemporaneous causal relationships are based on distinguishing between instantaneous causation from correlations and are represented through undirected graphs.

Let $\mathcal{Y} = (Y_1, \dots, Y_T)$ be a time series of n variables and length T . The joint distribution of the variables in \mathcal{Y} can be described by a graphical model $(G, \theta) \in \{\mathcal{G}, \Theta\}$, where G is a graph, θ is a vector of structural parameters, \mathcal{G} is the space of the graphs, and Θ is the parameter space. We assume the graph $G = (V, E)$ is an ordered pair of sets V and E . The vertex set V represents the set of random variables, and the edge set $E \subset V \times V$ is the set of pairs of vertices indicating the presence of structural relationships between the variables. We assume that the undirected (directed) graph $G \in \mathcal{G}$, which defines the contemporaneous (temporal) dependence among the variables, is a directed acyclic graph. The structural parameters are $\theta \equiv \{\mu, \Sigma\} \equiv \{\mu, \Psi^*, \Sigma_\varepsilon\}$, where μ is n vector of means of $Y_t = (Y_{1t}, \dots, Y_{nt})$, $\forall t$ and Σ is the covariance matrix of the observed time series that decomposes into $\{\Psi^*, \Sigma_\varepsilon\}$, where $\Psi^* = (\Psi_0^*, \Psi_1^*, \dots, \Psi_p^*)$, is the matrix of structural coefficients and Σ_ε is the structural error covariance matrix. Without loss of generality, and since the data has been detrended, we assume the data is generated by a stationary process ($\mu = 0$).

In equation (1), we set:

$$\Psi_s^* = (A_s \circ \Phi_s), \quad 0 \leq s \leq p \quad (2)$$

where for $s = 0$, Ψ_0^* is $n \times n$ structural coefficients of contemporaneous dependences. A_0 is a $n \times n$, 0-1 valued matrix (adjacency matrix) and Φ_0 is a $n \times n$ matrix of coefficients. For

2.2 Bayesian Graphical Model Approach

$1 \leq s \leq p$, Ψ_s^* is a $n \times n$ matrix of structural coefficients of temporal dependences, A_s is a $n \times n$ binary connectivity matrix and Φ_s is a $n \times n$ matrix of coefficients. The operator (\circ) is the element-by-element Hadaamard's product (i.e., $\Psi_{s,i,j}^* = A_{s,i,j} \Phi_{s,i,j}$). We refer to A_0 as the connectivity matrix of contemporaneous dependence and to A_s , $1 \leq s \leq p$, as the matrix of the temporal dependence.

The set of matrices A_s , $0 \leq s \leq p$, contain indicators for the existence of causal relationships such that $A_{s,i,j} = 1 \iff Y_{j,t-s} \rightarrow Y_{i,t}$, and $A_{s,i,j} = 0$ otherwise.

Elements in Φ_s , $0 \leq s \leq p$, are structural regression coefficients, such that $\Phi_{s,i,j} \in \mathbb{R}$ represents the value of the effect of $Y_{j,t-s}$ on $Y_{i,t}$. There is a one-to-one correspondence between Φ_s and Ψ_s^* conditionally on A_s :

$$\Psi_{s,i,j}^* = \begin{cases} \Phi_{s,i,j} & \text{if } A_{s,i,j} = 1 \\ 0 & \text{if } A_{s,i,j} = 0 \end{cases} \quad (3)$$

Based on the notation in (2), a graphical VAR model can be written as:

$$Y_t = (A_0 \circ \Phi_0)Y_t + \sum_{s=1}^p (A_s \circ \Phi_s)Y_{t-s} + \varepsilon_t \quad (4)$$

where $(A_s \circ \Phi_s)$ are the graphical model structural coefficient matrices whose non-zero elements describe the value associated with the contemporaneous and temporal dependences, respectively.

To complete the Bayesian network model, the prior distribution for Ψ^* is assumed normal and uninformative in the choice of hyperparameters, i.e., $\Psi^* \sim \mathcal{N}(\underline{\Psi}^*, \underline{V}_{\Psi})$.

Estimating (4) involves specification of the lag order, p of the causal structure, $A = (A_0, A_1, \dots, A_p)$, and the set of parameters, $\{\Psi_0^*, \Psi_1^*, \dots, \Psi_p^*, \Sigma_\varepsilon\}$ from Σ . A uniform prior is assumed for G on the set of complete graphs and, completing the specification, a conjugate Wishart prior distribution is assumed for $\Omega = \Sigma^{-1}$.

3. OIL NETWORK RESULTS

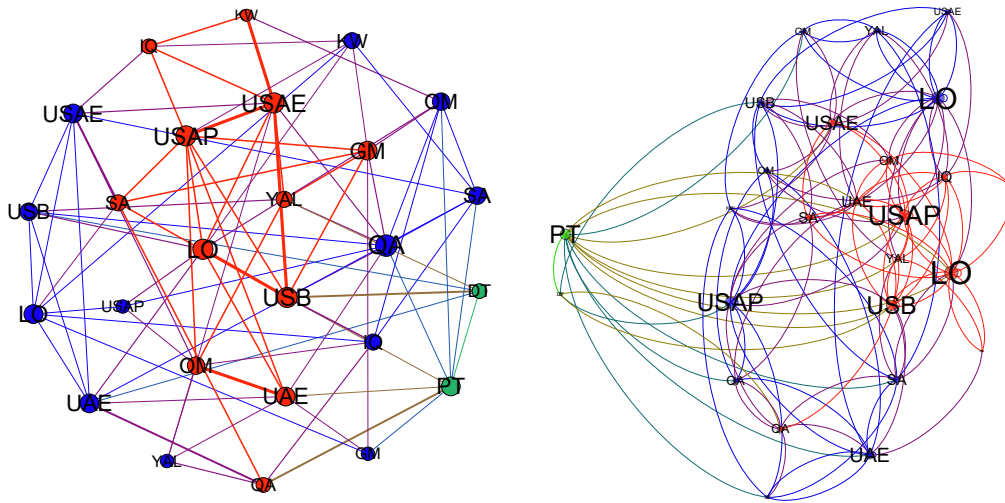


Figure 1: Contemporaneous (left panel) and lagged (right panel) oil production (red vertices) and rigs (blue vertices) networks. Green vertices correspond to the WTI oil price returns (PT) and the world economic activity index (DT). The size of each nodes increases with the number of linkages.

The results using BIC indicate that the optimal lag can be found at $p=1$. This aligns with the results in Espinasa et al. (2017), who also found that the information from further lags are already contained in the $s=1$ lag at the micro-level for multiple examples in their single-country analysis. Therefore, for the purpose of simplicity in the notation, we will drop the sub-index s when referring to lagged linkages.

Fig. 1 presents the contemporaneous (left panel) and lagged (right panel) linkages on the oil market considering oil production and rigs, after controlling for oil price returns PT_t (Cushing, OK WTI Spot Prices FOB in dollars per barrel), and the world economic activity index DT_t , built from Kilian (2009). For expository purposes only nodes with a total degree larger than 8 (threshold for graphical relevance) have been considered. The top-left panels in Fig. 8-9 provide a representation without this filter. Contemporaneous linkages in Fig. 1 are an undirected graph, whereas lagged linkages are directed graphs. Node sizes (and short name code sizes in the graphs) relate to the number of linkages for the corresponding node.

Following the results in the top-left panel of Fig. 1, the contemporaneous linkages network indicate the countries which are more likely affected by a joint shock. The central countries in the contemporaneous linkages networks are Eagle Ford, Permian, Bakken, and

2.2 Bayesian Graphical Model Approach

United Arab Emirates (USAE, USAP, USB, UAE) in terms of oil production (red vertices) and Eagle Ford, Bakken, Lower48, and United Arab Emirates (USAE, USB, LO, UAE) in terms of number of rigs (blue vertices). Also it is worth noting that the results from the model explored in Espinasa et al. (2017) are confirmed in this graphical representation, where production appears to be relatively independent of oil prices or the world economy, conditional on rigs, across most countries. This shows that production decisions are often-times made at the rig level (increasing or decreasing the number of rigs) rather than through increases in productivity per rig. UAE and Qatar are the only two countries/regions where current prices can drive increases in production independent of rig decisions (direct linkages between PT and red nodes in the top left panel of Fig. 1). The state of the world economy appears to be a factor affecting mostly rig decisions as well, with the exception of Bakken crude oil production, which was one of the first producers to use fracking due to geological advantages versus others such as Permian. An expected relationship is the linkage observed between PT and DT (world economy and oil prices). Finally, it appears that the US (across different regions) and the smaller countries in the Arabian Peninsula are most affected by other actors, while Saudi Arabia, despite its size, appears to be relatively less dependent from external influences (though there is a strong linkage between its rigs and oil prices, which aligns with a historically large adaptability of its production decisions to oil market dynamics). The lagged linkages network shows the main contributors to the transmission of shocks in the aftermath of a shock to one node of the network. The centrality of a node i in the network can be measured by its in, out, and total degree, which are defined, respectively:

$$d_i^+ = \sum_{j \in V} A_{i,j} \quad (5)$$

$$d_i^- = \sum_{j \in V} A_{j,i} \quad (6)$$

$$d_i = \sum_{j \in V} A_{i,j} + \sum_{j \in V} A_{j,i} \quad (7)$$

From the results in the right panel of Fig. 1, the oil producing countries (red vertices) with the largest number of linkages (i.e. the highest total degree) are Permian, Lower 48, and

Bakken (USAP, LO, USB). This indicates the relevance of the United States as a point of influence in the oil market, and, more especially, the biggest producer, the Permian region. While the remaining actors are very connected, and the network structure is complex (a conclusion expected due to the level of connectedness in oil markets), the non-U.S. producers have a smaller relevance in the lagged network, indicating that their role is mostly reactive. This can be seen as well in the small number of clockwise linkages between nodes that originate in those regions, compared to the United States. The price PT appears to be a major lagged factor, with most of the connections showing counter-clockwise linkages. Prices are reactive (with a lag) to production changes, rather than rig changes. This indicates that the market is not fully anticipating the changes in production due to changes in rigs. As mentioned in Espinasa et al. (2017), there is a lag between production increases and rig increases, due to physical and technological reasons. However, the network results indicate that PT reacts to lagged changes in production (which occur later than the changes in rigs), showing evidence of a delay in the digestion of the rig increase/decrease news (or perhaps imperfect information about the impact of rig changes on production changes). This could be an indication of potential opportunities for arbitrage due to imperfect information about the nature and dimension of lagged linkages between rigs and production. Permian, Lower 48, and United Arab Emirates (USB, LO, UAE) are central in terms of shocks to rigs (blue vertices).

The set of vertices V can be partitioned in oil production variables V^Y and rig variables V^R such that $V = V^Y \cup V^R$. The partition is used to define the two sub-graphs $G^Y = (V^Y, E^Y)$ and $G^R = (V^R, E^R)$ representing the oil production and the rig networks, respectively, with $E^Y = \{\{v_1, v_2\} \in E; v_1, v_2 \in V^Y\}$ and $E^R = \{\{v_1, v_2\} \in E; v_1, v_2 \in V^R\}$. We define similarly the bipartite graph $G^{RY} = (V, E^{RY})$ representing linkages between oil production and rigs, with $E^{RY} = \{\{v_1, v_2\} \in E; v_1 \in V^Y, v_2 \in V^R\}$. Decoupling the linkages in oil production, rigs and oil-rigs linkages is given in Fig. 8-9. This allows for relevant sub-analysis of the kind shown in Rousan et al. (2018), which is one possible partition of our proposed method (production-only sub-graph). This type of analysis is of relevance when exploring univariate relationships focusing exclusively on a factor of interest, such as production. However, some information can be lost when isolating a factor and excluding the

4.1 Rigs and Oil Market

information contained in cross-factor linkages, such as the one explored in Espinasa et al. (2017).

4. A DYNAMIC CONNECTEDNESS ANALYSIS

In order to assess the stability of the network relationships, a 60-month rolling window is used to extract the oil network sequentially by fitting the proposed Bayesian graphical VAR model.

4.1 Rigs and Oil Market

Let A_t denote the adjacency matrix estimate at time t of either lagged or contemporaneous networks, and $A_{i,j,t}$ denote its (i, j) -th element. The In- and out-connectedness and total connectedness measures, either lagged or contemporaneous, are defined as

$$d_{it}^+ = \sum_{j \in V} A_{i,j,t}, \quad d_{it}^- = \sum_{j \in V} A_{j,i,t}, \quad d_{it} = \sum_{j \in V} (A_{i,j,t} + A_{j,i,t}) \quad (8)$$

where in this case the sub-index t denotes the endpoint of the rolling window. This addition to the notation simply accommodates for the rolling nature of the stability experiment, rather than being an intrinsic part of the model. The results for d_{it} , $i = 1, \dots, 26$ and $t = 61, \dots, 215$ (from December 2004 to December 2017) are presented in Fig. 4, which shows the lagged (solid) and contemporaneous (dashed) total connectedness in the oil production and rig activity ($d_t = \sum_{i \in V} d_{it}$).

The top panel in Fig. 4 displays the total degree of contemporaneous (dashed) and lagged (solid) linkages (as a proportion of the total number of such linkages that is possible). As shown in this panel, countries (production and rig activity) are more connected through responsive lags (contagion effects) than through contemporaneous shocks. Also the level of connectedness has been increasing in the latter part of the sample. However, this increase is concentrated in the lagged linkages, indicating that there is an increased lagged action-reaction responsiveness, rather than a higher level of concurrent interconnectedness.

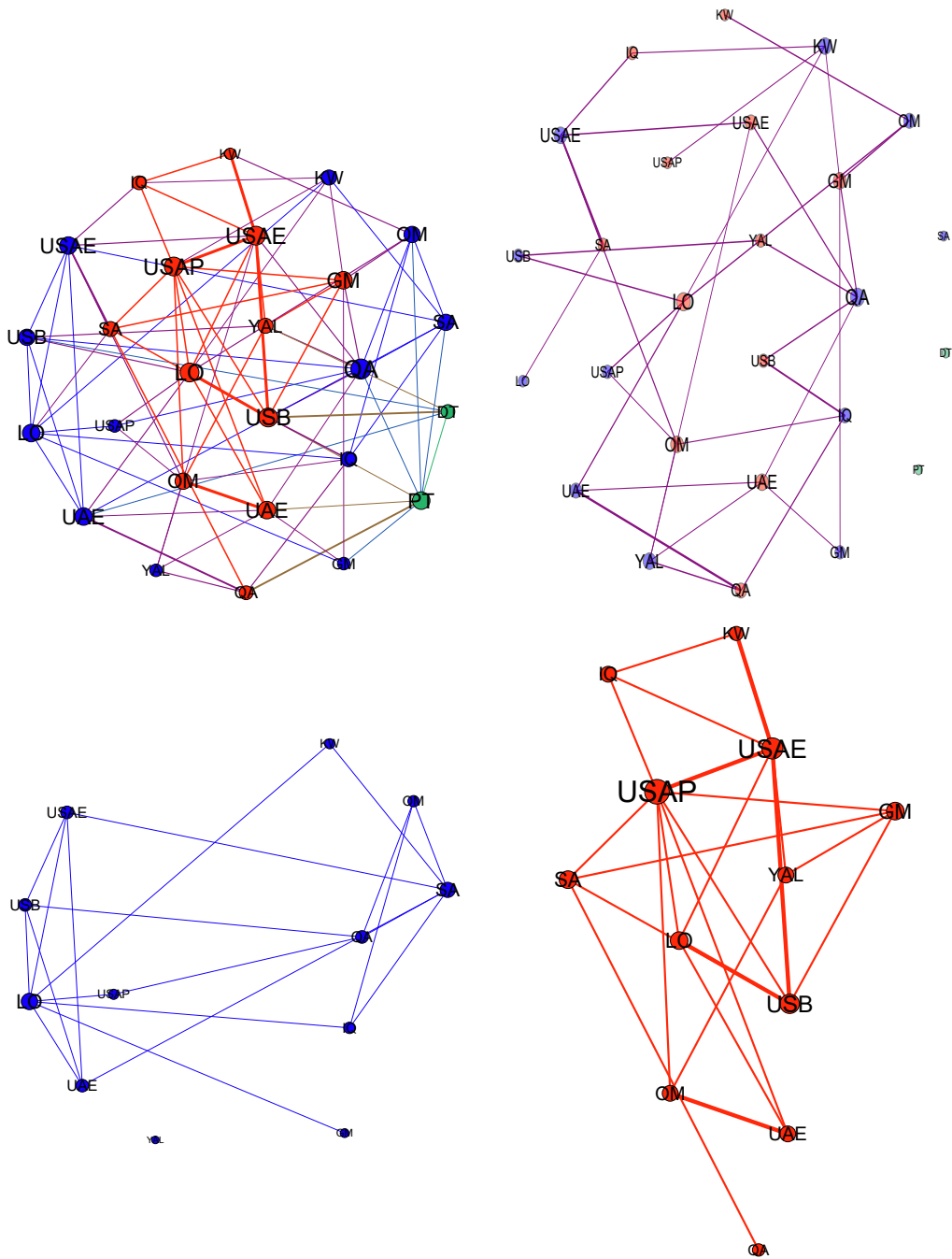


Figure 2: Contemporaneous linkages. Oil market network (top left panel). Production-rigs network (top right panel), production layer network (bottom left panel) and rigs layer network (bottom right panel).

4.1 Rigs and Oil Market

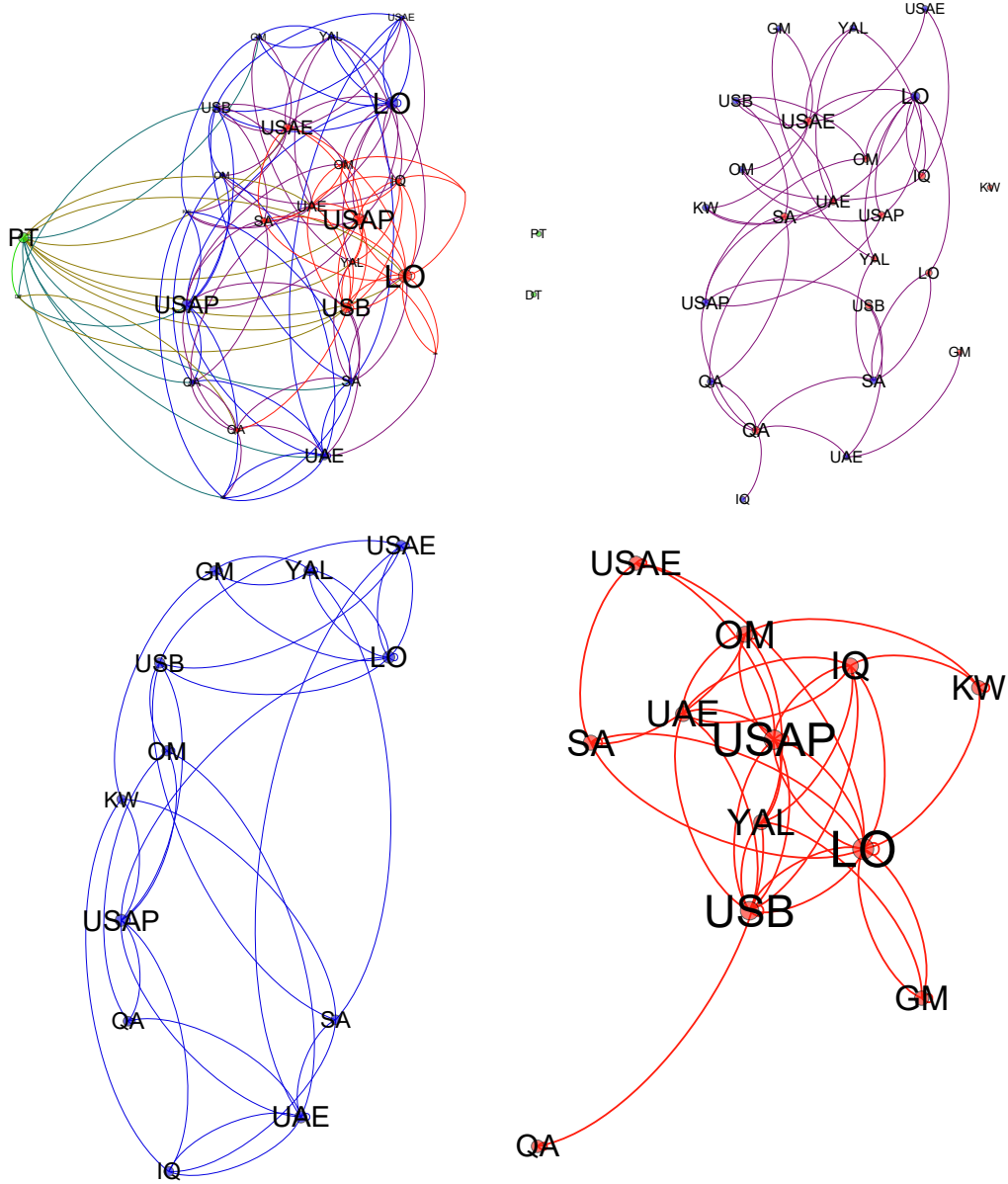


Figure 3: Lagged linkages. Oil market network (top left panel). Production-rigs network (top right panel), production layer network (bottom right panel) and rigs layer network (bottom left panel).

After 2011 there is a substantial change in the topology of the network. There is an increase of centrality of the nodes and of the number of rigs. More specifically, the top-right panel in Fig. 4 shows the mean and some key quantiles of the degree distribution. Year 2011 marks the beginning of some nodes with low degree becoming more connected, and the degree distribution becoming more concentrated about the mean. This is indicative of the higher level of interdependence of some of those actors, especially around the geopolitical contagion effects due to the protests starting in Egypt in 2011 and spreading through the region over the following months and years.

The Bottom-left panel in Fig. 4 shows the percentage of nodes which are *hubs*, which are defined as nodes with successors of high authority scores (solid line). This panel also shows the percentage of nodes which are *authorities*, which are defined as nodes with predecessors of high hub scores (dashed). There is an increase in the number of hubs (and the variability of this number) after 2011, indicating that there is heterogeneity in the linkages among oil activities, since a large majority of nodes have low degree, but a small number, known as hubs (important regions/countries), have a high number of linkages. Around 15% of the nodes fall into either category of hub or authority, placing them as central nodes of the oil economy at any point in time, while the rest can be considered more marginal nodes in the big picture of the oil economic network between the two geopolitical regions considered in this study.

When representing the oil production and extraction system as a network, the increased system fragility is reflected in a symmetric degree distribution with thinner tails (see Barabási and Albert, 1999, Acemoglu et al., 2012). Conversely, asymmetry and fat tails suggest heterogeneity in the linkages among regions/countries. In this case, the system is robust to random failures, but vulnerable to targeted attacks. Billio et al. (2016) exploited the heterogeneity of the network and consider entropy to build a systemic risk indicator. The bottom-right panel in Fig. 4 shows the largest and smallest eigenvalues of the adjacency matrix. The smallest eigenvalue gives a bound on the size of the largest independent set and of the smallest *clique* in the graph (defined as a subset of vertices such that any two vertices are adjacent). Since the smallest eigenvalue increases after 2011, this confirms an increase in the level of connectivity and a change in the topology of the network which ex-

4.1 Rigs and Oil Market

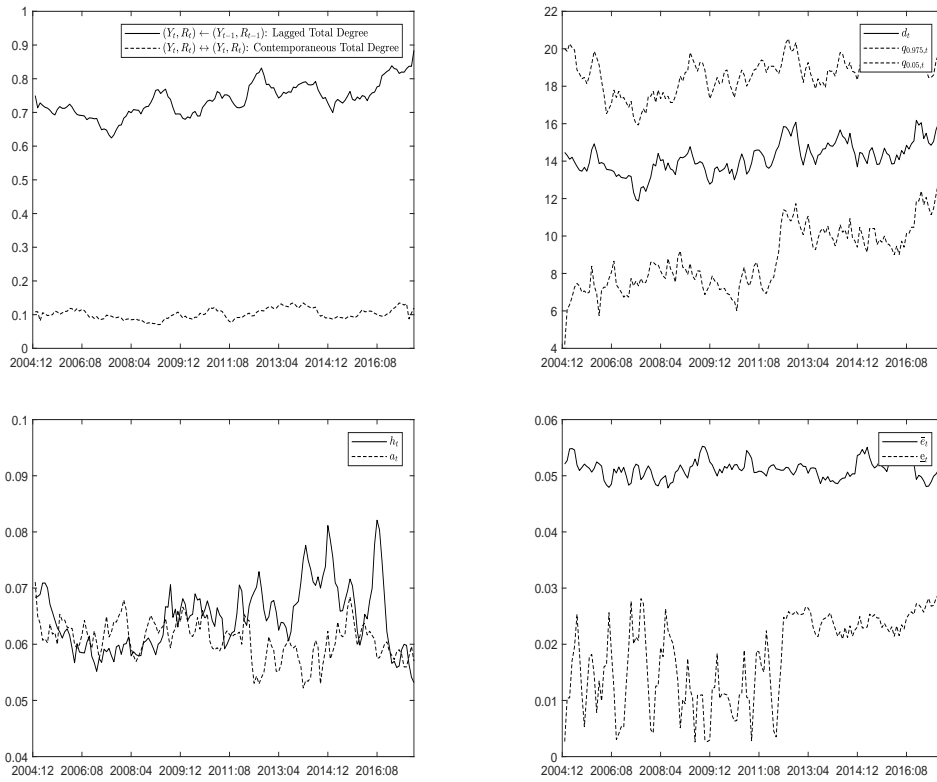


Figure 4: Top-left: Lagged (black lines) and contemporaneous (red line) connectedness over time for all regions/countries and variables. Top-right: mean (solid) and quantiles (dashed) of the degree distribution. Bottom-left: percentage of hubs (solid) that are nodes with successors of high authority scores, and percentage of authorities (dashed) that are nodes with predecessors of high hub scores. Bottom-right: the maximum (solid) and minimum (dashed) eigenvalues.

hibits larger groups of regions/countries highly connected between them. Again, this aligns with the increased correlation observed between geopolitical actors due to the Arab Spring events. The largest eigenvalue provides a lower bound on the maximum node degree, and in the application is quite stable over time.

Building on the notation in Equation 8, and expanding it to apply jointly to rigs and production, the connectedness measures are defined by

$$d_{it}^+ = \sum_{j \in V^k} A_{i,j,t}, \quad d_{it}^- = \sum_{j \in V^k} A_{j,i,t}, \quad d_t = \sum_{i,j \in V^k} (A_{i,j,t} + A_{j,i,t}) \quad (9)$$

while the total intra- and inter-connectedness for rigs, production, and across both rigs and production is measured by

$$d_t^Y = \sum_{i,j \in V^Y} (A_{i,j,t} + A_{j,i,t}) \quad (10)$$

$$d_t^R = \sum_{i,j \in V^R} (A_{i,j,t} + A_{j,i,t}) \quad (11)$$

$$d_t^{RY} = \sum_{i \in V^R, j \in V^Y} (A_{i,j,t} + A_{j,i,t}) + \sum_{i \in V^Y, j \in V^R} (A_{i,j,t} + A_{j,i,t}) \quad (12)$$

where V^Y is the set of oil producers and V^R is the set of rigs producers, with $V = V^R \cup V^Y$. The last measure, d_t^{RY} accounts for the cross-factor connectedness, which is a proposed new measure of oil connectedness at a micro-level. The normalized measures are reported as percentages of the maximum possible connectedness in Fig. 5. The level of connectedness is relatively stable over time. The cross-factor connectedness is very relevant, as shown in the second panel in Fig. 5, where approximately 35% of cross-factor lagged linkages are present at any point in time. In addition to this, approximately 5% of contemporaneous cross-factor linkages are present at any point in time. This combined cross-factor connectedness is evidence of the need of cross-factor modelling including both rigs and oil production to fully account for the oil relationships, as proposed initially in a static fashion in Espinasa et al. (2017). The top panel in Fig. 5 shows the degree of factor-specific autoregressive connectedness across the regions and countries. The peak in oil production autoregressive connectedness occurred around 2008, which aligns with the midst of the economic crisis. However, rigs have a very different dynamic, with higher autoregressive connectedness observed in the latter part of the sample. This disconnect between rig and production autoregressive behavior is also indicative of the need for joint modelling of these two key factors in the supply chain. The information content for each factor is relevant and not necessarily found in the other factor.

Additionally, from the decomposition of lagged connectedness in oil and rig connectedness (see Fig. 5) we find that: (1) The level of rig and oil connectedness measures (top panel; combined circles and solid-circles lines) increased between 2004 and 2017, nevertheless after 2011 the rig market connectedness is contributing most to the overall connectedness (more

4.1 Rigs and Oil Market

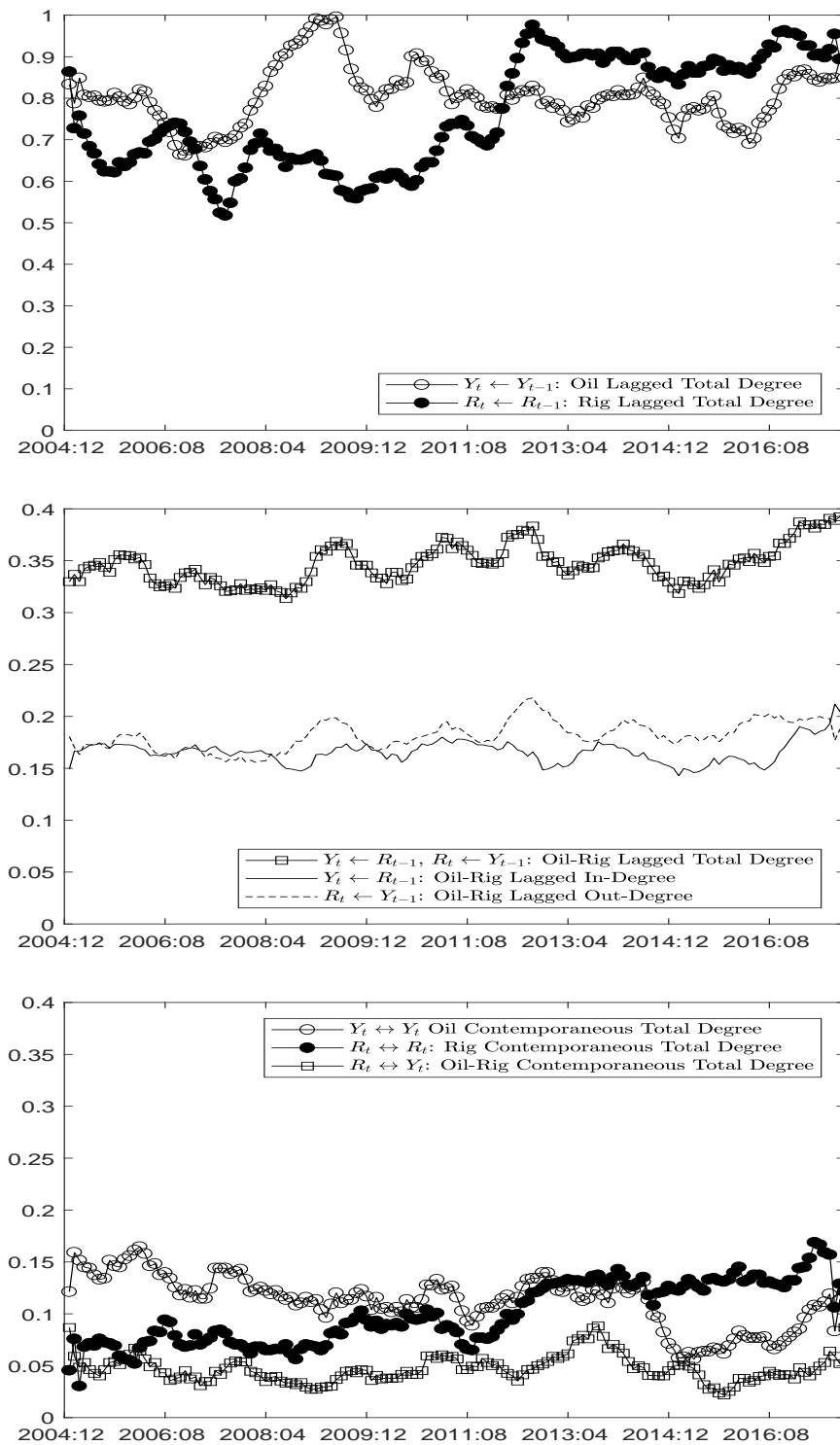


Figure 5: Lagged inter- and intra-connectedness over time (top panel). In- and out- inter-connectedness over time (middle panel). Contemporaneous inter- and intra-connectedness over time (bottom panel).

than the oil connectedness); (2) When decomposing the inter-connectedness between in- and out-degree (middle panel; solid and dashed lines), the out-degree is became slightly more relevant after January 2009. This suggests that oil production drove rig activity/decisions (in the lag) during the recovery years. In the last part of the sample (beginning of 2016) the in-degree has become more relevant, indicating that lagged rig activity decisions are affecting more the oil production changes (as producers reached maximum production capacity); and (3) From the decomposition of contemporaneous connectedness (bottom panel) an increase in connectedness is observable between rig activities across regions (solid circle), as well as a decrease in the connectedness between oil production across regions (circle). This again shows the relevance of the proposed approach, as oil production inter-connectedness may fade, as rig activity appears to take the lead driving the network connectedness.

Fig. 6 portrays across all panels the lagged connectedness levels between the economic activity index (DT) and the rest of the network across all regions and countries. There is a strong upward trend in the level of connectedness between both oil production and rig activity and economic activity, indicating that the cyclical nature of the oil markets has increased over time. This finding is of relevance in global hedging strategies.

Evidence of sharp reversal episodes can also be found in the top panel in Fig. 6. The connectedness levels between economic activity and both oil extraction and rigs oscillates significantly, showing sharp changes over time: (1) During the periods 2004-2007 and 2011-2013, the economic activity is more interconnected with rig activity than with oil production (top panel; solid circle line sitting above the circle line) ; (2) During the periods 2007-2011 and 2013-2015 the economic activity is more interconnected with oil production (top, circle line above solid circle line); and (3) After 2015 both oil extraction and oil production activities are equally interconnected with the economic activity, with a sharp increasing slope. These different level of connectedness with the economic activity index indicate that the global economy is a relevant variable (the levels are always relatively high for a single driving factor), but also that other factors may be missing, some of them of difficult quantification, such as supply shocks (new field explorations, drill authorizations, etc.), geopolitical (arab spring, political elections, etc.), or technological (spread of fracking, technological advances, etc.) factors, which can have evolving relative relevance over time.

4.1 Rigs and Oil Market

The middle and bottom panels in Fig. 6 show multiple episodes of reversal causation between economic activity and oil production and extraction. Though the general shape of the connectedness is not too dissimilar to that observed in the top panel, some additional effects can be observed in the bottom panels, where the overall levels (and decomposition) of the lagged connectedness between DT and each of rigs and oil production shows that the connectedness dynamics are very different over time, with periods where the economic activity is leading the overall connectedness (solid line defines the shape of the circle line), and others where the economic activity lags (dashed line driving the dynamics of the circle line). For example, during the 2008-2010 period, the economic activity is a key driver, with a significant spike of its impact on oil production. However rig decisions show less stable patterns around key periods, indicating a more stable relationship with the economic activity (decisions regarding rigs would be less driven by shorter-term factors, and more driven by longer-term prospects of demand than decisions about oil production, which are more shorter-term driven).

Fig. 7 contains the same analysis presented in 6, but this time for the price index (PT) as the central variable, rather than the economic activity index. The relevance of the price index is, again, time-varying, but with a strong level of connectedness with both oil production and rigs (both around 40%). This indicates that it is a relevant factor as well in the composition of the network, both in a contemporaneous and in a lagged fashion, as shown in Espinasa et al. (2017). The largest oscillation measured appears to affect the lagged relationship between prices and rig decisions, where a higher level of connectedness is observed around 2013. This coincides with the stabilization of the Arabian Peninsula region, leading to more fundamental factors (rather than geopolitical factors) being central to the decision-making around rigs. This short-lived sharp increase eventually stabilizes around long-term levels again around 2015.

Some additional graphical outcomes can be obtained from the proposed approach: (1) Fig. 8 portrays the instantaneous linkages across the full network (top left panel), and decomposes those into the cross-factor network (top-right panel; contemporaneous linkages between rigs and production), the rig-only contemporaneous network (bottom left panel) and the production-only contemporaneous network (bottom right panel; similar to the outputs

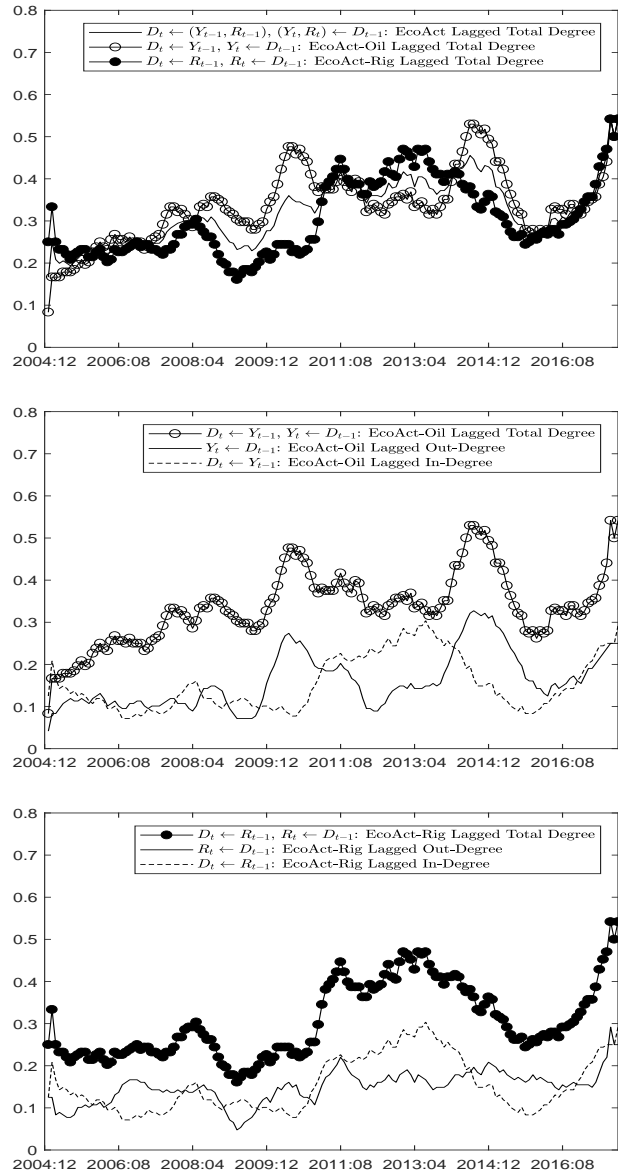


Figure 6: Lagged connectedness over time between economic activity and both rigs (top and bottom panel) and production (top two panels), both synchronous (top panel) and lagged (bottom two panels).

4.1 Rigs and Oil Market

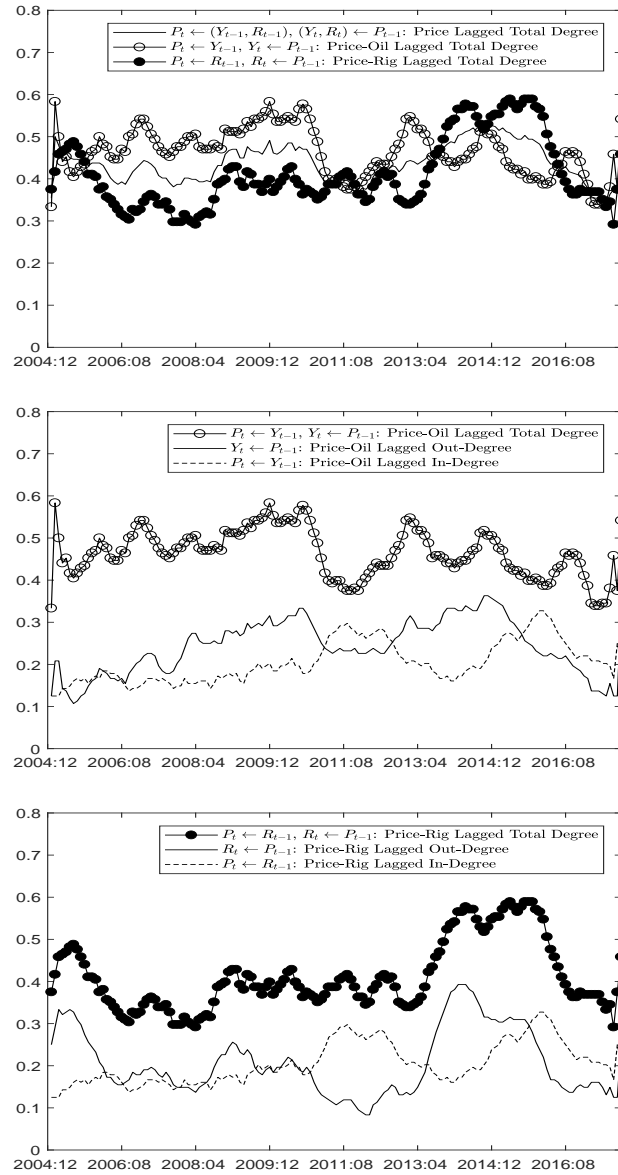


Figure 7: Lagged connectedness over time between oil prices and both rigs (top and bottom panel) and production (top two panels), both synchronous (top panel) and lagged (bottom two panels).

from Rousan et al. (2018)). The regional connectedness of rig decisions is very strong, with the United States producers clustered together, and, with the exception of UAE, the remaining producers also clustered together. The UAE again shows stronger linkages with the United States than other countries; (2) Fig. 9 shows, in the same order of panels, the lagged relationships. It can be seen that the networks are denser, showing that lagged linkages are more prevalent than contemporaneous ones, both in single-factor and cross-factor subnetworks; and (3) Figs. 10-12 show the lagged network over time, at 16 different, equally spaced points in the sample, with 10 months between each one. Though dense in nature, these graphs allow for identification of dynamics over time that may be occurring intra- or inter-factor.

5. CONCLUSION

This paper extends the work of Espinasa et al. (2017) and Rousan et al. (2018), and proposes a unified approach that combines the micro-level information described in the former reference with the network approach formalized in the latter reference. A Bayesian Graphical VAR model is proposed, which avoids some shortcomings of the conditional or pairwise Granger-causality models, which can lead to spurious linkages and to overestimation of the level of connectedness in network and graphical models. A full network is constructed to assess not only the intra-factor relationships, but also the cross-factor relationships that drive oil markets. An illustrating example includes rigs and oil production of two key interconnected producers, namely the United States (across multiple regions) and key middle eastern producers. Additionally, a world economic index and an oil price index are introduced to moderate relationships, since they have been shown to be relevant in the micro-level relationships across factors (Espinasa et al., 2017). Some key findings identified in this manuscript include: (1) Several central regions are identified in the oil network. They are relevant both in terms of oil production and rig activity, and include 3 regions in the United States and the UAE; (2) The overall relevance of the rig level as connecting (conditioning) vertices between the global factors (oil prices and the economic index) and the oil production. This

4.1 Rigs and Oil Market

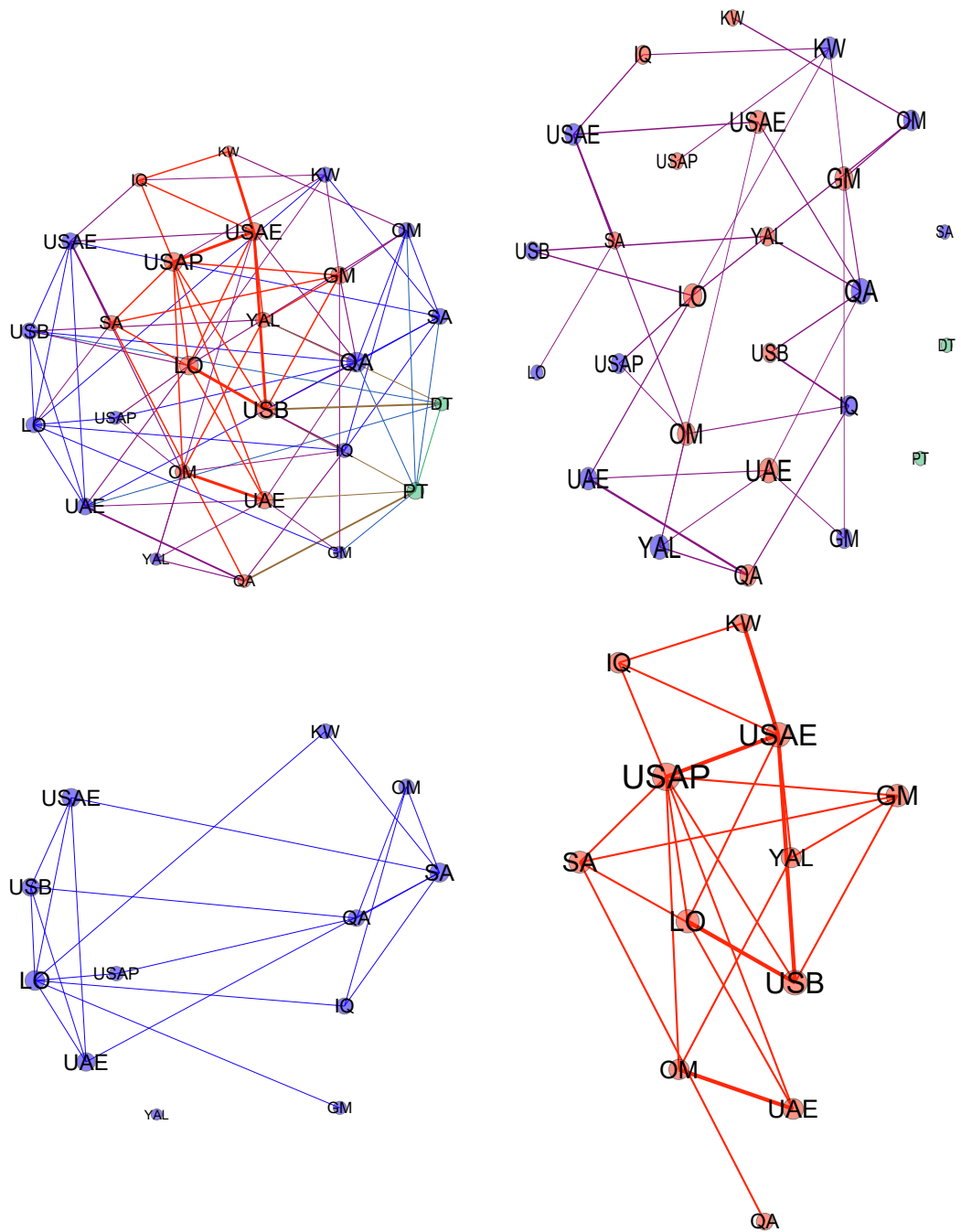


Figure 8: Contemporaneous linkages. Full network (top left). Production-rigs network (top right), rig layer network (bottom left) and production layer network (bottom right).

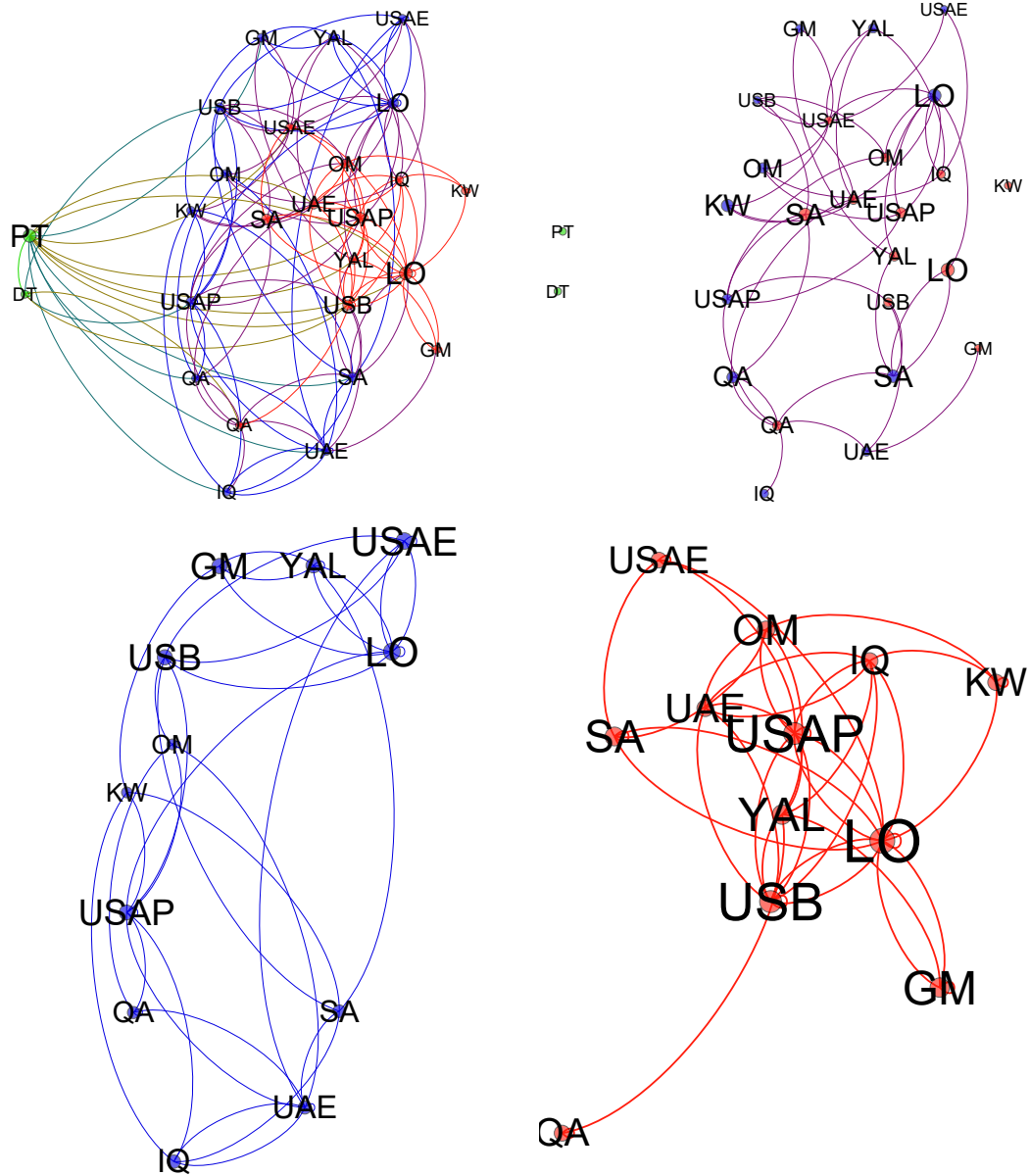


Figure 9: Lagged linkages. Oil market network (top left). Production-rigs network (top right), production layer network (bottom right) and rigs layer network (bottom left).

4.1 Rigs and Oil Market

shows that excluding the rigs from the network effectively removes what would become a key latent variable driving the full network dynamics; (3) The relevance of Saudi Arabia as a major actor may be overstated in terms of its connectedness, with other regions or countries showing a higher relevance in the network; (4) Cross-factor relationships are highly relevant. Rigs and production show high levels of connectivity, which, again, aligns with the proposed approach and conclusions in (Espinasa et al., 2017); (5) Some changes in the topology of the network can be found after 2011, coinciding with geopolitical events in the middle eastern countries within the sample; and (6) The cyclical nature of oil markets has increased over time, as shown in the upward trend of the connectedness between both oil production and rig activity with the economic activity index.

The approach proposed in this paper can serve as building block and unifying point between static, micro-level models, which expand the number of factors at the expense of imposing static linkages or single-country analyses, and macro-level models, which allow for dynamic linkages but are limited in focus to single factors.

REFERENCES

- Acemoglu D, Carvalho VM, Ozdaglar A, Tahbaz-Salehi A. 2012. The network origins of aggregate fluctuations. *Econometrica* **80**: 1977–2016.
- Ahelegbey D, Billio M, Casarin R. 2016a. Bayesian graphical models for structural vector autoregressive processes. *Journal of Applied Econometrics* **31**: 357–386.
- Ahelegbey D, Billio M, Casarin R. 2016b. Sparse graphical multivariate autoregression: A bayesian approach. *Annals of Economics and Statistics* **123/124**: 1–30.
- Barabási AL, Albert R. 1999. Emergence of scaling in random networks. *Science* **286**: 509–512.
- Battiston S, Gatti DD, Gallegati M, Greenwald B, Stiglitz J. 2012. Liaisons dangereuses: Increasing connectivity, risk sharing, and systemic risk. *Journal of Economic Dynamics and Control* **38(8)**: 1121–1141.
- Bernard AB, Moxnes A. 2018. Networks and trade. Working paper.
- Bianchi D, Billio M, Casarin R, Guidolin M. 2018. Modelling systemic risk with markov switching graphical sur models. *Journal of Econometrics* Forthcoming.
- Billio M, Casarin R, Costola M, Pasqualini A. 2016. An entropy-based early warning indicator for systemic risk. *Journal of International Financial Markets, Institutions and Money* **45**: 42–59.
- Billio M, Getmansky M, Lo A, Pellizzon L. 2012. Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of Financial Economics* **104**: 535–559.
- Bollobas B. 2001. *Random Graphs*. New York: Cambridge University Press.
- Bollobas B. 2002. *Modern Graph Theory*. New York: Springer-Verlag.
- Bonacich P. 2007. Some unique properties of eigenvector centrality. *Social Networks* **4**: 555–564.

- Brillinger DR. 1996. Remarks concerning graphical models for time series and point processes. *Revista de Econometria* **16**: 1–23.
- Broner F, Ventura J. 2016. Rethinking the effects of financial globalization. *The quarterly journal of economics* **131**: 1497–1542.
- Brunnermeier M, Pedersen L. 2009. Market liquidity and funding liquidity. *Review of Financial Studies* **22(6)**: 2201–2238.
- Carvalho C, West M. 2007. Dynamic matrix-variate graphical models. *Bayesian Analysis* **2**: 69–98.
- Corander J, Villani M. 2006. A Bayesian approach to modelling graphical vector autoregressions. *Journal of Time Series Analysis* **27(1)**: 141–156.
- Dawid A, Lauritzen S. 1993. Hyper-markov laws in the statistical analysis of decomposable graphical models. *The Annals of Statistics* **21**: 1272–1317.
- Diebold F, Yilmaz K. 2014. On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics* **182(1)**: 119–134.
- Dunbar R, Arnaboldi V, Conti M, Passarella A. 2015. The structure of online social networks mirrors those in the offline world. *Social Networks* **43**: 39–47.
- Economou A, Agnolucci P, Fattouh B, de Lipis V. 2017. *A structural Model of the World Oil Market*. The Oxford Institute for Energy Studies.
- Espinasa R, Manzano O, ter Horst E, Molina G, Rigobon R. 2017. A micro-based model for world oil market. *Energy Economics* **66**: 431–449.
- Goldenberg A, Zheng AX, Feinberg SE, Airoidi EM. 2010. *A Survey of Statistical Network Models (Foundations and Trends(r) in Machine Learning)*. Now Publishers Inc.
- Hautsch N, Schaumburg J, Schienle M. 2014. Financial network systemic risk contributions. *Review of Finance* .
- Kilian L. 2009. Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. *American Economic Review* **99(3)**: 1053–1069.

- Lauritzen S. 1996. *Graphical Models*. Clarendon Press, Oxford.
- Maluck J, Donner RV. 2015. A network of networks perspective on global trade. *PLoS One* **10**: e0133310.
- Puliga M, Flori A, Pappalardo G, Chessa A, Pammolli F. 2016. The accounting network: How financial institutions react to systemic crisis. *PLoS One* **11**: e0162855.
- Rodriguez A, Lenkoski A, Dobra A. 2011. Sparse covariance estimation in heterogeneous samples. *Electronic Journal of Statistics* **5**: 981.
- Rousan SA, Sbia R, Tas BKO. 2018. A dynamic network analysis of the world oil market: Analysis of opec and non-opec members. *Energy Economics* **75**: 28–41.
- Wang H. 2010. Sparse seemingly unrelated regression modelling: Applications in finance and econometrics. *Computational Statistics and Data Analysis* **54**: 2866–2877.
- Wang H, Reeson C, Carvalho CM. 2011. Dynamic financial index models: Modeling conditional dependencies via graphs. *Bayesian Analysis* **6**: 639–664.
- Wang H, West M. 2009. Bayesian analysis of matrix normal graphical models. *Biometrika* **96**: 821–834.
- Whittaker J. 1990. *Graphical Models in Applied Multivariate Statistics*. Chichester, UK: John Wiley and Sons.

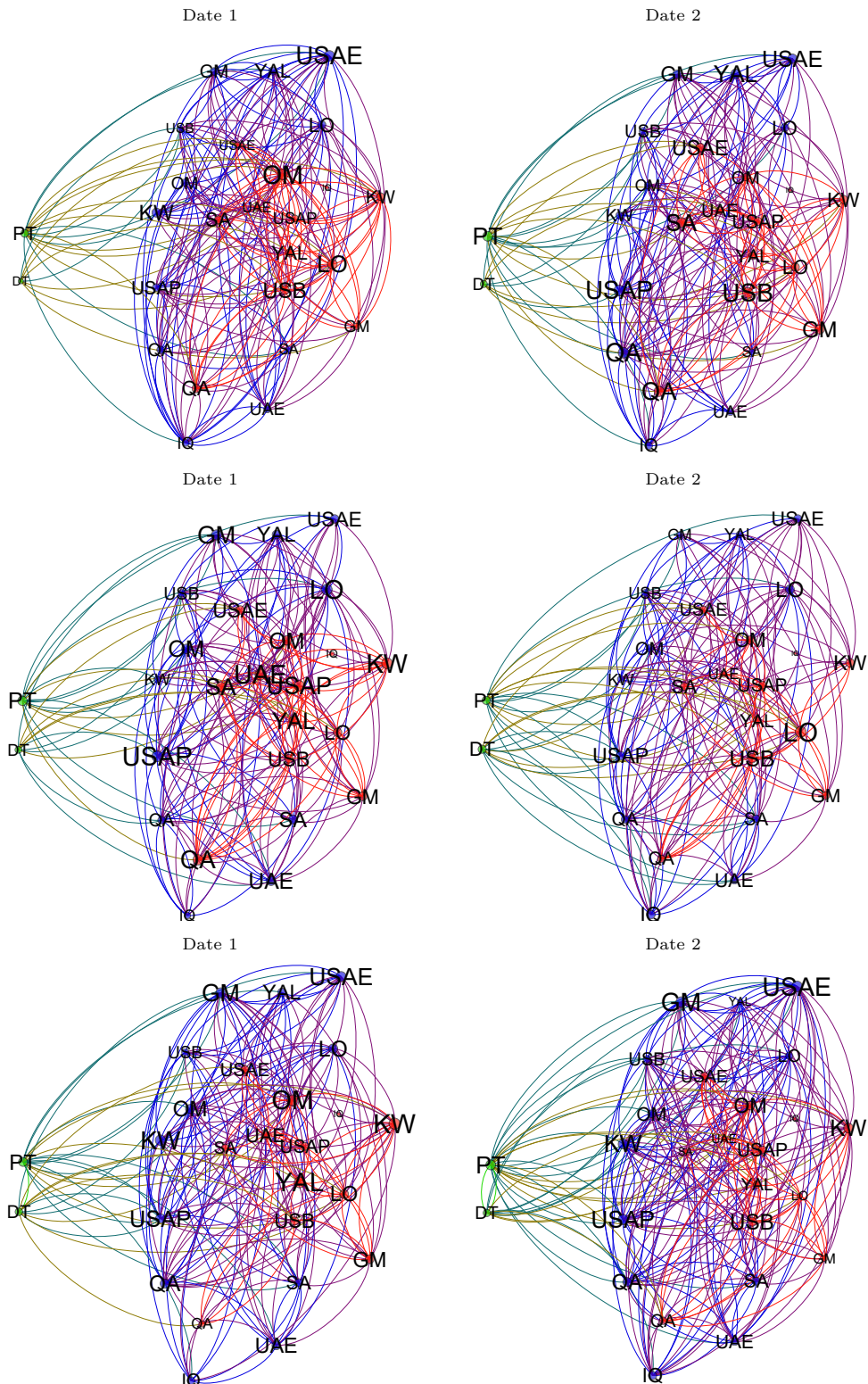


Figure 10: Lagged linkages over time (every 10 months). Measurements 1-6.

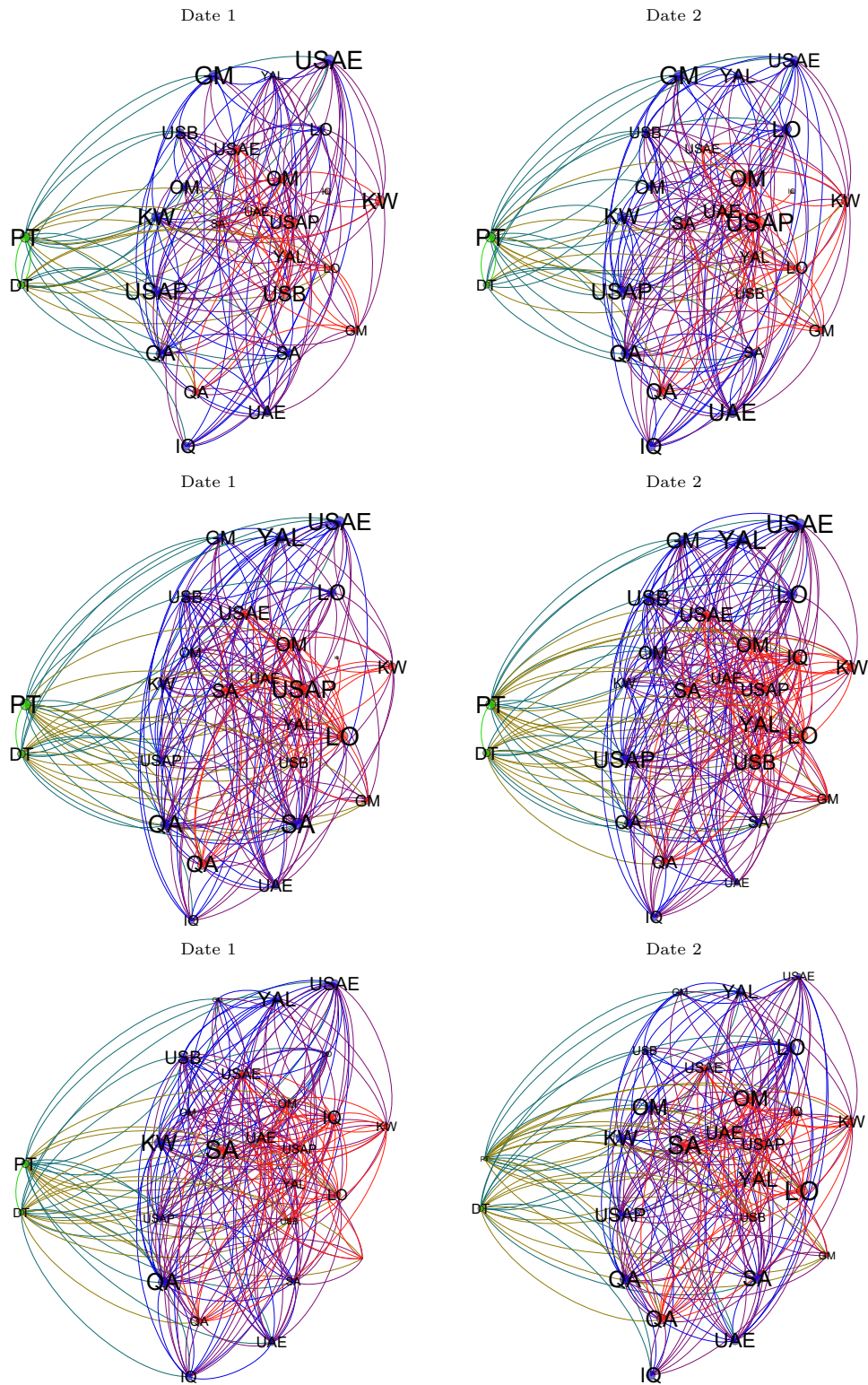


Figure 11: Lagged linkages over time (every 10 months). Measurements 7-12.

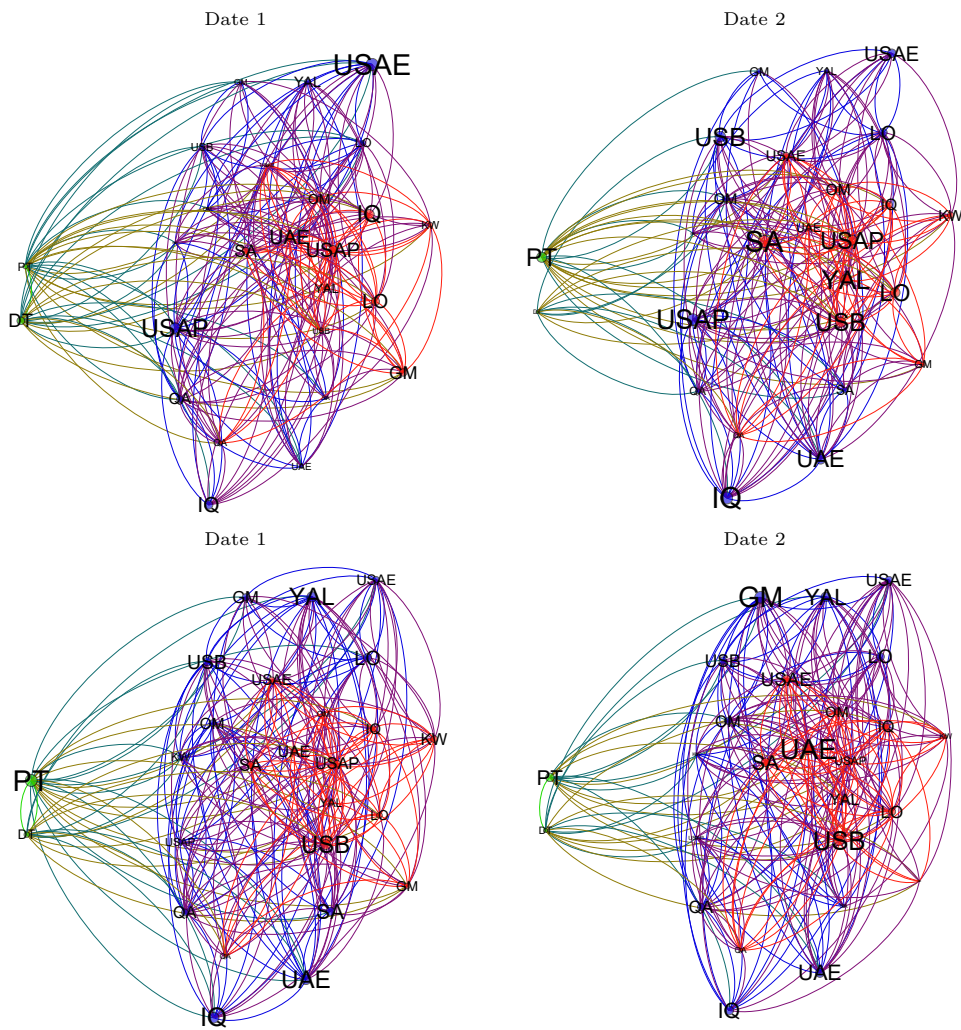


Figure 12: Lagged linkages over time (every 10 months). Measurements 13-16.