

OTDK-dolgozat

2025

NEUMANN JÁNOS EGYETEM

SZEKTORÁLIS KÍNÁLATI SOKKOK IDENTIFIKÁCIÓJA
NAGY BAYESI VAR MODELLEKKEL

IDENTIFICATION OF SECTORAL SUPPLY SHOCKS:
A LARGE BAYESIAN VAR APPROACH

Szentistványi István János

Konzulens: Gánics Gergely

Kézirat lezárásának dátuma: 2024.04.15.

Abstract

This paper employs a large Bayesian vector autoregression framework to identify sectoral supply shocks using US producer price indices at the industry level. In this approach, all sectoral variables are directly included in the model to provide a more systemic view on the interdependence of sectoral and aggregate dynamics. The identification restrictions take the form of ranking restrictions based on input-output tables, which are constructed to treat import contributions as a separate sector. Four sectoral shocks are identified and their impact on various aggregate variables is assessed. The results show that some of the identified sectoral shocks, such as the ones related to Machinery and Computer & electronic products industries, have a significant impact on aggregate industrial production and unemployment, while Import supply shocks have a significant impact on the consumer price index. The results confirm the findings of structural macro models that sectoral fluctuations have aggregate effects by propagation through the production network. The robustness of the results is confirmed by testing different model specifications.

Contents

1	Introduction	1
2	Methods	2
2.1	The reduced form VAR(p) model	2
2.2	Estimation	3
2.3	Impulse response functions	3
3	Identification	5
3.1	Sectoral supply and demand shocks	5
3.2	Ranking restrictions	6
3.3	Identification algorithm	7
4	Data	8
4.1	Input-output tables	8
4.2	Constructing the requirements tables	9
4.3	Ranking restrictions in input-output tables	10
4.4	Sectoral and aggregate time series	12
5	Results	13
5.1	Main results from smaller models ($n = 20$)	14
5.2	Comparison with large models ($n = 50$)	16
5.3	Choice of the restriction horizon	18
5.4	Impact of sectoral supply shocks on the CPI	19
6	Conclusions	22
	References	23
A	Appendix	27
A.1	Key equations of the VAR(p) system	27
A.2	Derivation of the natural conjugate prior	28
A.3	Details on the identification algorithm	30
A.4	Computational properties	31
A.5	Further results	31
A.6	Data sources and IO-NAICS matching	33

List of Tables

1	Ranking restrictions on the 43×43 direct requirements matrix	11
2	Ranking restrictions on the 13×13 direct requirements matrix	11
A1	Computational properties of the identification	31
A2	Correspondence of BEA input-output and NAICS categories	34
A3	Summary of other time series	35

List of Figures

1	IRFs for VAR(3), $n = 20$, $h = 1$, Direct Requirements (1–2), Industrial production, Unemployment	15
2	IRFs for VAR(3), $n = 20$, $h = 4$, Total Requirements (1–2), Industrial production, Unemployment	16
3	IRFs for VAR(3), $n = 50$, $h = 1$, Direct Requirements (2–0), Industrial production, Unemployment	17
4	IRFs for VAR(3), $n = 50$, $h = 4$, Total Requirements (2–0), Industrial production, Unemployment	18
5	IRFs for VAR(3), $n = 20$, $h = 4$, Direct Requirements (1–2), Industrial production, Unemployment	19
6	IRFs for VAR(3), $n = 20$, $h = 1$, Total Requirements (1–2), Industrial production, Unemployment	20
7	IRFs for VAR(3), $n = 20$, $h = 1$, Direct Requirements (1–2), CPI	21
8	IRFs for VAR(3), $n = 20$, $h = 4$, Total Requirements (1–2), CPI	21
A1	IRFs for VAR(3), $n = 15$, $h = 1$, Direct Requirements (1–2), Industrial production, Unemployment	32
A2	IRFs for VAR(3), $n = 45$, $h = 1$, Direct Requirements (2–0), Industrial production, Unemployment	33

1 Introduction

In recent years, increasing attention has been paid to the effects of sector-specific shocks on aggregate fluctuations, as suggested for example by [Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi \(2012\)](#). In this paper, we identify several sectoral supply shocks jointly in a large Bayesian vector autoregressive (VAR) model to investigate the aggregate effects of sectoral shocks. Large Bayesian VAR models have long been used for forecasting since [Bańbura, Giannone, and Reichlin \(2010\)](#), and more recently for identifying multiple *aggregate* shocks through the novel method of [Chan, Matthes, and Yu \(2023\)](#). To our knowledge, the only non-structural econometric model that incorporates production networks and the identification of sectoral shocks is the factor-augmented VAR (FAVAR) approach of [De Graeve and Schneider \(2023\)](#), which uses a partial identification strategy and employs the FAVAR method of [Bernanke, Boivin, and Elias \(2005\)](#).

In this paper, we argue that the identification of sectoral shocks using a large Bayesian VAR is possible without resorting to dimension-reducing methods such as the FAVAR approach of [De Graeve and Schneider \(2023\)](#). In this way, all sectoral variables are directly included in the model and it is possible to identify multiple sectoral shocks jointly, which enables a more systemic view on the interdependence of sectoral and aggregate dynamics. This extended partial identification scheme also leads to a more granular restriction set, which helps us to separate a sectoral shock from the aggregate fluctuations of the system more efficiently. It may be extended to a full identification in smaller models where all possible sectoral and aggregate shocks are accounted for.

The structural identification we use in this paper takes the form of ranking restrictions based on input-output tables, similar to [De Graeve and Schneider \(2023\)](#). We use domestic direct and total requirements tables, while also accounting for import contributions separately. The model is constructed using sectoral producer price index data and identified by the ranking of rows in the input-output table, corresponding to downstream supply shocks. The importance of production networks is also recognized in some structural general equilibrium models, for example [Ghassibe \(2021b\)](#), who suggests that a large part of the aggregate impact of monetary policy is related to the spillover effects through input-output linkages. Our approach fits well with macro models dealing with production networks, as it could be used to find econometric evidence for sectoral phenomena such as multiplier and network spillover effects based only on agnostic assumptions.

Results show that sectoral shocks – when identified by ranking restrictions from input-output tables – have aggregate effects on key macro variables such as total industrial production, unemployment and the consumer price index (CPI). In this paper, we partially identify 4 sectoral shocks and show that, for example, the machinery and computer parts industry has a considerable negative impact on industrial production in most model specifications and a positive impact on unemployment only in some specifications. The impact of these shocks on the aggregate CPI is less significant, and most models show that only import supply shocks have a direct and visible contribution to the increase in CPI. The choice of the restriction horizon affects the dynamics and the shape of the impulse response functions (IRFs), but has no impact on our main results. After controlling for the choice of restriction horizon, there is no significant difference when using direct or total requirements input-output tables.

The remainder of the paper is organized as follows. Section 2 briefly summarizes the basic methods for estimating a large Bayesian VAR using the natural conjugate prior setup, with further details provided in the Appendix. Section 3 discusses the issue of identification using ranking restrictions and, in particular, input-output tables. Section 4 summarizes the data used in this paper, including the derivation of input-output requirements tables that include a separate entry for imports. Section 5 presents the results of the analysis in the form of impulse-response functions of the identified sectoral shocks. Section 6 concludes and includes avenues for further research.

2 Methods

2.1 The reduced form VAR(p) model

We consider a standard homoskedastic VAR of order p . Following the notation used by [Chan \(2020\)](#), $\mathbf{y}_t = (y_{1t}, \dots, y_{nt})'$ is the $n \times 1$ vector of dependent variables at time t . Then, the VAR(p) equation is given by:

$$\mathbf{y}_t = \mathbf{b} + \mathbf{A}_1 \mathbf{y}_{t-1} + \dots + \mathbf{A}_p \mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t, \quad (1)$$

where $\mathbf{b}$ is an $n \times 1$ vector of intercepts, $\mathbf{A}_1, \dots, \mathbf{A}_p$ are $n \times n$ coefficient matrices and $\boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$. The system of variables is described by n equations, where each equation has $k = np + 1$ regression coefficients, resulting in a total of $nk = n(np + 1)$ coefficients.

2.2 Estimation

In their paper on large Bayesian VARs, Bańbura et al. (2010) discuss the importance of shrinkage priors, which help eliminate the problems due to the high number of parameters: $T \ll n(np+1)$, where T is the number of observations. Later, Chan (2020) surveys three common approaches to formulating priors in large Bayesian VARs, providing the derivation of the posterior and computational properties. Alternatives range from the standard Minnesota prior of Litterman (1986) to the natural conjugate and the independent normal and inverse-Wishart priors, which are each less restrictive on the structure of the estimated covariance matrix Σ . The choice depends on computational constraints, as less restrictive setups are computationally more expensive. These prior setups are similar in shrinking elements of $\mathbf{A}_1, \dots, \mathbf{A}_p$ to 1 on the diagonal of the first lag, and to zero otherwise. Modified constructions shrink coefficients to 1 only in case of random walk variables, and to zero in case of mean-reverting variables. Also, the covariance matrix of the coefficients is shrunk in a way that the variance decreases at higher lags, and the rate of decrease is higher for lags of other variables. See the details on prior hyperparameter selection in Section A.2 in the Appendix.

In this paper, we choose to implement the natural conjugate prior, precisely because of the computational burdens of the independent normal and inverse-Wishart prior in large VARs. While the natural conjugate prior solution imposes a relatively restrictive structure on the covariance matrix of β coefficients, its computational tractability makes it an efficient choice in large models, while it is also more flexible than the standard Minnesota prior. In this case, the posterior distribution of parameters is also a joint normal-inverse-Wishart distribution, so the parameter values can be sampled directly, without the need to use the Gibbs-sampler algorithm. Based on the prior density and the likelihood function, parameters of the posterior need to be calculated only once. Then, sampling from the distribution is computationally feasible even with large models.

2.3 Impulse response functions

Impulse response functions can be derived using the infinite order Vector Moving Average representation of the VAR(p) process. According to the Wold-theorem, such VMA(∞) representation exists when the VAR is stationary – where stability implies stationarity. If all eigenvalues of the companion form coefficient matrix $\mathbf{F}$ are smaller than 1 in absolute

value, the VAR is stable. In order to derive impulse-response functions, first we consider the following VAR(1) process: $\mathbf{y}_t = \mathbf{b} + \mathbf{A}\mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t$. In this case, the Wold-coefficients $\mathbf{C}(L)$ are simply the powers of the $\mathbf{A}$ coefficient matrix. That is,

$$\begin{aligned} C_0 &= \mathbf{I}_n \\ C_1 &= \mathbf{A} \\ C_2 &= \mathbf{A}^2 \\ &\vdots \end{aligned}$$

This can be derived easily, as we only need to find a $\mathbf{C}(L)$ matrix such that $\mathbf{y}_t = \mathbf{b} + \mathbf{A}\mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t$ can be written as $\mathbf{y}_t = \mathbf{C}(L)\mathbf{b} + \mathbf{C}(L)\boldsymbol{\varepsilon}_t = \boldsymbol{\mu} + \mathbf{C}(L)\boldsymbol{\varepsilon}_t$. The coefficients $\mathbf{C}(L)$ can be calculated as shown in Equations 2–4:

$$\mathbf{y}_t = \mathbf{b} + \mathbf{A}\mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t, \quad (2)$$

$$(\mathbf{I} - \mathbf{A})\mathbf{y}_t = \mathbf{b} + \boldsymbol{\varepsilon}_t, \quad (3)$$

$$\mathbf{C}(L) = (\mathbf{I} - \mathbf{A})^{-1} \quad (4)$$

simply using the identity $(\mathbf{I} - \mathbf{A})(\mathbf{I} - \mathbf{A})^{-1} = \mathbf{I}$. This is possible as the stability condition implies exactly that $(\mathbf{I} - \mathbf{A})$ is invertible. To generalize the above process, we can rewrite the VAR(p) process as a VAR(1) using the companion form, which is constructed in the following way (the size of each matrix is displayed next to each element):

$$\mathbf{Z}_t^{np \times 1} = \nu^{np \times 1} + \mathbf{F}^{np \times np} \mathbf{Z}_{t-1}^{np \times 1} + \mathbf{E}_t^{np \times 1} \quad (5)$$

$$\mathbf{Z}_t = \begin{bmatrix} \mathbf{y}_t \\ \mathbf{y}_{t-1} \\ \vdots \\ \mathbf{y}_{t-p+1} \end{bmatrix}, \quad \nu = \begin{bmatrix} \mathbf{b} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \dots & \mathbf{A}_{p-1} & \mathbf{A}_p \\ \mathbf{I}_n & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_n & & \mathbf{0} & \mathbf{0} \\ \vdots & & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{I}_n & \mathbf{0} \end{bmatrix}, \quad \mathbf{E}_t = \begin{bmatrix} \boldsymbol{\varepsilon}_t \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix} \quad (6)$$

The coefficient matrix of this transformed VAR(p) is the $\mathbf{F}$ from the companion form, which means that the IRF matrices (the Wold-coefficients) for each given horizon are the corresponding powers of $\mathbf{F}$, from which we only take the upper left $n \times n$ sub-matrix.

3 Identification

3.1 Sectoral supply and demand shocks

When a distinction is made between supply and demand shocks, the identification algorithm needs to take into account the differences between the two, usually by controlling for price and quantity changes simultaneously with specific sign-restrictions, see the conventional solutions in the structural VAR literature in [Canova and Nicoló \(2002\)](#) for example. In order to identify negative supply shocks in this setup, we would impose restrictions that given a positive price shock, the eventual impact on production levels should be negative.

In the present identification algorithm, it is infeasible to include both price and quantity variables for each sector, due to the size of the model. However, as [Acemoglu, Akcigit, and Kerr \(2016\)](#) argue, whether sectoral shocks are supply or demand-induced can be inferred from the observed propagation mechanisms to other industries, which are determined by the structure of the input-output tables. They state that demand shocks are mostly characterized by upstream propagation, impacting the industries that provide the inputs for the origin sector. On the other hand, sectoral supply shocks are characterized by a downstream propagation mechanism and they mostly impact the industries which use the outputs of the origin sector as their input.

In the IO (input-output) tables, this is given by the structure of columns and rows respectively: for example, to observe the downstream propagation mechanisms of a specific industry, we can see from the elements of that row which other industries rely mostly on the output of the given sector. Note, that [De Graeve and Schneider \(2023\)](#) also use this same concept and method for separately identifying supply and demand related shocks from ranking the rows or columns on the input-output tables respectively. However, in their analysis, they choose to discuss both types of shocks simultaneously for each specification. In this paper, we solely focus on identifying sectoral supply shocks.

If we accept the above assumption, then we can distinguish supply and demand shocks without relying on price and quantity data simultaneously. This translates to our identification algorithm in the following way: impulse-response functions that describe the impact of a specific sectoral shock (a column of the IRF matrix) must follow the same ranking as the corresponding row in the IO table. Impulse-responses of a sectoral shock are also normalized so that the impact of the shock on the origin industry price index is

positive, which along with the structure that resembles the IO tables is characteristic of a *negative* supply shock.

3.2 Ranking restrictions

We follow an identification strategy using ranking restrictions, which are set-type restrictions on IRFs that most frequently impose that either 1) the impact of a shock is higher on variable i than on variable j , or 2) shock i has a higher impact on the same variable than shock j . They are similar to sign restrictions, and [Amir-Ahmadi and Drautzburg \(2021\)](#) provide a systematic review of ranking restrictions, discussing both theory and applications. Before that, [De Graeve and Karas \(2014\)](#) discuss restrictions based on heterogeneity in panel-VAR settings, where rankings between IRF values are imposed in the restriction set. In theory, the algorithm of [Rubio-Ramírez, Waggoner, and Zha \(2010\)](#) can be used to find admissible IRFs that fulfill ranking restrictions. We can use the SVAR representation,

$$\mathbf{B}_0^{-1}\mathbf{y}_t = \mathbf{B}_0^{-1}\mathbf{b} + \mathbf{B}_0^{-1}\mathbf{A}(L)\mathbf{y}_{t-1} + \mathbf{u}_t \quad (7)$$

where $\mathbf{u}_t = \mathbf{B}_0^{-1}\boldsymbol{\varepsilon}_t$ are orthogonal structural shocks, $\mathbf{u}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and $\mathbf{B}_0\mathbf{B}_0' = \boldsymbol{\Sigma}$. Let $\mathbf{L}$ be the lower triangular matrix from the Cholesky decomposition of $\boldsymbol{\Sigma}$ such that $\mathbf{L}\mathbf{L}' = \boldsymbol{\Sigma}$ and $\mathbf{L} = \mathbf{B}_0$:

$$\mathbf{y}_t = \boldsymbol{\mu} + \mathbf{C}(L)\mathbf{L}\mathbf{u}_t. \quad (8)$$

Then, [Rubio-Ramírez et al. \(2010\)](#) use random draws of $\mathbf{Q}$ matrices to right-multiply $\mathbf{C}(L)\mathbf{L}$ by $\mathbf{Q}$. This can be done as for orthogonal matrices, $\mathbf{Q}\mathbf{Q}' = \mathbf{I}$ and so, $\mathbf{L}\mathbf{Q}\mathbf{Q}'\mathbf{L}' = \boldsymbol{\Sigma}$ also. The matrix $\mathbf{Q}$ is uniformly drawn from the set of $n \times n$ orthogonal matrices $\mathcal{O}(n)$ using the QR -decomposition of a random normal $n \times n$ matrix. $\mathbf{Q}$ is repeatedly drawn for each parameter draw $(\mathbf{A}, \boldsymbol{\Sigma})$ a given number of times, and for each draw, IRFs are checked against our ranking restrictions and the ones satisfying our restrictions are retained.

3.3 Identification algorithm

While ranking restrictions work similarly to sign restrictions, it seems that they are relatively more strict in general. For example, when applied to a single shock, the clustered or full ranking of IRFs of the shock on variables requires all of them to have the same sign – imposing a strict sign restriction by construction. Also, intuition suggests that for only one element of the IRF matrix, the imposed sign restriction is fulfilled more often by random draws than if that element is required to be the maximum in value. It is not surprising to see, that imposing a large number of sign restrictions when n is large is infeasible using the standard [Rubio-Ramírez et al. \(2010\)](#) algorithm, but this is also true for only a small number of ranking restrictions, as shown by [Chan et al. \(2023, p16.\)](#) for a few additional ranking restrictions. To solve this issue, [Chan et al. \(2023\)](#) suggest a modified algorithm, which makes use of the fact that any permutation of the orthogonal matrix $\mathbf{Q}$ can be used to represent the same structural shocks by relabeling columns and switching signs. They efficiently scan through all permutations by constructing an auxiliary $\mathbf{T}$ matrix using only trivial calculations. See Section A.3 in the Appendix for more details. This matrix is then used to determine whether the set of admissible IRFs for a given $\mathbf{Q}$ is empty or not. If not, they use the row and column indices of nonzero elements in this $\mathbf{T}$ matrix to draw uniformly from this set.

In identifying sectoral shocks using this algorithm, we come across some issues when following this identification strategy. In case of sign restrictions, it is easy to ensure that two shocks are distinct by ex-ante defining the restrictions such that 1) the two shocks have signed impact on at least two common variables and 2) that there is at least one variable for which the signed impact of the two shocks is different. This is also true for ranking restrictions when combined with other sign restrictions, as formally defined by [Chan et al. \(2023\)](#). However, a problem arises when we consider sectoral shocks identified solely by ranking restrictions. We cannot easily justify the combination of ranking and sign restrictions in the sectoral model, as we assume the same sign of impact from e.g. aggregate variables on all sectors. Then, it is not immediately clear whether two shocks are guaranteed to be distinct in case of only ranking restrictions.

It is possible that one column of the IRF matrix fulfills the ranking restrictions of a particular shock, and the negative of the same column fulfills the restrictions of another shock. This is visible from the $\mathbf{T}$ matrix (there are multiple non-zero elements in a

column) and a theoretically correct and conservative solution is to ex-post eliminate those $\mathbf{T}$ matrices that exhibit this pattern. See Eq. 9 for an example of this issue, corresponding to a possible $\mathbf{T}$ in a $n = 20$ model with 4 identified shocks. In this 4×20 matrix

$$\mathbf{T}^{4 \times 20} = \begin{bmatrix} 0 & 0 & 0 & \dots & 1 & 0 & -1 \\ 1 & 0 & 1 & \dots & 0 & 0 & 0 \\ 0 & 1 & -1 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 & 0 \end{bmatrix}, \quad (9)$$

we can see that each row has at least one non-zero element, meaning we were able to find columns in the IRF matrix that fulfill the restrictions for a given shock. However, $\mathbf{T}_{2,3} = 1$ and $\mathbf{T}_{3,3} = -1$ indicate that the third column (including its opposite) of the IRF matrix simultaneously fulfills the restrictions for two separate shocks. Such matrices must be discarded if we require the identified shocks to be distinct from each other.

4 Data

4.1 Input-output tables

Identification is based on US input-output tables, provided annually by the Bureau of Economic Analysis (BEA). We use the raw tables for the year 2019, consisting of 71 industries and 73 commodities. We use the 2019 tables because the COVID years starting from 2020 are not included in the estimation. The BEA provides both a domestic-only and imports-included version for most tables, but we choose to add a separate import column and row to the domestic-only table, separating the import channel of downstream supply shocks and treating imports as a distinct source of sectoral shocks. Therefore, the total requirements table needs to be calculated manually from the make, use and import contribution tables. The year in which the input-output tables are reported would not change our results significantly, as studies suggest that input-output coefficients are relatively stable over time. This stability extends to the distribution and magnitude of coefficients as well, and naturally implies that the position of non-zero elements and their relative importance is even more stable (Sawyer, 1992; Torres-González & Yang, 2019).

4.2 Constructing the requirements tables

Following the notations and method of BEA for the derivation of domestic direct and total requirements table, we use the following matrices and vectors. In the industry-by-commodity make matrix V a column shows for a given commodity the amount produced in each industry. In the commodity-by-industry use table U , the column shows for a given industry the amount of each commodity it uses. The import contribution matrix M is also commodity-by-industry and each column shows for a given industry the amount of each imported commodity. Let q be a column vector showing the total output of each commodity and g be a column vector showing the total output of each industry. Also, let U_d be the domestic use table, where $U_d = U - M$.

Then, the following identities apply, where i denotes a conformable column vector of ones, serving as a summation vector.

$$g = Vi, \quad (10)$$

$$q = U_d i + e_d, \quad (11)$$

where e_d is the vector of domestically produced final demand purchases for each commodity, given by $e_d = e - e_i$. Here, e_i is the vector of imported final demand purchases and e is the vector of total final demand purchases, including both imports and domestic production. Then, the commodity-by-industry domestic direct requirements matrix B_d and the market share matrix (transformation matrix between commodities and industries) D is given by:

$$B_d = U_d [\text{diag}(g)]^{-1}, \quad (12)$$

$$D = V [\text{diag}(q)]^{-1}. \quad (13)$$

We can add the rows and columns to both B_d and D corresponding to the import contributions using $m = iM$ such that the new import row of B_d consists of the vector m/g' and a 0 at the end, whereas the import column of B_d simply contains all zeros. Adding imports to the D matrix is even more straightforward, as there is no correspondence between import as a separate entity and the other industries and commodities. Therefore, we can add a row and column of zeros to D , except for the bottom right element, which is 1. The direct requirements matrix is $D_{imp}B_{imp}$. Then, we can proceed with the calculation of the

total requirements matrix using the newly constructed B_{imp} and D_{imp} . This matrix comes from the solution of the system of linear equations $g = D_{imp}(B_{imp}g + e_d^*)$, where e_d^* is the previously defined e_d vector supplemented by the sum of e_i as its last element, accounting for all import final demand purchases across commodities. The industry-by-industry total requirements matrix, also called the Leontief-inverse is the term $(I - D_{imp}B_{imp})^{-1}$ from the equation

$$g = (I - D_{imp}B_{imp})^{-1}D_{imp}e_d^*. \quad (14)$$

By this construction, the bottom right element of the total requirements matrix (representing the import sector's contribution to itself) is 1, but the same element in the direct requirements matrix was 0. This value has no real economic meaning, as in this case, the contribution of imports in producing imports has no meaningful interpretation. Therefore, this element is only used for the ranking restrictions in the identification scheme. Similar to other sectors, where the diagonal element is of key importance in the ranking restrictions, we can assume that a supply shock originating from imports will have the highest impact first on the import price index. This way, we can safely substitute the 0 in the last element of the direct requirements matrix to 1 after all calculations.

4.3 Ranking restrictions in input-output tables

In Tables 1 and 2, we give examples of the relevant subsets of the direct requirements matrix to illustrate our identifying restrictions. Since we estimate models of different sizes – with either 13 or 43 sectoral variables – the main downstream contributions of the sectors may change, hence the differences between the two tables. Both tables show the input-output coefficients with respect to the four sectoral shocks we identify in our models: Machinery, Computers and Electronic Products, Food and Imports. For each of these tables, the element A_{ij} shows the downstream contribution of the industry in row i to the industry in column j .

The identifying restrictions are based on the ranking of these values: the largest A_{ij} elements for each row i . Small specifications are identified by a 1–2 scheme, i.e. we are interested in the strictly largest element in a row, then additionally the second and third largest as well, but the ranking between the latter two does not matter. In large models however, we choose to implement a restriction set in which we find the two strictly largest elements in a row with no restrictions on the rest of the ranking. We call this the 2–0

scheme. This ranking is then checked against the impulse response functions, where (since the indices are swapped) the element in the row of industry j and the column of shock i is the impact of the corresponding sectoral shock on industry j .

Table 1: Subset of the 43×43 direct requirements table to illustrate ranking restrictions in model specifications of size $n = 50$ and $n = 45$ for the four sectoral shocks. The index of each sector is shown in parentheses, see Table A2 in the Appendix for the names and codes of each sector. The largest three values in each row are highlighted in bold font. For models with $n = 50$ and $n = 45$, the 2–0 restriction scheme applies.

Industry (ind.)	Mining (2)	Machinery (8)	Computer (9)	Electrical equipment (10)	Motor vehicles (11)	Other transportation equipment (12)	Food (15)	Petroleum and coal (20)	Residential care (41)	Accommodation (42)	Imports
Machinery (8)	0.0201	0.0479	0.0007	0.0084	0.0260	0.0084	0.0016	0.0001	0.0006	0.0001	0
Computer (9)	0.0001	0.0098	0.0599	0.0224	0.0082	0.0261	0.0015	0.0003	0.0002	0.0003	0
Food (15)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.1580	0.0008	0.0163	0.0188	0
Imports	0.0491	0.1270	0.0742	0.1088	0.1698	0.1507	0.0620	0.2521	0.0204	0.0262	1

Table 2: Subset of the 13×13 direct requirements table to illustrate ranking restrictions in model specifications of size $n = 20$ and $n = 15$ for the four sectoral shocks. The index of each sector is shown in parentheses, see Table A2 in the Appendix for the names and codes of each sector. The largest three values in each row are highlighted in bold font. For models with $n = 20$ and $n = 15$, the 1–2 restriction scheme applies.

Industry (ind.)	Oil and gas (1)	Wood products (4)	Primary metals (6)	Fabricated metal (7)	Machinery (8)	Computer (9)	Food (15)	Petroleum and coal (20)	Food and beverage stores (24)	Air transportation (26)	Imports
Machinery (8)	0.0161	0.0026	0.0062	0.0091	0.0479	0.0007	0.0016	0.0001	0.0004	0.0000	0
Computer (9)	0.0001	0.0052	0.0034	0.0076	0.0098	0.0599	0.0015	0.0003	0.0001	0.0000	0
Food (15)	0.0000	0.0003	0.0000	0.0000	0.0000	0.0000	0.1580	0.0008	0.0017	0.0015	0
Imports	0.0399	0.0778	0.1619	0.0780	0.1270	0.0742	0.0620	0.2521	0.0143	0.0916	1

Note that the decision on which variables to include in the model can significantly change the ranking in the input-output structure. For a given industry, if some of its downstream links are not included in a model, its top contributions may have to be derived from originally less important industries. For example, see the differences between the two tables: the top contributions of the machinery sector to other industries differ because

the Mining and Motor vehicles industries are not included in the smaller models. Instead, the top contributions of Machinery turn out to be on Oil & gas extraction and Fabricated metals. This highlights the role of the *diagonal element*, which is the link of the origin sector to itself. These are always present in the IO structure regardless of model size, and in most cases are the largest element in the row or column. However, if the results show similar IRF dynamics in both smaller and large models, it suggests that the most important ranking restriction is the largest element in each row.

4.4 Sectoral and aggregate time series

In this paper, we use monthly sectoral producer price index (PPI) data of the US economy to identify sectoral supply shocks. The time series cover the NAICS (North American Industry Classification System) industry categories, and thus need to be matched to the industry groups represented by the 71×71 input-output table. [Baqaee and Farhi \(2022\)](#) provide a solution for matching these two categories. Most matches can be found mechanically using the official correspondence tables for NAICS and IO industries, but in some cases, where the same input-output group is matched to several NAICS industries, the final matching is done manually, based on total industrial production figures (looking for the maximum contribution of the smaller industry groups) or as a last resort, qualitatively similar industry descriptions. This way, we found 42 reliable matches, which allows us to use a 43×43 subset of the IO table for identification, when including the additional Imports category, which represent a separate quasi-industry on its own. This artificial import industry group is then matched to an import price index provided by the Bureau of Labor Statistics (BLS), based on all imported goods.

The two main aggregate macro time series of interest are industrial production and unemployment. These two are included in all specifications, but in extended models ($n = 20$ and $n = 50$), we use five additional aggregate variables: CPI, 5-year inflation expectations, federal surplus and deficit, personal savings and consumer expenditure. Note that the original unit of unemployment is level (thousands of persons) instead of rates. Our main results are presented for industrial production and unemployment, but in Section 5.4, we also discuss the impact of sectoral shocks on the CPI. All variables are normalized to zero mean and unit standard deviation, which makes it possible to compare and rank impulse-response functions by their absolute magnitude. Before estimation, the trans-

formed series are available monthly from January 2004 to December 2019. Years of the pandemic are intentionally left out of the sample to minimize non-stationarity distortions in the parameter estimation. See Section A.6 in the Appendix for more details on variable sources, transformations and the selection of sectoral and aggregate time series for different model specifications. Note, that further observations are lost depending on the choice of lag parameter in the VAR model.

In all main specifications, we measure the aggregate implications of sectoral shocks by checking the impulse-responses on the core macro variables, industrial production, unemployment and CPI. We choose to concentrate on these variables only in order to maintain the focus of the analysis. It is true for all macro variables of the model, that no structural restrictions are imposed on them in the identification scheme. It would be possible to include aggregate impacts in the restriction set, doing so however should be well-founded by economic theory. In practice, it is possible to add sign-restrictions that require that a sectoral shock has for example negative impact on industrial production at a later horizon than the between-sector ranking restrictions. Using different horizons for the two sets of restrictions would be consistent with the mechanisms of sectoral shock propagation: it impacts the directly linked industries first, then the aggregate impact is realized later through these direct and indirect linkages.

5 Results

We present impulse response functions for different specifications: models of different sizes (mainly $n = 20$ and $n = 50$), depending on the set of sectoral and aggregate variables; where restrictions are imposed based on either the direct requirements matrix for smaller horizons of the IRF curve or the total requirements matrix for the longer horizon. All IRFs displayed show the aggregate impact of the four sectoral shocks on our two main macro variables, industrial production and unemployment. In Section 5.1 we present smaller ($n = 20$) models, then we compare results with large models ($n = 50$) in Section 5.2. Additionally, we investigate the effects of switching the restriction horizon for each type of requirements matrix used for identification. Finally, we discuss the impact of sectoral shocks on the consumer price index in Section 5.4. In Section A.5 of the Appendix we also show results from models where $n = 15$ or $n = 45$. All specifications assume a VAR(3) model: after preliminary estimations, the results are not affected by the number of lags

included, but the computational costs of estimation are greatly increased when more lags are included. Also, for monthly data, the 3-period lags correspond to observations through one quarter. The origin sectors in these examples are: Machinery, Computer and electronic products, Food and Imports.

The algorithm draws a certain number of reduced-form parameters and for each draw of Σ and $\mathbf{A}$, it checks 10 000 orthogonal $\mathbf{Q}$ matrices and the impulse response functions against the restriction set. See the computational properties of each model specification in Table A1 of the Appendix. For each specification, 1000 admissible draws from the posterior distribution are kept where the IRFs satisfy our restriction set. The following plots show the 68% credible set of the IRF distribution, and the black line shows the median¹ of this distribution. The IRFs are calculated only in terms of growth rates. In this way, a positive IRF for a given time horizon means that an identified sectoral shock causes a change in growth rates for that time horizon. The results are significant if the impulse response plots show 68% credible intervals that do not contain zero at one or more horizons.

5.1 Main results from smaller models ($n = 20$)

First, we present the results of specifications where $n = 20$ with the smaller set of sectoral time series and the expanded set of macro variables. We expect impulse responses with a negative sign for the impact of a negative sectoral supply shock on industrial production and a positive sign for the impact on unemployment. Figures 1 and 2 show that for this model size, identification based on both direct and total requirements tables yields results that are in line with our expectations for the sectoral shocks related to the Machinery and Computer industries. In all small-model specifications, the imposed ranking restrictions follow the 1–2 scheme described in Section 4.3. We suggest that this specification is sufficient to identify a sectoral shock. Then, our results can be interpreted as the impact of sectoral fluctuations on aggregate variables.

¹Kilian and Lütkepohl (2017) summarize the literature on the two major shortcomings of interpreting median IRFs in a meaningful way, namely 1) that the median response function does not directly correspond to any of the admissible parameter estimates, and 2) that the median is an inappropriate statistical measure of a vector valued random variable. In this paper our research hypothesis and the results are not affected by this issue, as there is no focus on any central values of the posterior IRF distribution, only the specific credible interval is important in determining whether the sectoral shocks have significant impact on the aggregate variables.

In most cases, we can see a positive impact of an import supply shock on industrial production, and even negative impacts on unemployment, contrary to our expectations and the IRFs observed in case of other sectoral shocks. While at this point, we have no sound economic explanation for this unexpected result, this issue deserves further investigation. However, this is beyond the scope of this paper.

Figure 1: VAR(3), $n = 20$, $h = 1$, Direct Requirements (1–2)

Impulse response functions of sectoral shocks on Industrial production [Panel (a)] and Unemployment [Panel (b)] up to a 12-month horizon in a $n = 20$ VAR(3) model. The identified sectoral shocks are Machinery, Computer and electronic products, Food and Imports. The restrictions are imposed based on the direct requirements matrix at the 1-month horizon, following the 1–2 restriction structure: the impact of the sectoral shock is largest on the top 1 downstream receiving industry, while the second and third largest need not have strict ranking. The blue area is the equal-tailed 68% credible set and the black line is the median of the posterior distribution of IRFs.

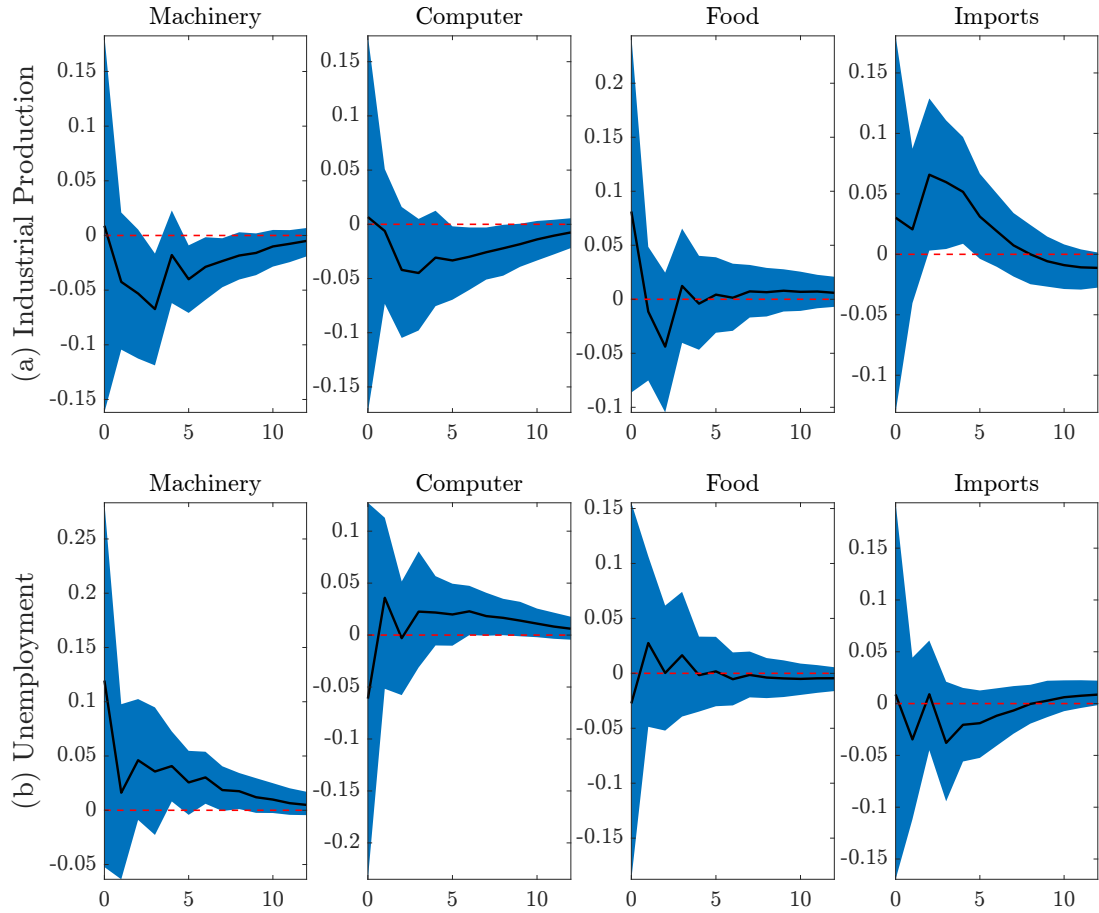
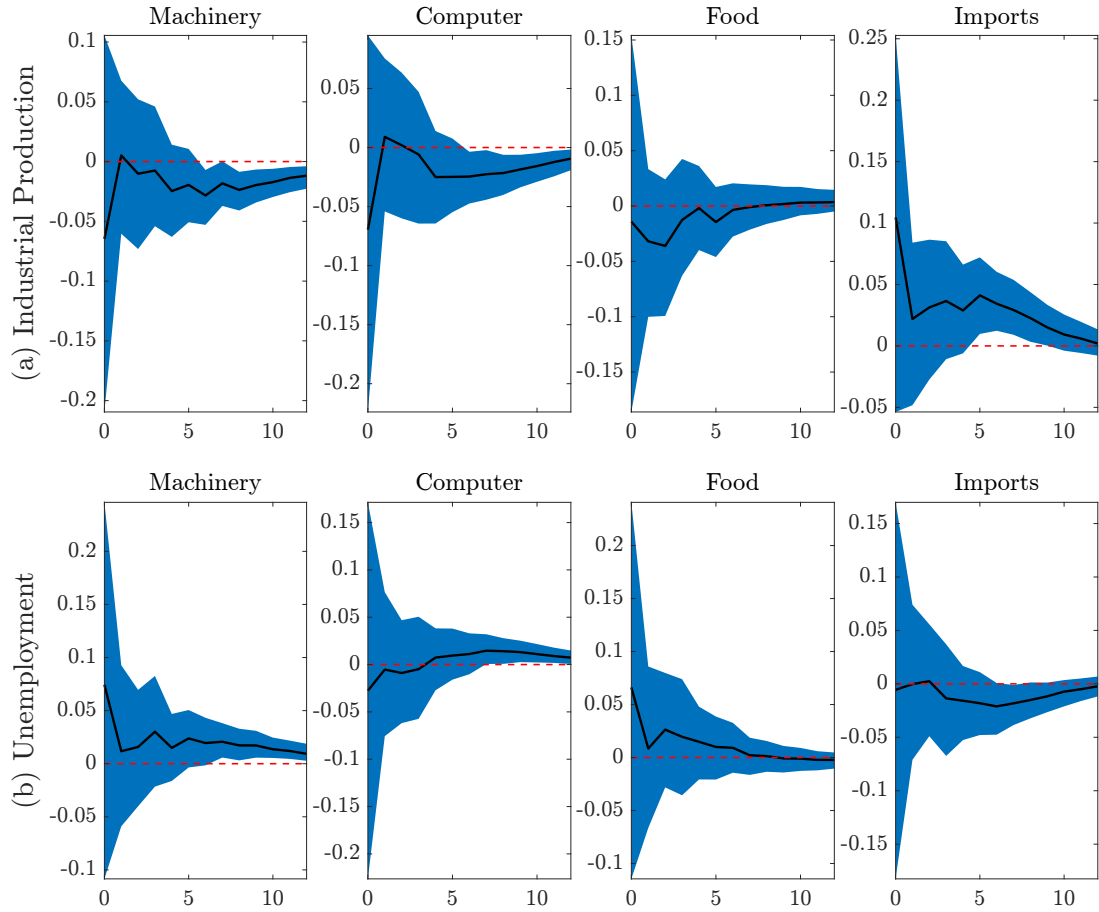


Figure 2: VAR(3), $n = 20$, $h = 4$, Total Requirements (1–2)

Impulse response functions of sectoral shocks on Industrial production [Panel (a)] and Unemployment [Panel (b)] up to a 12-month horizon in a $n = 20$ VAR(3) model. The identified sectoral shocks are Machinery, Computer and electronic products, Food and Imports. The restrictions are imposed based on the total requirements matrix at the 4-month horizon, following the 1–2 restriction structure: the impact of the sectoral shock is largest on the top 1 downstream receiving industry, while the second and third largest need not have strict ranking. The blue area is the equal-tailed 68% credible set and the black line is the median of the posterior distribution of IRFs.



5.2 Comparison with large models ($n = 50$)

We compare our results of the small model with the IRFs of the large model, where $n = 50$, which contains all 42+1 sectoral price information and all seven macro variables. The 1–2 restriction scheme turned out to be infeasible in this large model. The difficulty of imposing a large restriction set on the model is to be expected due to the high degree of variation in the impulse response values in a large matrix. Therefore, we use a 2–0 restriction scheme defined in Section 4.3. Expected results are the same as before, namely that a negative sectoral supply shock has a negative impact on industrial production and

a positive impact on unemployment. In large models, we choose a different shrinkage parameter in the prior selection process following the insight of [Bańbura et al. \(2010\)](#) that more powerful shrinkage of coefficients is needed in large models. See Section A.2 for more details on prior parameter selection. We find that the resulting signs of IRF curves are similar to the specifications in the previous section. Figure 3 shows a negative impact of the Machinery sectoral shocks on industrial production for the direct requirements matrix. Figure 4 also shows a positive impact on unemployment for the specification based on the total requirements matrix.

Figure 3: VAR(3), $n = 50$, $h = 1$, Direct Requirements (2-0)

Impulse response functions of sectoral shocks on Industrial production [Panel (a)] and Unemployment [Panel (b)] up to a 12-month horizon in a $n = 50$ VAR(3) model. The identified sectoral shocks are Machinery, Computer and electronic products, Food and Imports. The restrictions are imposed based on the direct requirements matrix on a 1-month horizon, following the 2-0 restriction structure: the impact of the sectoral shock is largest on the top 2 downstream receiving industries with strict ranking. The blue area is the equal-tailed 68% credible set and the black line is the median of the posterior distribution of IRFs.

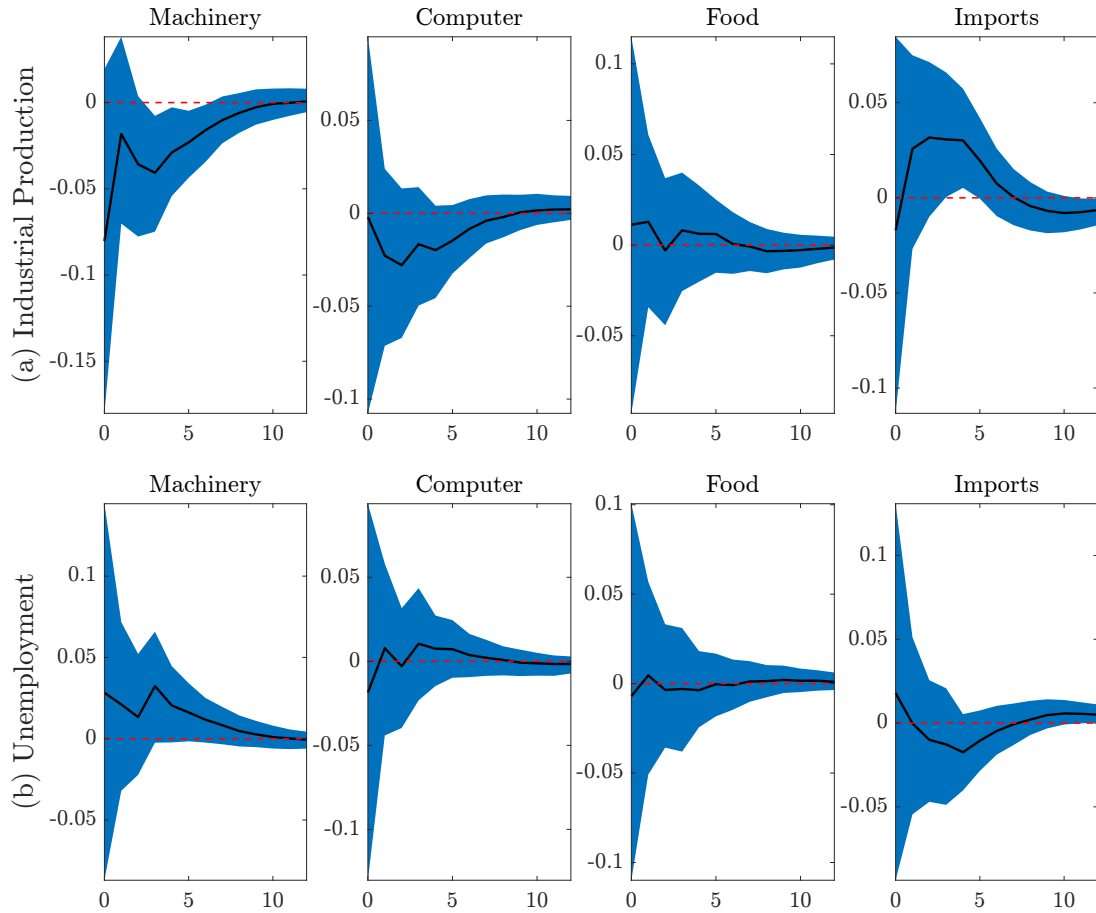
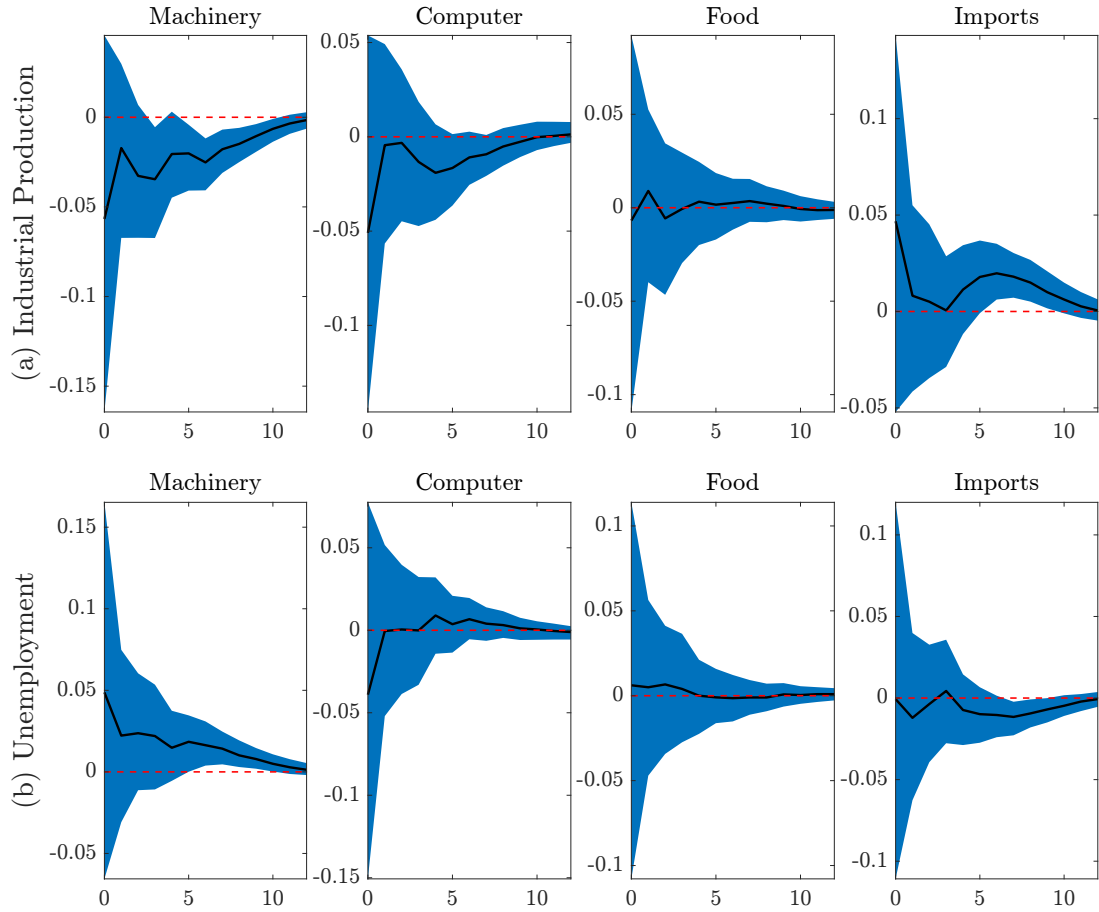


Figure 4: VAR(3), $n = 50$, $h = 4$, Total Requirements (2-0)

Impulse response functions of sectoral shocks on Industrial production [Panel (a)] and Unemployment [Panel (b)] up to a 12-month horizon in a $n = 50$ VAR(3) model. The identified sectoral shocks are Machinery, Computer and electronic products, Food and Imports. The restrictions are imposed based on the total requirements matrix at the 4-month horizon, following the 2-0 restriction structure: the impact of the sectoral shock is largest on the top 2 downstream receiving industries with strict ranking. The blue area is the equal-tailed 68% credible set and the black line is the median of the posterior distribution of IRFs.

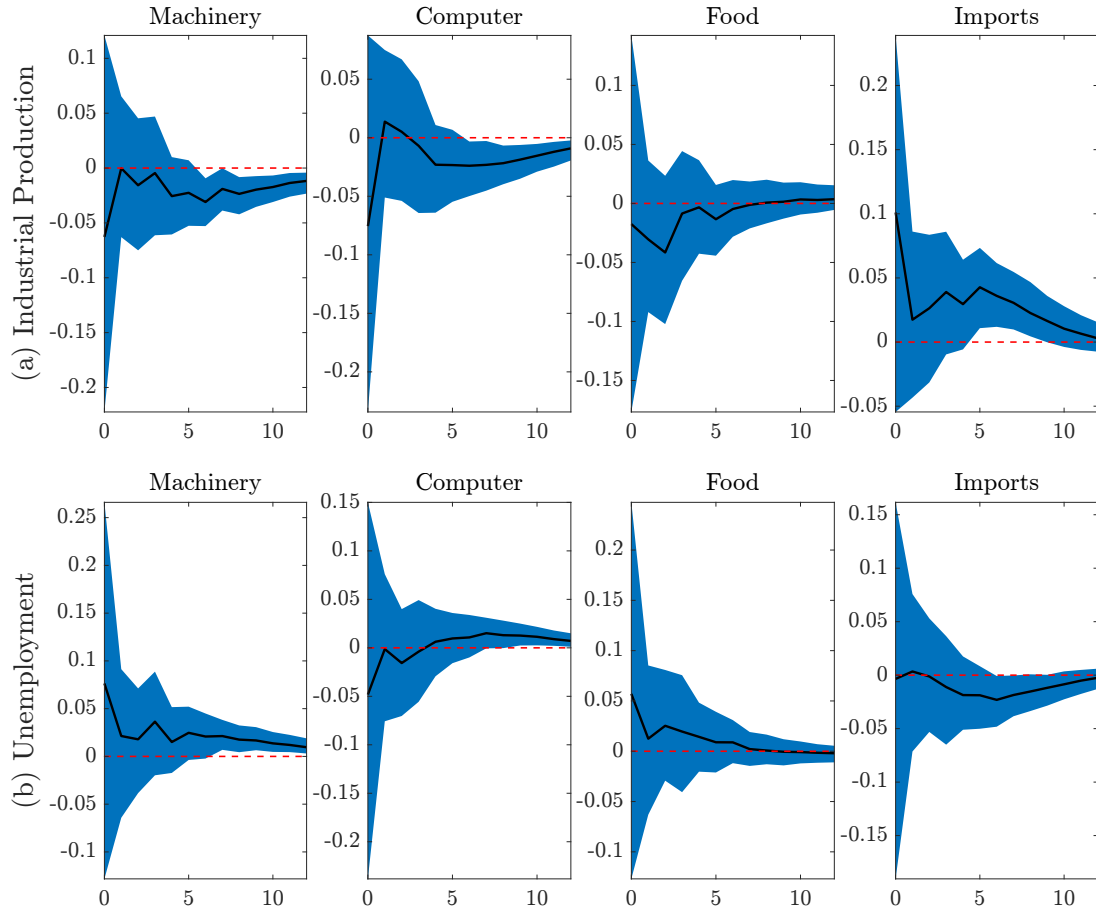


5.3 Choice of the restriction horizon

In order to disentangle differences in results by the choice of restriction horizons and the input-output matrix (total or direct requirements), we present IRFs for the $n = 20$ model, similar to the one in Figure 1 and 2, only switching the restriction horizons. This way, restrictions are imposed on the 4-month horizon when the direct requirements matrix is used and on the 1-month horizon for the total requirements table. Figures 5 and 6 show that results are similar for each chosen horizon, irrespective of the input-output matrix, which implies that it is the choice of restriction horizons which influences somewhat the overall dynamics of the impulse response curves, not the choice of input-output matrix.

Figure 5: VAR(3), $n = 20$, $h = 4$, Direct Requirements (1–2)

Impulse response functions of sectoral shocks on Industrial production [Panel (a)] and Unemployment [Panel (b)] up to a 12-month horizon in a $n = 20$ VAR(3) model. The identified sectoral shocks are Machinery mfg., Computer parts mfg., Food products mfg. and Imports. The restrictions are imposed based on the direct requirements matrix at the 4-month horizon, following the 1–2 restriction structure: the impact of the sectoral shock is largest on the top 1 downstream receiving industry, while the second and third largest need not have strict ranking. The blue area is the equal-tailed 68% credible set and the black line is the median of the posterior distribution of IRFs.

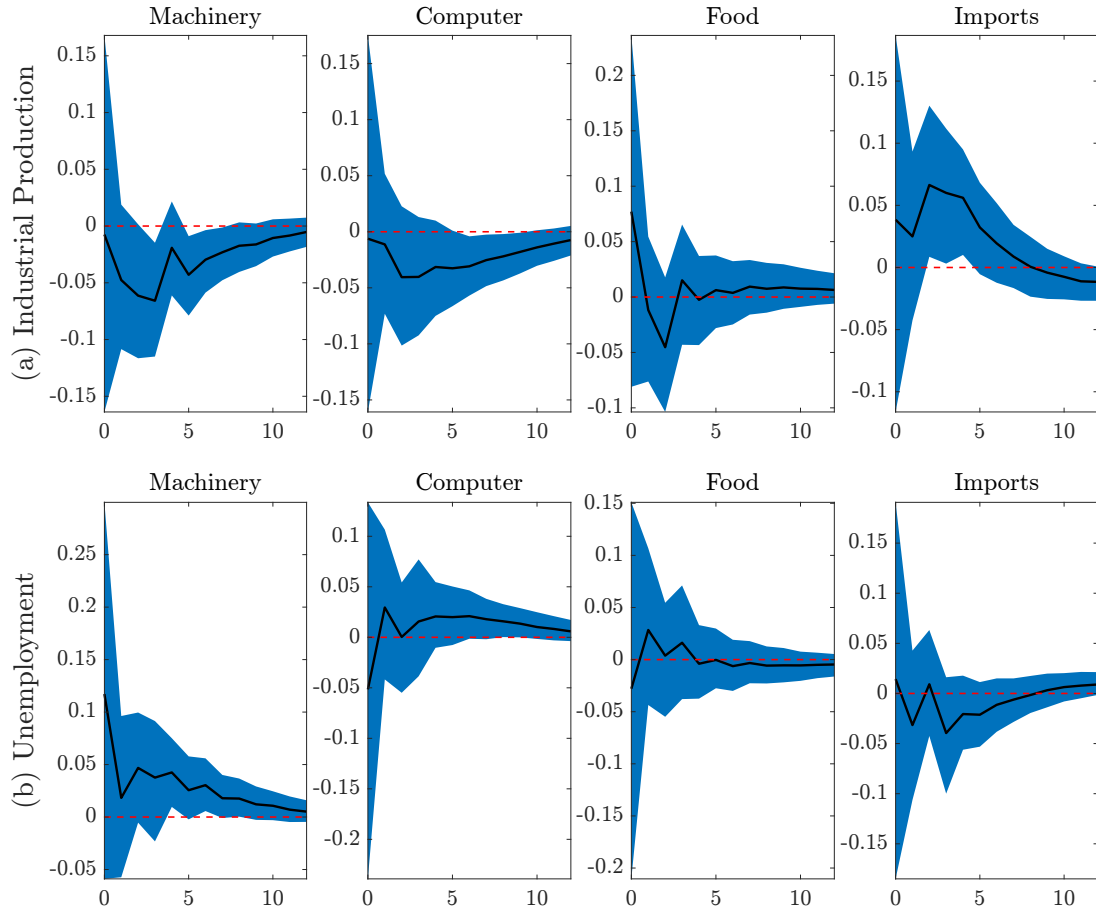


5.4 Impact of sectoral supply shocks on the CPI

While the primary focus regarding the aggregate consequences of sectoral shocks is on industrial production and unemployment, we also analyze the impact of these shocks on CPI. The results are interesting, because the propagation mechanism from producer to consumer prices is not trivial, especially given that sectoral producer conditions may only have limited contributions to the aggregate dynamics of prices. However, in case of important (large) sectors of the economy that are presumably closer to the final demand purchases in the production chain, we may expect that increased production costs due to a negative supply shock will *eventually* lead to increasing consumer prices. Therefore,

Figure 6: VAR(3), $n = 20$, $h = 1$, Total Requirements (1–2)

Impulse response functions of sectoral shocks on Industrial production [Panel (a)] and Unemployment [Panel (b)] up to a 12-month horizon in a $n = 20$ VAR(3) model. The identified sectoral shocks are Machinery, Computer and electronic products, Food and Imports. The restrictions are imposed based on the total requirements matrix at the 1-month horizon, following the 1–2 restriction structure: the impact of the sectoral shock is largest on the top 1 downstream receiving industry, while the second and third largest need not have strict ranking. The blue area is the equal-tailed 68% credible set and the black line is the median of the posterior distribution of IRFs.



in Figures 7 and 8 we present IRFs of the impact on aggregate CPI from the same two model specifications as in Figures 1 and 2 before: the $n = 20$ model with a (1–2) restriction identification scheme. Results show that only import supply shocks have a positive significant impact on consumer prices, the effect of other sectoral shocks cannot be seen in these specifications. Note that by treating imports as a separate source of sectoral shocks in the construction of the input-output tables, we essentially create an artificial sector that is considerably larger than other sectors. Since imports have a significant contribution to many sectors, it is expected to see that supply shocks related to imports will have a greater impact on CPI than any other sector.

Figure 7: VAR(3), $n = 20$, $h = 1$, Direct Requirements (1–2), CPI

Impulse response functions of sectoral shocks on CPI up to a 12-month horizon in a $n = 20$ VAR(3) model. The identified sectoral shocks are Machinery, Computer and electronic products, Food and Imports. The restrictions are imposed based on the direct requirements matrix at the 1-month horizon, following the 1–2 restriction structure: the impact of the sectoral shock is largest on the top 1 downstream receiving industry, while the second and third largest need not have strict ranking. The blue area is the equal-tailed 68% credible set and the black line is the median of the posterior distribution of IRFs.

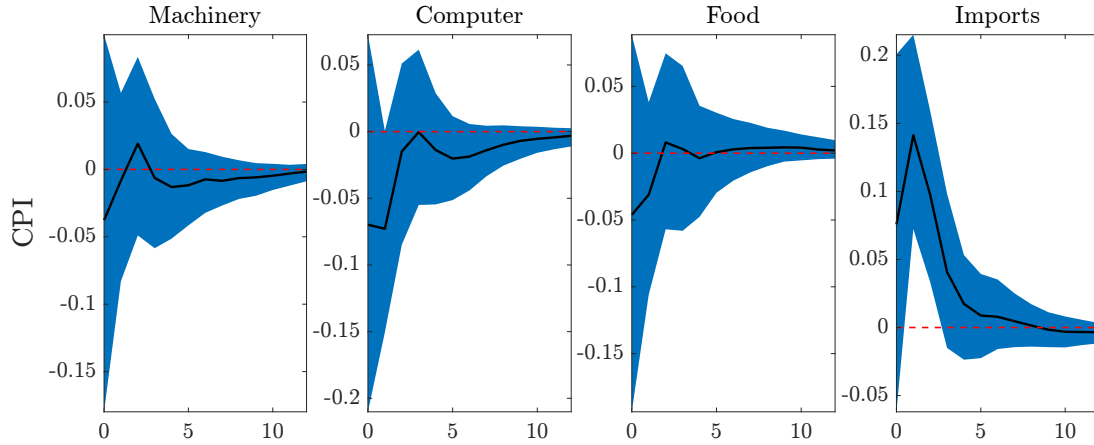
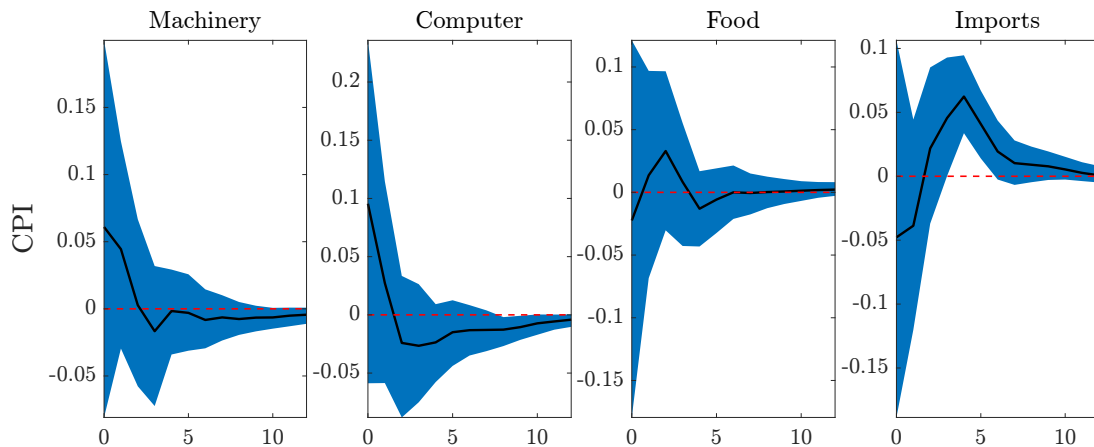


Figure 8: VAR(3), $n = 20$, $h = 4$, Total Requirements (1–2), CPI

Impulse response functions of sectoral shocks on CPI up to a 12-month horizon in a $n = 20$ VAR(3) model. The identified sectoral shocks are Machinery, Computer and electronic products, Food and Imports. The restrictions are imposed based on the total requirements matrix at the 4-month horizon, following the 1–2 restriction structure: the impact of the sectoral shock is largest on the top 1 downstream receiving industry, while the second and third largest need not have strict ranking. The blue area is the equal-tailed 68% credible set and the black line is the median of the posterior distribution of IRFs.



6 Conclusions

In this paper, we provided a strategy for identifying multiple sector-specific supply shocks jointly using a large Bayesian VAR approach, where the structural restrictions in the form of ranking restrictions are derived from input-output tables. In our model specifications, we partially identified 4 sectoral shocks – Machinery, Computer and electronic products manufacturing, Food and Imports – as opposed to the strategy of [De Graeve and Schneider \(2023\)](#) who only identify one sectoral shock at a time. We used the approach of [Chan et al. \(2023\)](#) to solve the computational intractability of large models when identified by the standard [Rubio-Ramírez et al. \(2010\)](#) algorithm for set-type restrictions. Therefore, our main contribution to the literature is the joint application of two very recent methods: that of identifying sectoral shocks and the efficient estimation algorithm that makes it feasible to do so in large Bayesian VARs.

We also discussed in details the possible issues that arise when applying this novel identification method to a purely ranking restriction scheme, such as our large sectoral model. We investigated whether sectoral shocks have a significant impact on aggregate macro variables like industrial production, unemployment and CPI. We found that in most model specifications, regardless of model size, the choice of restriction horizons or the type of input-output tables, sectoral shocks related to the Machinery and Computer manufacturing industries showed a significant negative impact on aggregate industrial production. Thus, we confirmed theoretical findings of models such as [Acemoglu et al. \(2012\)](#), namely that sectoral supply shocks have a significant impact on aggregate macro variables.

In future research, besides extending this large Bayesian VAR model to cover more sectoral shocks or to impose tighter restrictions on a smaller number of shocks, it is possible to compare the results of the large Bayesian setup with the factor-augmented models. For this purpose, the FAVAR model of [De Graeve and Schneider \(2023\)](#) can be extended to include not only one, but several sectoral shocks simultaneously as observed factors. This extended partial identification strategy can be used to capture the main sources of sectoral fluctuations in a systemic way, allowing a more direct comparison of the BVAR and FAVAR techniques in terms of shock identification.

As [Chan et al. \(2023\)](#) note, an interesting extension of their new algorithm is the inclusion of zero restrictions in the model. This would be a particularly practical and relevant extension in the case of IO tables, as many sectors have minimal or almost no exposure

to some industries. The algorithm of [Arias, Rubio-Ramírez, and Waggoner \(2018\)](#) would already be theoretically applicable in this large sectoral model without problems, but would most likely be computationally intractable, similar to the regular algorithm of [Rubio-Ramírez et al. \(2010\)](#) for set-like restrictions. Therefore, it would make sense to use an identification algorithm that can simultaneously use zero restrictions and retain the computationally favorable properties of the algorithm of [Chan et al. \(2023\)](#).

The future research plan also includes theoretical considerations on the production network, in particular on the propagation time of shocks through input-output links. This concept has far-reaching consequences for the practice of shock identification through input-output tables. Both theoretical and empirical research suggests that the average propagation time of the production network is relatively large, i.e. the propagation of industry-specific shocks is not instantaneous: this is investigated, for example, in the empirical study by [Rácz \(2023\)](#). In this case, the on-impact IRFs should be checked using an input-output structure that captures short-term spillover channels in the system, such as the direct requirements matrix. Moreover, the use of different exponents of the matrix for each IRF horizon can be explored depending on the estimated propagation speeds. This is why we proposed to investigate the differences between the use of direct and total requirements matrices for identification.

In addition, there are some interesting ideas in production network research related to monetary policy and fiscal multipliers, for example in [Ghassibe \(2021a, 2021b\)](#) and [Ghassibe and Zanetti \(2022\)](#). All three papers present theoretical models of the state-dependent nature and complexity of monetary policy transmission (role of production networks) and the fiscal multiplier, and also provide econometric evidence using the simple local projections method of [Jordà \(2005\)](#). The systemic analysis of sectoral variables made possible by large Bayesian VARs may prove to be a suitable tool to extend the econometric evidence for these theoretical insights. For example, large sectoral models may be used to capture the involvement of production networks in the propagation of an aggregate fiscal stimulus; or even the local multiplier effects of sectoral fiscal shocks. Also, in general, large Bayesian VARs could be used to investigate findings of macro models built on production networks, such as [La'O and Tahbaz-Salehi \(2022\)](#); [Luo and Villar \(2023\)](#).

References

- Acemoglu, D., Akcigit, U., & Kerr, W. (2016). Networks and the Macroeconomy: An Empirical Exploration. *NBER Macroeconomics Annual*, 30(1), 273–335. Retrieved 2023-12-22, from <https://ideas.repec.org//a/ucp/macann/doi10.1086-685961.html>
- Acemoglu, D., Carvalho, V. M., Ozdaglar, A., & Tahbaz-Salehi, A. (2012). The Network Origins of Aggregate Fluctuations. *Econometrica*, 80(5), 1977–2016. Retrieved 2023-12-14, from <https://www.jstor.org/stable/23271439>
- Amir-Ahmadi, P., & Drautzburg, T. (2021). Identification and inference with ranking restrictions. *Quantitative Economics*, 12(1), 1–39. Retrieved 2023-08-31, from <https://onlinelibrary.wiley.com/doi/abs/10.3982/QE1277>
- Arias, J. E., Rubio-Ramírez, J. F., & Waggoner, D. F. (2018). Inference Based on Structural Vector Autoregressions Identified with Sign and Zero Restrictions: Theory and Applications. *Econometrica*, 86(2), 685–720. Retrieved 2023-10-13, from <https://www.jstor.org/stable/44955981>
- Baqae, D., & Farhi, E. (2022, May). Supply and Demand in Disaggregated Keynesian Economies with an Application to the COVID-19 Crisis. *American Economic Review*, 112(5), 1397–1436. Retrieved 2023-12-28, from <https://www.aeaweb.org/articles?id=10.1257/aer.20201229>
- Bañbura, M., Giannone, D., & Reichlin, L. (2010). Large Bayesian vector auto regressions. *Journal of Applied Econometrics*, 25(1), 71–92. Retrieved 2023-08-23, from <https://onlinelibrary.wiley.com/doi/abs/10.1002/jae.1137>
- Bernanke, B. S., Boivin, J., & Elias, P. (2005, February). Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach*. *The Quarterly Journal of Economics*, 120(1), 387–422. Retrieved 2023-09-11, from <https://doi.org/10.1162/0033553053327452>
- Canova, F., & Nicoló, G. D. (2002, September). Monetary disturbances matter for business fluctuations in the G-7. *Journal of Monetary Economics*, 49(6), 1131–1159. Retrieved 2024-04-11, from <https://www.sciencedirect.com/science/article/pii/S0304393202001459>
- Chan, J. C. C. (2020). Large Bayesian Vector Autoregressions. In P. Fuleky (Ed.), *Macroeconomic Forecasting in the Era of Big Data: Theory and Practice* (pp. 95–125). Cham: Springer International Publishing. Retrieved 2023-12-14, from

https://doi.org/10.1007/978-3-030-31150-6_4

- Chan, J. C. C., Matthes, C., & Yu, X. (2023). Large Structural VARs with Multiple Sign and Ranking Restrictions. Retrieved from <https://joshuachan.org/papers/SVAR-sign.pdf>
- De Graeve, F., & Karas, A. (2014, August). Evaluating Theories of Bank Runs with Heterogeneity Restrictions. *Journal of the European Economic Association*, 12(4), 969–996. Retrieved 2023-08-31, from <https://doi.org/10.1111/jeea.12080>
- De Graeve, F., & Schneider, J. D. (2023, August). Identifying sectoral shocks and their role in business cycles. *Journal of Monetary Economics*. Retrieved 2023-08-31, from <https://www.sciencedirect.com/science/article/pii/S0304393223000971>
- Ghassibe, M. (2021a). *Endogenous production networks and non-linear monetary transmission*. Retrieved 2023-12-02, from https://ghassibem.github.io/Ghassibe_endognet_21oct21_.pdf
- Ghassibe, M. (2021b). Monetary policy and production networks: an empirical investigation. *Journal of Monetary Economics*, 119, 21–39. Retrieved 2023-11-20, from <https://www.sciencedirect.com/science/article/pii/S0304393221000179>
- Ghassibe, M., & Zanetti, F. (2022). State dependence of fiscal multipliers: the source of fluctuations matters. *Journal of Monetary Economics*, 132, 1–23. Retrieved 2023-11-20, from <https://www.sciencedirect.com/science/article/pii/S0304393222001155>
- Jordà, Ò. (2005, March). Estimation and Inference of Impulse Responses by Local Projections. *American Economic Review*, 95(1), 161–182. Retrieved 2023-12-28, from <https://www.aeaweb.org/articles?id=10.1257/0002828053828518>
- Kilian, L., & Lütkepohl, H. (2017). Identification by Sign Restrictions. In *Structural Vector Autoregressive Analysis* (pp. 421–490). Cambridge University Press.
- La’O, J., & Tahbaz-Salehi, A. (2022). Optimal Monetary Policy in Production Networks. *Econometrica*, 90(3), 1295–1336. Retrieved 2023-08-25, from <https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA18627>
- Litterman, R. B. (1986). Forecasting with Bayesian Vector Autoregressions: Five Years of Experience. *Journal of Business & Economic Statistics*, 4(1), 25–38. Retrieved 2023-12-22, from <https://www.jstor.org/stable/1391384>

- Luo, S., & Villar, D. (2023, July). Propagation of shocks in an input-output economy: Evidence from disaggregated prices. *Journal of Monetary Economics*, 137, 26–46. Retrieved 2023-08-24, from <https://www.sciencedirect.com/science/article/pii/S030439322300048X>
- Rubio-Ramírez, J. F., Waggoner, D. F., & Zha, T. (2010). Structural Vector Autoregressions: Theory of Identification and Algorithms for Inference. *The Review of Economic Studies*, 77(2), 665–696. Retrieved 2023-10-13, from <https://www.jstor.org/stable/40587642>
- Rácz, O. M. (2023, March). *Network Origins of Aggregate Dynamics* [SSRN Scholarly Paper]. Rochester, NY. Retrieved 2024-01-04, from <https://papers.ssrn.com/abstract=4392276>
- Sawyer, J. A. (1992, January). Forecasting with Input–Output Matrices: Are the Coefficients Stationary? *Economic Systems Research*, 4(4), 325–348. Retrieved 2024-04-11, from <https://doi.org/10.1080/09535319200000030>
- Torres-González, L. D., & Yang, J. (2019, October). The persistent statistical structure of the US input–output coefficient matrices: 1963–2007. *Economic Systems Research*, 31(4), 481–504. Retrieved 2024-04-11, from <https://doi.org/10.1080/09535314.2018.1561425>

A Appendix

A.1 Key equations of the VAR(p) system

For Bayesian estimation, there are two useful representations of the VAR(p) system. The first stacks the VAR(p) over $t = 1, \dots, T$ using the following representation:

$$\mathbf{y}_t = \mathbf{X}_t \boldsymbol{\beta} + \boldsymbol{\varepsilon}_t, \quad (15)$$

where $\mathbf{X}_t = \mathbf{I}_n \otimes [1, \mathbf{y}'_{t-1}, \dots, \mathbf{y}'_{t-p}]$, an $n \times nk$ matrix, and $\boldsymbol{\beta} = \text{vec}([\mathbf{b}, \mathbf{A}_1, \dots, \mathbf{A}_p]')$, an $nk \times 1$ vector of coefficients. Stacking by t , such as $\mathbf{y} = (\mathbf{y}'_1, \dots, \mathbf{y}'_T)'$, we obtain:

$$\mathbf{y} = \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad (16)$$

where the matrix of regressors $\mathbf{X} = (\mathbf{X}'_1, \dots, \mathbf{X}'_T)'$ is $Tn \times nk$ and $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_T \otimes \boldsymbol{\Sigma})$.

The second representation stacks the dependent variables into a $T \times n$ matrix $\mathbf{Y}$, in which the t -th row is $\mathbf{y}'_t$. Then, $\mathbf{Z}$ is a $T \times k$ matrix of regressors, such that the t -th row is $\mathbf{x}'_t = (1, \mathbf{y}'_{t-1}, \dots, \mathbf{y}'_{t-p})$. The coefficient matrix $\mathbf{A} = (\mathbf{b}, \mathbf{A}_1, \dots, \mathbf{A}_p)'$ is $k \times n$. This provides the following representation of VAR(p):

$$\mathbf{Y} = \mathbf{Z} \mathbf{A} + \mathbf{U}, \quad (17)$$

where $\mathbf{U}$ is $T \times n$, where the t -th row is $\boldsymbol{\varepsilon}'_t$. The two representations are connected by the following equalities: $\mathbf{y} = \text{vec}(\mathbf{Y}')$, $\boldsymbol{\beta} = \text{vec}(\mathbf{A})$ and $\boldsymbol{\varepsilon} = \text{vec}(\mathbf{U}')$. Then,

$$\text{vec}(\mathbf{U}) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma} \otimes \mathbf{I}_T). \quad (18)$$

The two equivalent representations of the VAR(p) imply the following two likelihood functions:

$$(\mathbf{y} | \boldsymbol{\beta}, \boldsymbol{\Sigma}) \sim \mathcal{N}(\mathbf{X} \boldsymbol{\beta}, \mathbf{I}_T \otimes \boldsymbol{\Sigma}), \quad (19)$$

$$p(\mathbf{y} | \boldsymbol{\beta}, \boldsymbol{\Sigma}) = (2\pi)^{-\frac{Tn}{2}} |\boldsymbol{\Sigma}|^{-\frac{T}{2}} e^{-\frac{1}{2}(\mathbf{y} - \mathbf{X} \boldsymbol{\beta})' (\mathbf{I}_T \otimes \boldsymbol{\Sigma}^{-1}) (\mathbf{y} - \mathbf{X} \boldsymbol{\beta})}, \quad (20)$$

$$(\mathbf{Y} | \mathbf{A}, \boldsymbol{\Sigma}) \sim \mathcal{N}(\mathbf{Z} \mathbf{A}, \boldsymbol{\Sigma} \otimes \mathbf{I}_T), \quad (21)$$

$$p(\mathbf{Y} | \mathbf{A}, \boldsymbol{\Sigma}) = (2\pi)^{-\frac{Tn}{2}} |\boldsymbol{\Sigma}|^{-\frac{T}{2}} e^{-\frac{1}{2} \text{tr}(\boldsymbol{\Sigma}^{-1} (\mathbf{Y} - \mathbf{Z} \mathbf{A})' (\mathbf{Y} - \mathbf{Z} \mathbf{A}))}, \quad (22)$$

A.2 Derivation of the natural conjugate prior

The natural conjugate prior is used due to the computational burdens of the independent normal and inverse-Wishart prior setup. In the present case, with $n = 50$, matrix inversions and multiplications quickly lead to computational limitations, despite the fact that [Chan \(2020\)](#) provides an efficient method so that the matrix inversion of size $nk \times nk$ does not become completely infeasible. First, the inverse need not be calculated directly, instead, backward and forward substitution can be used to solve the corresponding linear system of equations. Second, sampling from the posterior distribution of β does not require the calculation of that inverse either. However, the independent normal and inverse-Wishart prior setup still proves to be infeasible with large n and relatively large T , because of certain matrix multiplications in each step of the Gibbs-sampler. More specifically, when the parameter $\mathbf{K}_\beta^{-1}$ of the β posterior distribution is estimated, the matrix multiplication $\mathbf{X}'(\mathbf{I}_T \otimes \Sigma^{-1})\mathbf{X}$ is of size $(nk \times Tn) \times (Tn \times Tn) \times (Tn \times nk)$. While exploiting the properties of quick sparse matrix multiplication helps, this still takes a lot of time when e.g. $n = 50$, $T = 189$, and that is in each iteration of the sampler that may be repeated many thousand times.

For the natural conjugate prior, the second representation of the $\text{VAR}(p)$ is used. In general, an $n \times n$ random matrix Ω follows an inverse-Wishart distribution with shape parameter $v > 0$ and scale matrix $\mathbf{S}$ if its density function is:

$$f(\Omega|v, \mathbf{S}) = \frac{|\mathbf{S}^{v/2}|}{2^{nv/2}\Gamma_n(v/2)} |\Omega|^{-\frac{v+n+1}{2}} e^{-\frac{1}{2}\text{tr}(\mathbf{S}\Omega^{-1})}, \quad (23)$$

where Γ_n is the multivariate gamma function. Then, $\Omega \sim \mathcal{IW}(v, \mathbf{S})$ with the important property of $\mathbb{E}\Omega = \mathbf{S}/(v - n - 1)$ for $v > n + 1$.

An $m \times n$ random matrix $\mathbf{W}$ and an $n \times n$ random matrix Ω are said to have a normal-inverse-Wishart distribution with parameters $\mathbf{M}, \mathbf{P}, \mathbf{S}$ and v if $(\text{vec}(\mathbf{W})|\Omega) \sim \mathcal{N}(\text{vec}(\mathbf{M}), \Omega \otimes \mathbf{P})$ and $\Omega \sim \mathcal{IW}(v, \mathbf{S})$. Then we can write the joint distribution as $(\text{vec}(\mathbf{W}), \Omega) \sim \mathcal{NIW}(\text{vec}(\mathbf{M}), \mathbf{P}, v, \mathbf{S})$. It can be derived that kernel of the normal-inverse-Wishart density function is the following:

$$f(\mathbf{W}, \Omega|\mathbf{M}, \mathbf{P}, v, \mathbf{S}) \propto |\Omega|^{-\frac{v+m+n+1}{2}} e^{-\frac{1}{2}\text{tr}(\Omega^{-1}(\mathbf{W}-\mathbf{M})'\mathbf{P}^{-1}(\mathbf{W}-\mathbf{M}))} e^{-\frac{1}{2}\text{tr}(\Omega^{-1}\mathbf{S})}. \quad (24)$$

We use the following normal-inverse-Wishart prior on parameters $(\mathbf{A}, \Sigma)$:

$$\Sigma \sim \mathcal{IW}(v_0, \mathbf{S}_0), \quad (\text{vec}(\mathbf{A}) | \Sigma) \sim \mathcal{N}(\text{vec}(\mathbf{A}), \Sigma \otimes \mathbf{V}_\mathbf{A}). \quad (25)$$

The joint density function of the distribution $(\text{vec}(\mathbf{A}), \Sigma) \sim \mathcal{NIW}(\text{vec}(\mathbf{A}_0), \mathbf{V}_\mathbf{A}, v_0, \mathbf{S}_0)$ is:

$$p(\mathbf{A}, \Sigma) \propto |\Sigma|^{-\frac{v_0+n+k+1}{2}} e^{-\frac{1}{2}\text{tr}(\Sigma^{-1}\mathbf{S}_0)} e^{-\frac{1}{2}\text{tr}(\Sigma^{-1}(\mathbf{A}-\mathbf{A}_0)'\mathbf{V}_\mathbf{A}^{-1}(\mathbf{A}-\mathbf{A}_0))}. \quad (26)$$

Combining the likelihood and the prior, by Bayes's theorem, we get the posterior, which is the kernel of the normal-inverse-Wishart density function:

$$p(\mathbf{A}, \Sigma | \mathbf{Y}) \propto p(\mathbf{A}, \Sigma) p(\mathbf{Y} | \mathbf{A}, \Sigma) \quad (27)$$

$$\propto |\Sigma|^{-\frac{v_0+n+k+T+1}{2}} e^{-\frac{1}{2}\text{tr}(\Sigma^{-1}\hat{\mathbf{S}})} e^{-\frac{1}{2}\text{tr}[\Sigma^{-1}(\mathbf{A}-\hat{\mathbf{A}})'\mathbf{K}_\mathbf{A}(\mathbf{A}-\hat{\mathbf{A}})]}. \quad (28)$$

That is,

$$(\mathbf{A}, \Sigma | \mathbf{Y}) \sim \mathcal{NIW}(\hat{\mathbf{A}}, \mathbf{K}_\mathbf{A}^{-1}, v_0 + T, \hat{\mathbf{S}}), \quad (29)$$

where

$$\mathbf{K}_\mathbf{A} = \mathbf{V}_\mathbf{A}^{-1} + \mathbf{Z}'\mathbf{Z}, \quad (30)$$

$$\hat{\mathbf{A}} = \mathbf{K}_\mathbf{A}^{-1} (\mathbf{V}_\mathbf{A}^{-1}\mathbf{A}_0 + \mathbf{Z}'\mathbf{Y}), \quad (31)$$

$$\hat{\mathbf{S}} = \mathbf{S}_0 + \mathbf{A}_0'\mathbf{V}_\mathbf{A}^{-1}\mathbf{A}_0 + \mathbf{Y}'\mathbf{Y} - \hat{\mathbf{A}}'\mathbf{K}_\mathbf{A}\hat{\mathbf{A}}. \quad (32)$$

Following [Chan \(2020\)](#), we set the values for the natural conjugate prior as: $\mathbf{A}_0 = \mathbf{0}$, the two hyperparameters of the covariance matrix $\mathbf{V}_\mathbf{A}$ are assumed to be $c_2 = 10^2$ and $c_1 = 0.2^2$ for the smaller models ($n = 15$ and $n = 20$). However, we use $c_1 = 0.1^2$ for large models ($n = 45$ and $n = 50$), which results in a more powerful shrinkage of coefficients, which is particularly useful in large models ([Bańbura et al., 2010](#)). Also, $v_0 = n + 3$, and $\mathbf{S}_0 = \text{diag}(s_1^2, \dots, s_n^2)$, where s_i^2 is the standard OLS estimate of the error variance for the i -th equation of the VAR(p).

Properties of the natural conjugate prior allow direct sampling from the joint posterior distribution of $(\mathbf{A}, \Sigma)$ without the need to use Gibbs-sampler (or other MCMC methods). The direct consequence of this is that the sample is truly independent, as opposed to

MCMC draws, where close sample values are correlated. Drawing $\mathbf{A}$ and $\mathbf{\Sigma}$ from the normal-inverse-Wishart distribution is done by first drawing $\mathbf{\Sigma}$ from the marginal inverse-Wishart distribution, then to draw the vector form of $\mathbf{A}$ conditional on $\mathbf{\Sigma}$. To summarize, a computationally feasible algorithm for estimating model with natural conjugate priors follows these steps:

1. Set prior parameters $(\mathbf{A}_0, \mathbf{V}_\mathbf{A}, v_0, \mathbf{S}_0)$.
2. Formulate likelihood $p(\mathbf{Y}|\mathbf{A}, \mathbf{\Sigma})$ and prior density $p(\mathbf{A}, \mathbf{\Sigma})$.
3. Calculate parameters of the posterior distribution: $(\mathbf{K}_\mathbf{A}, \hat{\mathbf{A}}, \hat{\mathbf{S}})$.
4. Draw from the distribution $\mathcal{NIW}(\hat{\mathbf{A}}, \mathbf{K}_\mathbf{A}^{-1}, v_0 + T, \hat{\mathbf{S}})$.
5. Given a draw of $(\mathbf{A}, \mathbf{\Sigma})$, calculate Wold IRFs.
6. Identify shocks based on the restriction set. Keep parameter draws and IRFs that fulfill the restrictions.
7. Repeat steps (4)–(6) to account for estimation and identification uncertainty.

A.3 Details on the identification algorithm

For identification, the modified algorithm suggested by [Chan et al. \(2023\)](#) is based on the fact that uniformly drawn $\mathbf{Q}$ matrices from $\mathcal{O}(n)$ are invariant to right-multiplication by any other orthogonal matrix, such as the permutation matrix $\mathbf{P}$ and column sign switch matrix $\mathbf{D}$. Then, the set $\mathcal{Q}(\mathbf{\Sigma}, \mathbf{Q})$ consists of all possible permutations and sign switches of $\mathbf{LQ}$ for a given $\mathbf{\Sigma}$ matrix such that $\mathbf{E} = \mathbf{LQPD} \in \mathcal{Q}(\mathbf{\Sigma}, \mathbf{Q})$. Then, for any $\mathbf{Q}^* = \mathbf{QPD}$, $\mathbf{Q}^* \sim \mathcal{U}(\mathcal{O}(n))$ and $\mathbf{LQ}^*\mathbf{Q}^{*\prime}\mathbf{L}' = \mathbf{\Sigma}$. They provide an algorithm that efficiently checks whether it is true for any element $\mathbf{E}$ of $\mathcal{Q}(\mathbf{\Sigma}, \mathbf{Q})$ that $\mathbf{C}(L)\mathbf{E}$ satisfies all restrictions. To do that, they construct an auxiliary matrix $\mathbf{T}$ of size $m \times n$ (m is the number of identified shocks, while there are n variables in total) where each element $\mathbf{T}_{ij}$ shows whether column j – or its negative – of the IRF matrix satisfies the restrictions corresponding to shock i . $\mathbf{T}$ is constructed quickly, then, it can be checked easily whether the set of admissible orthogonal matrices for any $\mathbf{Q}$ draw is empty or not. If there is a row in $\mathbf{T}$ containing all zeros, then there is a structural shock for which no column fits all restrictions, the set is empty and $\mathbf{Q}$ is discarded. If the set is non-empty, we can uniformly draw from it by choosing unused column indices for each shock so that the column satisfies all restrictions.

To summarize, step (6) in the estimation algorithm described above consists of the following sub-steps given a parameter draw of $(\mathbf{A}, \mathbf{\Sigma})$ and the Wold IRFs:

1. Draw an orthogonal matrix $\mathbf{Q} \sim \mathcal{U}(\mathcal{O}(n))$.
2. Using the algorithm of [Chan et al. \(2023\)](#), construct the $\mathbf{T}$ matrix.
3. Check if the set of admissible IRFs is empty or not.
4. If the set is empty, draw another $\mathbf{Q}$ matrix.
5. If not, draw uniformly from this set and calculate IRFs.
6. Repeat steps (1)–(5) many times to account for identification uncertainty.

A.4 Computational properties

Table A1: Number of admissible IRFs for each specification, for a given number of parameter and $\mathbf{Q}$ matrix draws and the time required to complete the loop. The specification details include model size, the identification scheme, the selected restriction horizon and the type of input-matrix.

Matlab version: R2023b, Parallel workers for *parfor*: 8

CPU: AMD Ryzen 7 4800HS, 8 cores (16 threads), 2.9 GHz. RAM: 16GB

Model size	Identification scheme	Restriction horizon	Type of input-output table	Parameter draws	Q draws	Admissible IRFs	Time
$n = 20$	1-2	$h = 1$	DR	15000	10000	1785	0h47m
$n = 20$	1-2	$h = 1$	TR	15000	10000	1763	0h51m
$n = 20$	1-2	$h = 4$	TR	15000	10000	1207	0h43m
$n = 20$	1-2	$h = 4$	DR	15000	10000	1195	0h45m
$n = 50$	1-1	$h = 1$	DR	15000	10000	1678	3h48m
$n = 50$	1-1	$h = 4$	TR	30000	10000	1168	7h35m
$n = 15$	1-2	$h = 1$	DR	25000	10000	1096	1h06m
$n = 45$	1-1	$h = 1$	DR	15000	10000	1411	3h02m

A.5 Further results

Here we present some results on alternative model specifications: where $n = 15$ or $n = 45$. These models only include the industrial production and unemployment aggregate variables besides the 13 and 43 sectoral time series respectively. There are no major differences in results compared to the main models in Section 5, so we only include a small sample of specifications here: ranking restrictions are derived from the direct requirements matrix and the restrictions are imposed on horizon $h = 1$ in both cases. We conclude that results are robust with respect to the set of aggregate variables included in the model.

Figure A1: VAR(3), $n = 15$, $h = 1$, Direct Requirements (1–2)

Impulse response functions of sectoral shocks on Industrial production [Panel (a)] and Unemployment [Panel (b)] up to a 12-month horizon in a $n = 15$ VAR(3) model. The identified sectoral shocks are Machinery, Computer and electronic products, Food and Imports. The restrictions are imposed based on the direct requirements matrix at the 1-month horizon, following the 1–2 restriction structure: the impact of the sectoral shock is largest on the top 1 downstream receiving industry, while the second and third largest need not have strict ranking. The blue area is the equal-tailed 68% credible set and the black line is the median of the posterior distribution of IRFs.

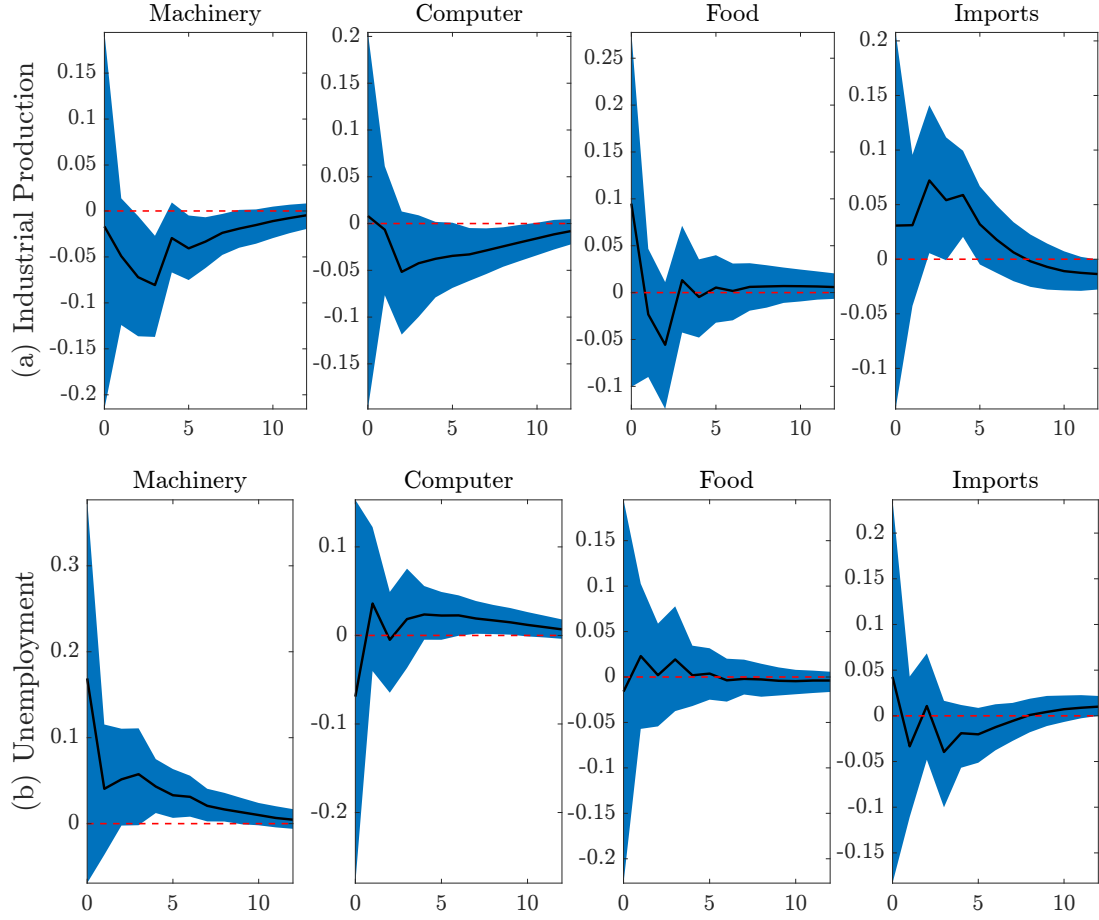
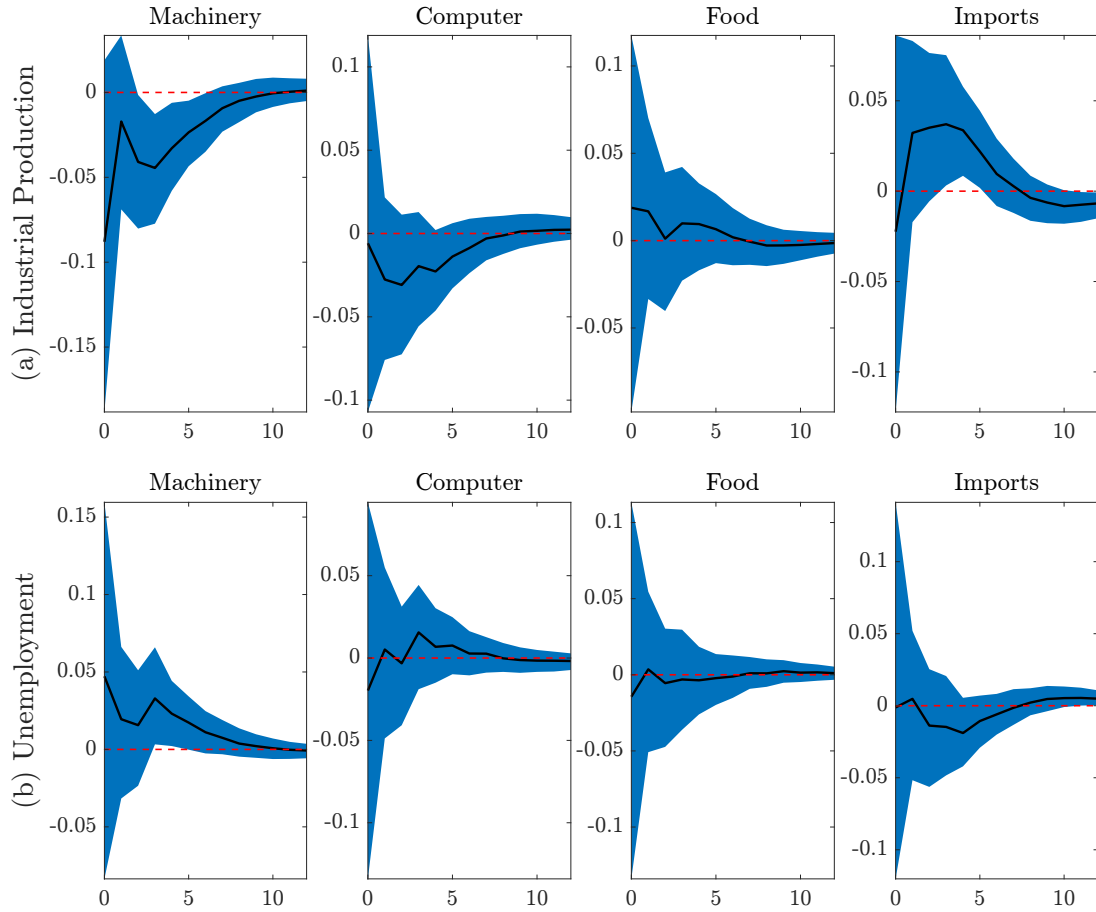


Figure A2: VAR(3), $n = 45$, $h = 1$, Direct Requirements (2-0)

Impulse response functions of sectoral shocks on Industrial production [Panel (a)] and Unemployment [Panel (b)] up to a 12-month horizon in a $n = 45$ VAR(3) model. The identified sectoral shocks are Machinery, Computer and electronic products, Food and Imports. The restrictions are imposed based on the direct requirements matrix on a 1-month horizon, following the 2-0 restriction structure: the impact of the sectoral shock is largest on the top 2 downstream receiving industries with strict ranking. The blue area is the equal-tailed 68% credible set and the black line is the median of the posterior distribution of IRFs.



A.6 Data sources and IO-NAICS matching

Table A2 shows the list of the 42 final matches between the input-output industry categories and NAICS codes, with the corresponding titles of the PPI time series obtained from the BEA. All of these variables were downloaded in index unit of measurement and were transformed by taking log differences.

Table A2: Correspondences found between input-output industry categories and NAICS producer price index series. The third and fifth columns show the category names for both types respectively. The last column is an indicator value that shows which industries were included in smaller models ($n = 15$ and $n = 20$).

ind.	IO	IO Industry Category	PPI	NAICS Category	Small
1	211	Oil and gas extraction	211	Oil and gas extraction	✓
2	212	Mining, except oil and gas	212	Mining (except oil and gas)	
3	22	Utilities	221	Utilities	
4	321	Wood products	321	Wood product manufacturing	✓
5	327	Nonmetallic mineral products	327	Nonmetallic mineral product manufacturing	
6	331	Primary metals	331	Primary metal mfg	✓
7	332	Fabricated metal products	332	Fabricated metal product mfg	✓
8	333	Machinery	333	Machinery manufacturing	✓
9	334	Computer and electronic products	334	Computer and electronic product mfg	✓
10	335	Electrical equipment, appliances, and components	335	Electrical equipment and appliance mfg	
11	3361MV	Motor vehicles, bodies and trailers, and parts	3361	Motor vehicle manufacturing	
12	3364OT	Other transportation equipment	3364	Aerospace product and parts manufacturing	
13	337	Furniture and related products	337	Furniture and related product mfg	
14	339	Miscellaneous manufacturing	339	Miscellaneous mfg	
15	311FT	Food and beverage and tobacco products	311	Food mfg	✓
16	313TT	Textile mills and textile product mills	313	Textile mills	
17	315AL	Apparel and leather and allied products	315	Apparel manufacturing	
18	322	Paper products	322	Paper manufacturing	
19	323	Printing and related support activities	323	Printing and related support activities	
20	324	Petroleum and coal products	324	Petroleum and coal products manufacturing	✓
21	325	Chemical products	325	Chemical mfg	
22	326	Plastics and rubber products	326	Plastics and rubber products mfg	
23	441	Motor vehicle and parts dealers	441	Motor vehicle and parts dealers	
24	445	Food and beverage stores	445	Food and beverage retailers	✓
25	4A0	Other retail	444	Building material and garden supply retailers	✓
26	481	Air transportation	481	Air transportation	✓
27	482	Rail transportation	482	Rail transportation	
28	483	Water transportation	483	Water transportation	
29	484	Truck transportation	484	Truck transportation	✓
30	487OS	Other transportation and support activities	488	Transportation support activities	
31	513	Broadcasting and telecommunications	517	Telecommunications	
32	514	Data processing, internet publishing, and other information services	5182	Data processing and related services	
33	521CI	Federal Reserve banks, credit intermediation, and related activities	5221	Depository credit intermediation	
34	523	Securities, commodity contracts, and investments	523	Security, commodity contracts and like activity	
35	524	Insurance carriers and related activities	524	Insurance carriers and related activities	
36	532RL	Rental and leasing services and lessors of intangible assets	5321	Automotive equipment rental and leasing	
37	5411	Legal services	5411	Legal services	
38	5412OP	Miscellaneous professional, scientific, and technical services	5413	Architectural, engineering and related services	
39	621	Offices of physicians	6211	Offices of physicians	
40	622	Hospitals	622	Hospitals	
41	623	Nursing and residential care facilities	6231	Nursing care facilities	
42	721	Accommodation	721	Accommodation	

The above table covers all 42 industry-specific PPI time series that were obtained from the Bureau of Economic Analysis. This is the set of sectoral time series we used in the large models ($n = 45$ or $n = 50$). The last column of the table indicates whether a particular sector is also included in smaller models ($n = 15$ and $n = 20$). All other variables: the import price index and 7 macro time series are listed in Table A3, where we indicate the source and code of the variables, along with their original unit of measurement and any additional transformations that were applied on these series.

Table A3: Summary of the remaining variables, including the import price index and 7 macro time series. Sources, units of measurement and transformations on these series are also reported.

Series title	Source	Series ID	Unit of measurement	Transformation
Monthly import price index for BEA End Use, All commodities	BEA/BLS	EIUIR	Index	Log diff.
Unemployment Level	FRED	UNEMPLOY	Thousands of persons	Log diff.
Industrial Production	FRED	INDPRO	Index	Log diff.
Consumer Price Index	FRED	CPIAUCSL	Index	Log diff.
5-Year, 5-Year Forward Inflation Expectation Rate	FRED	T5YIFRM	Change, percent	None
Federal Surplus or Deficit	FRED	MTSDS133FMS	Percent change	None
Personal Saving Rate	FRED	PSAVERT	Change, percent	None
Real Personal Consumption Expenditures	FRED	DPCERAM1M225NBEA	Percent change from preceding period	None