The Network Effects of Fiscal Adjustments (Web Appendix)

Edoardo Briganti,* Carlo A. Favero† and Madina Karamysheva‡

June 2018

Abstract

A large and increasing body of empirical evidence has established that fiscal adjustments based on government spending cuts are less costly in terms of losses in output growth than those based on tax increases. We show that the propagation of fiscal adjustment plans through the industrial network can in theory explain this evidence and that it does so in practice for the US economy. The heterogenous effects of tax-based and expenditure-based adjustments might depend on the difference in their propagation channels in the network of industries. A tax based adjustment plan is mainly a supply shock which propagates downstream (from supplier industries to customer industries) while an expenditure based plan is a demand shock which propagates upstream (from customer industries to supplier industries). Empirical investigation of these channels on US data based on Spatial Vector Autoregressions reveals that tax based plans propagate through the network with an average output multiplier of close to -2, while the propagation of expenditure based plans does not lead to any statistically significant effect on growth.

Keywords: industrial networks, fiscal adjustment plans, output. **JEL codes :** E60, E62.

*Edoardo Briganti: Bocconi University; edoardo.briganti@unibocconi.it

†Carlo A. Favero: Deutsche Bank Chair, IGIER-Bocconi, and CEPR; carlo.favero@unibocconi.it.

‡Madina Karamysheva: Assistant Professor NRU Higher School of Economics, Moscow; mkaramysheva@hse.ru

1 Web Appendix

1.1 The effect of Tax and Expenditure Adjustments: a simple illustration

Consider the simple case in which we have three industries. We describe the network by triplets $\{i, j, k\}$ where j is a supplier of industry i and k is a customer of industry i. The network structure is the following simplified one: $\{2, 3, 1\}, \{1, 2, 3\}, \{3, 1, 2\}$.

Assume that $u(c_1, c_2, c_3, l) = \gamma_l \prod_{i=1}^{3} c_i^{1/3}$. Sector's i production function is $y_i = e^{z_i} l_i^{\alpha_i^l} x_{ij}^{\alpha_{ij}}$. Also set $z_i = 0$ for $\forall i$

Market clearing condition for sector i is $y_i = c_i + x_{ki} + G_i$. Combining

$$a_{ij} = \frac{p_j x_{ij}}{p_i y_i}, \quad \alpha_i^l = \frac{w\,(1+\tau)\,l_i}{p_i y_i}$$

and

$$\frac{p_i c_i}{\beta_i} = \frac{p_j c_j}{\beta_j}$$

to eliminate prices we get

$$a_{ij} = \frac{c_i x_{ij}}{c_j y_i}$$

$$\alpha_i^l = \frac{w\,(1+\tau)\,l_i}{p_i y_i}$$

using the fact that

$$p_i c_i = \beta_i\,(wl - T)$$

get

$$\alpha_i^l = \frac{w\,(1+\tau)\,l_i c_i}{\beta_i\,(wl - T)\,y_i}$$

$w = 1, \beta_i = \beta = 1/3, l = 1$

$$\alpha_i^l = \frac{3\,(1+\tau)\,l_i c_i}{(1-T)\,y_i}$$

Substituting these expressions into the production function, we obtain

$$y_i = l_i^{\alpha_i^l} x_{ij}^{\alpha_{ij}} = \left(\frac{\alpha_i^l\,(1-T)\,y_i}{3\,(1+\tau)\,c_i}\right)^{\alpha_i^l} \left(\frac{\alpha_{ij} c_j y_i}{c_i}\right)^{\alpha_{ij}}$$

Taking into account the fact that $\alpha_i^l + \alpha_{ij} = 1$ for our example, simplify expression above to

$$c_i = \left(\frac{\alpha_i^l\,(1-T)}{3\,(1+\tau)}\right)^{\alpha_i^l} (\alpha_{ij} c_j)^{\alpha_{ij}} = \left(\frac{\alpha_i^l}{3}\right)^{\alpha_i^l} (\alpha_{ij})^{\alpha_{ij}} \left(\frac{(1-T)}{(1+\tau)}\right)^{\alpha_i^l} (c_j)^{\alpha_{ij}}$$

Let $\left(\frac{\alpha_i^l}{3}\right)^{\alpha_i^l} (\alpha_{ij})^{\alpha_{ij}} = \Omega_{ij}$, then

$$c_i = \Omega_{ij} \left(\frac{1-T}{1+\tau}\right)^{\alpha_i^l} c_j^{\alpha_{ij}}, i = 1, 2, 3$$

Solving simultaneously the three equations, we obtain:

$$c_i = \widetilde{\Omega}_i \left(\frac{1-T}{1+\tau}\right)^{\eta_i}$$

where $\widetilde{\Omega}_i$ - some constant and

$$\eta_i = \frac{\alpha_i^l + \alpha_j^l a_{i,j} + \alpha_k^l a_{ij} a_{jk}}{1 - a_{ij} a_{jk} a_{ki}}$$

Taking the log differential of expression for c_i we get

$$d \ln c_i = \eta_i \left[d \ln(1-T) - d \ln(1+\tau)\right]$$

using the fact that

$$\mathbf{d} \ln \mathbf{c} = \mathbf{d} \ln \mathbf{y}$$

$$d \ln y_i = \eta_i \left[d \ln(1-T) - d \ln(1+\tau)\right]$$

so

$$d \ln y_1 = \frac{\alpha_1^l + \alpha_2^l a_{1,2} + \alpha_3^l a_{12} a_{23}}{1 - a_{12} a_{23} a_{31}} \left[d \ln(1-T) - d \ln(1+\tau)\right]$$

$$d \ln y_2 = \frac{\alpha_2^l + \alpha_3^l a_{2,3} + \alpha_1^l a_{23} a_{31}}{1 - a_{23} a_{31} a_{12}} \left[d \ln(1-T) - d \ln(1+\tau)\right]$$

$$d \ln y_3 = \frac{\alpha_3^l + \alpha_1^l a_{3,1} + \alpha_2^l a_{31} a_{12}}{1 - a_{31} a_{12} a_{23}} \left[d \ln(1-T) - d \ln(1+\tau)\right]$$

Consider now the case of an Expenditure adjustments expressed in nominal terms as $d\widetilde{G}_1, d\widetilde{G}_2, d\widetilde{G}_3$. As in the case of Tax adjustments we set $\beta_1 = \beta_2 = \beta_3 = 1/3$. For simplicity we assume that governement expenditures are fully financed by lump-sum taxation and we set payroll tax to zero $\tau = 0$.

Utility function is $u(c_1, c_2, c_3, l) = \gamma_l \prod_{i=3}^{3} c_i^{1/3}$.Unit cost function can be written as:

$$C_i(p, w) = \mu_i w^{\alpha_i^l} p_j^{a_{ij}}$$

where $\mu_i = \left(\frac{\alpha_i^l}{a_{ij}}\right)^{a_{ij}} + \left(\frac{a_{ij}}{\alpha_i^l}\right)^{\alpha_i^l}$. In equilibrium we have

$$p_i = C_i(p, w) = \mu_i w^{\alpha_i^l} p_j^{a_{ij}}$$

Since $w = 1$we can solve the last equation for price

$$p_i = \gamma^{\overline{\frac{1}{1 - a_{ij} a_{jk} a_{ki}}}}$$

where $\gamma = \mu_i \mu_j^{a_{ij}} \mu_k^{a_{ij} a_{jk}}$. Taking into account the fact that prices do not respond to expenditre adjustments, we consider nominal values, denoted by $\tilde{}$.

$$d\widetilde{y}_i = d\widetilde{c}_i + a_{ki} d\widetilde{y}_k + d\widetilde{G_i}$$

From the household optimisation problem we have

$$\widetilde{c}_i = \frac{1}{(1+\lambda)3} - \frac{\widetilde{G}_i + \widetilde{G}_j + \widetilde{G}_k}{(1+\lambda)3}$$

Taking differential and combining it with resource constraint will leads to

$$d\widetilde{y}_i = -\frac{d\widetilde{G}_i + d\widetilde{G}_j + d\widetilde{G}_k}{(1+\lambda)3} + a_{ki} d\widetilde{y}_k + d\widetilde{G}_i, \forall i = 1, 2, 3$$

Solving this system of equations leads to

$$d\widetilde{y}_i = \frac{1}{1 - a_{ij} a_{jk} a_{ki}} \left\{ \begin{array}{c} d\widetilde{G}_i + a_{ki} a_{jk} d\widetilde{G}_j + a_{ki} d\widetilde{G}_k \\ -\frac{1 + a_{ki} + a_{ki} a_{jk}}{(1+\lambda)3} [d\widetilde{G}_i + d\widetilde{G}_j + d\widetilde{G}_k] \end{array} \right\}$$

1.2 Detailed derivations of the I-O matrix A

The Bureau of Economic Analysis (BEA) provides 4 requirement tables. In particular, we are interested in an industry by industry total requirememnt table.
The construction of the total equirement table is detailed over page 12-8 (page 8 of chapter 12) of the "Concepts and Methods of the U.S. Input-Output accounts" - a guide released by the BEA, which provides a full explanation of the industrial network data.
Consider a generic industry, say Z, whose total output is denoted with y. Since supply and demand must coincide, y is equal to F - final uses - plus x - demand from other industries which use the output of industry Z as input:

$$y = F + x.$$

Now we define the coefficient matrix A as:

$$A = \frac{x}{y},$$

that is, the share of industry Z output used as production input by the other industries.
Therefore, we have $x = A \cdot y$, and plugging it into the previous equation we have:

$$y = F + A \cdot y,$$

whose close-form solution is:

$$y = (I - A)^{-1} F = f(F).$$

In the I-O terminology used by the BEA, function f, which links final uses with the industry output, is called total requirement table. In economic theory we usually refer to such a transformation as the Leontief Inverse matrix.
In order to construct such a table, the BEA starts from storing raw data into two tables: the Make Table and the Use Table. The empirical counterpart of $(I - A)^{-1}$ is constructed in several steps, illustrated by the BEA guide.

The first step consists of reshaping the Use table, which is a non-symmetric commodity-by-industry table. The Use table shows the uses of commodities by intermediate and final users. Differently from the Make table, the rows in the Use table present the commodities or products, and the columns display the industries and final users that utilize them. The sum of the entries in a row is the output of that commodity. The columns show the products consumed by each industry and the three components of value added, compensation of employees, taxes on production and imports less subsidies, and gross operating surplus. Value added is the difference between an industry's output and the cost of its intermediate inputs. Total value added is equal to GDP. The sum of the entries in a column is that industry's output. We can derive the analytic form of the Use table, by introducing a specific terminolgy:

- INP_j^i = Commodity j used as input by industry i. This is the generic element of the Use table.
- SALES_j = Total output of industry j,

we rewrite the generic element of the Use table - assuming for simplicity that the number of commodities and industries is three ($n = 3$) - in this way:

$$\mathrm{USE} = \begin{bmatrix} \mathrm{INP}_1^1 & \mathrm{INP}_1^2 & \mathrm{INP}_1^3 \\ \\ \mathrm{INP}_2^1 & \mathrm{INP}_2^2 & \mathrm{INP}_2^3 \\ \\ \mathrm{INP}_3^1 & \mathrm{INP}_3^2 & \mathrm{INP}_3^3 \end{bmatrix}.$$

At this point we can derive a commodity by industry direct requirement table by dividing each industry's input by its corresponding total industry output. We denote such a matrix with letter B and we can express its generic element using the previous notation in this way:

$$B_{ij} = \frac{\mathrm{INP}_i^j}{\mathrm{SALES}_j},$$

where i denotes the row and j the column of matrix B. Therefore, the analytic form of matrix B is :

$$\mathrm{B} = \begin{bmatrix} \frac{\mathrm{INP}_1^1}{\mathrm{SALES}_1} & \frac{\mathrm{INPUT}_1^2}{\mathrm{SALES}\ _2} & \frac{\mathrm{INP}_1^3}{\mathrm{SALES}_3} \\ \\ \frac{\mathrm{INP}_2^1}{\mathrm{SALES}_1} & \frac{\mathrm{INPUT}_2^2}{\mathrm{SALES}\ _2} & \frac{\mathrm{INP}_2^3}{\mathrm{SALES}_3} \\ \\ \frac{\mathrm{INP}_3^1}{\mathrm{SALES}_1} & \frac{\mathrm{INPUT}_3^2}{\mathrm{SALES}\ _2} & \frac{\mathrm{INP}_3^3}{\mathrm{SALES}_3} \end{bmatrix}.$$

The BEA guide provides also a numerical example - with 3 industries ($n = 3$) - which we report here for the sake of clarity:

	1	2	3	Final demand	Total Commodity Output
1	50	120	120	40	330
2	180	30	60	130	400
3	50	150	50	20	270
Scrap	1	3	1	0	5
VA	47	109	34	/	190
Total Industry Output	328	412	265	190	/

Consider the first row: 50 units of commodity 1 are used by industry 1, 120 are used by industry 2 and 120 are used by industry 3; 40 units of commodity 1 are demanded as final product, therefore, the overall production of commodity 1 amounts to 50 plus 120 plus 120 plus 40: 330 units.
At the same time, we can derive the direct requirement table by following the instructions explained above:

	1	2	3
1	0.152	0.291	0.453
2	0.549	0.073	0.226
3	0.152	0.364	0.189
Scrap	0.003	0.007	0.004
VA	0.143	0.265	0.128
Total	1	1	1

The first element of the first row is obtained by dividing 50 by 328, for instance. The second element of the first row is obtained dividing 120 by 412 and so on and so forth.
By removing scrap and value added from the above table, we obtain a symmetric commodity-by-industry matrix, denoted with B, whose generic elements are described above:

$$B = \begin{bmatrix} 0.152 & 0.291 & 0.453 \\ 0.549 & 0.073 & 0.226 \\ 0.152 & 0.364 & 0.189 \\ \end{bmatrix}.$$

At this point put aside for a while the direct requirement matrix just derived, and

focus on the Make table, which shows the production of commodities by industries. The rows present the industries, and the columns display the commodities that the industries produce. Looking across a row, all the commodities produced by that industry are identified, and the sum of the entries is that industry's output. Looking down a column, all the industries producing that commodity are identified, and the sum of the entries is the output of that commodity.
As we did previously, we now introduce a useful notation, which allows to better interpret what we are computing:

- Y_j = Total production of commodity j.
- OUT_j^i = Commodity j produced by industry i
- NSR_i^{-1} = The inverse of the non-scrap ratio of industry i,

The analytical form of the Make table is the following:

$$\mathrm{MAKE} = \begin{bmatrix} \mathrm{OUT}_1^1 & \mathrm{OUT}_2^1 & \mathrm{OUT}_3^1 \\ \\ \mathrm{OUT}_1^2 & \mathrm{OUT}_2^2 & \mathrm{OUT}_3^2 \\ \\ \mathrm{OUT}_1^3 & \mathrm{OUT}_2^3 & \mathrm{OUT}_3^3 \end{bmatrix}$$

At this point, we divide each row for the total commodity output to obtain the

market share matrix, which shows the proportion of commodity output produced by each industry, whose analytical form is the following:

$$\mathrm{MS} = \begin{bmatrix} \dfrac{\mathrm{OUT}_1^1}{Y_1} & \dfrac{\mathrm{OUT}_2^1}{Y_2} & \dfrac{\mathrm{OUT}_1^3}{Y_3} \\ \\ \dfrac{\mathrm{OUT}_1^2}{Y_1} & \dfrac{\mathrm{OUT}_2^2}{Y_2} & \dfrac{\mathrm{OUT}_3^2}{Y_3} \\ \\ \dfrac{\mathrm{OUT}_1^3}{Y_1} & \dfrac{\mathrm{OUT}_2^3}{Y_2} & \dfrac{\mathrm{OUT}_3^3}{Y_3} \end{bmatrix}$$

Again, we show a sample Make table, with $n = 3$:

	1	2	3	Scrap	Total Industry Output
1	300	25	0	3	328
2	30	360	20	2	4412
3	0	15	250	0	265
Total Commodity Output	330	400	270	5	/

Consider the first column which corresponds to industry 1 output: industry 1 makes 300 of commodity 1, 30 of commodity 2 and it does not produce commodity 3. Overall, industry 1 makes 300 plus 30 plus 0, 330 of total commodity output. Following the instructions described above, we derive the market share table:

	1	2	3
1	0.909	0.063	0
2	0.091	0.900	0.074
3	0	0.038	0.926
Total Commodity Output	1	1	1

The third step is to make adjustments for scrap. The I-O accounts include a commodity for scrap, which is a byproduct of industry production. No industry produces scrap on demand; rather, it is the result of production to meet other demands. In order to make the I-O model work correctly - that is, not requiring industry output because of a demand for scrap inputs- we have to eliminate scrap as a secondary product. At the same time, we must also keep industry output at the same level. This adjustment is accomplished by calculating the ratio of nonscrap output to industry output for each industry and then applying these ratios to the market shares matrix in order to account for total industry output. More precisely, the non-scrap ratio is defined as follows:

$$(\text{Non-scrap ratio})_i = \frac{\text{Industry i output}}{\text{Industry i output - scrap i}} = \text{NSR}_i$$

Therefore, using the numbers from the previous example, we have:

	Tot.Ind.Out.	Scrap	Δ	Non-Scrap Ratio
1	328	3	325	0.991
2	412	2	410	0.995
3	265	0	265	1

The market shares matrix is adjusted for scrap by dividing each row coefficient by the non-scrap ratio for that industry. In the resulting transformation matrix, called W, the implicit commodity output of each industry has been increased.
We might write the generic element of the adjusted market share matrix W in this way:

$$(\text{Market share adjusted})_{ij} = W_{ij} = \frac{\mathrm{OUT}^i_j \cdot NSR_i^{-1}}{Y_j},$$

whose analytical form is:

$$\mathrm{W} = \begin{bmatrix} \frac{\mathrm{OUT}^1_1 \cdot NSR_1^{-1}}{Y_1} & \frac{\mathrm{OUT}^1_2 \cdot NSR_1^{-1}}{Y_2} & \frac{\mathrm{OUT}^1_3 \cdot NSR_1^{-1}}{Y_3} \\ \frac{\mathrm{OUT}^2_1 \cdot NSR_2^{-1}}{Y_1} & \frac{\mathrm{OUT}^2_2 \cdot NSR_2^{-1}}{Y_2} & \frac{\mathrm{OUT}^2_3 \cdot NSR_2^{-1}}{Y_3} \\ \frac{\mathrm{OUT}^3_1 \cdot NSR_3^{-1}}{Y_1} & \frac{\mathrm{OUT}^3_2 \cdot NSR_3^{-1}}{Y_2} & \frac{\mathrm{OUT}^3_3 \cdot NSR_3^{-1}}{Y_3} \end{bmatrix}$$

The resulting transformation matrix W of our example is:

$$W = \begin{bmatrix} 0.917 & 0.063 & 0 \\ 0.091 & 0.904 & 0.074 \\ 0 & 0.038 & 0.926 \\ & & \end{bmatrix}.$$

We now have all the elements to compute a symmetric direct requirement table. Recall now that the transformation matrix is an industry by commodity table, while the direct requirement table B is a commodity by industry table. Therefore, by multiplying them, we can construct a symmetric industry-by-industry direct requirement table, denoted with WB.

$$WB = \begin{bmatrix} \dfrac{\text{OUT}_1^1 \cdot NSR_1^{-1}}{Y_1} & \dfrac{\text{OUT}_2^1 \cdot NSR_1^{-1}}{Y_2} & \dfrac{\text{OUT}_3^1 \cdot NSR_1^{-1}}{Y_3} \\ \\ \dfrac{\text{OUT}_1^2 \cdot NSR_2^{-1}}{Y_1} & \dfrac{\text{OUT}_2^2 \cdot NSR_2^{-1}}{Y_2} & \dfrac{\text{OUT}_3^2 \cdot NSR_2^{-1}}{Y_3} \\ \\ \dfrac{\text{OUT}_1^3 \cdot NSR_3^{-1}}{Y_1} & \dfrac{\text{OUT}_2^3 \cdot NSR_3^{-1}}{Y_2} & \dfrac{\text{OUT}_3^3 \cdot NSR_3^{-1}}{Y_3} \end{bmatrix} \cdot \begin{bmatrix} \dfrac{\text{INP}_1^1}{\text{SALES}_1} & \dfrac{\text{INPUT}_1^2}{\text{SALES }_2} & \dfrac{\text{INP}_1^3}{\text{SALES}_3} \\ \\ \dfrac{\text{INP}_2^1}{\text{SALES}_1} & \dfrac{\text{INPUT}_2^2}{\text{SALES }_2} & \dfrac{\text{INP}_2^3}{\text{SALES}_3} \\ \\ \dfrac{\text{INP}_3^1}{\text{SALES}_1} & \dfrac{\text{INPUT}_3^2}{\text{SALES }_2} & \dfrac{\text{INP}_3^3}{\text{SALES}_3} \end{bmatrix}$$

The generic element of matrix WB is the following:

$$WB_{ij} = \sum_{s=1}^{3} W_{is} \cdot B_{sj}$$

For instance let's derive the analytic form of the second element of the first row:

$$WB_{12} = \frac{\dfrac{\text{OUT}_1^1 \cdot NSR_1^{-1}}{Y_1} \cdot \ \text{INP}_1^2 + \dfrac{\text{OUT}_2^1 \cdot NSR_1^{-1}}{Y_2} \cdot \text{INP}_2^2 + \dfrac{\text{OUT}_3^1 \cdot NSR_1^{-1}}{Y_3} \cdot \text{INP}_3^2}{SALES_2}$$

$$= \frac{\text{SALES}_{1\to 2}}{\text{SALES}_2}.$$

Again, let's derive another element: WB_{21}:

$$WB_{21} = \frac{\dfrac{\text{OUT}_1^2 \cdot NSR_2^{-1}}{Y_1} \cdot \ \text{INP}_1^1 + \dfrac{\text{OUT}_2^2 \cdot NSR_2^{-1}}{Y_2} \cdot \text{INP}_2^1 + \dfrac{\text{OUT}_3^2 \cdot NSR_2^{-1}}{Y_3} \cdot \text{INP}_3^1}{SALES_1}$$

$$= \frac{\text{SALES}_{2\to 1}}{\text{SALES}_1}.$$

Therefore, the analytic form of matrix WB is:

$$WB = \begin{bmatrix} \frac{\text{SALES}_{1\to1}}{\text{SALES}_1} & \frac{\text{SALES}_{1\to2}}{\text{SALES}_2} & \frac{\text{SALES}_{1\to3}}{\text{SALES}_3} \\ \\ \frac{\text{SALES}_{2\to1}}{\text{SALES}_1} & \frac{\text{SALES}_{2\to2}}{\text{SALES}_2} & \frac{\text{SALES}_{2\to3}}{\text{SALES}_3} \\ \\ \frac{\text{SALES}_{3\to1}}{\text{SALES}_1} & \frac{\text{SALES}_{3\to2}}{\text{SALES}_2} & \frac{\text{SALES}_{3\to3}}{\text{SALES}_3} \end{bmatrix}.$$

Notice that this matrix coincide with our theoretical matrix A transposed. Recall infact that from the profit maximization problem we obtained:

$$a_{ij} = \frac{p_j \cdot x_{ij}}{p_i \cdot y_i} = \frac{\text{SALES}_{j\to i}}{\text{SALES}_i}$$

where x_{ij} is the quantity of good employed by sector i and supplied by industry j (as usual, i is the number of the row while j the number of the column):

$$A = \begin{bmatrix} a_{11} = \frac{\text{SALES}_{1\to1}}{\text{SALES}_1} & a_{12} = \frac{\text{SALES}_{2\to1}}{\text{SALES}_1} & a_{13} = \frac{\text{SALES}_{3\to1}}{\text{SALES}_1} \\ \\ a_{21} = \frac{\text{SALES}_{1\to2}}{\text{SALES}_2} & a_{22} = \frac{\text{SALES}_{2\to2}}{\text{SALES}_2} & a_{23} = \frac{\text{SALES}_{3\to2}}{\text{SALES}_2} \\ \\ a_{31} = \frac{\text{SALES}_{1\to3}}{\text{SALES}_3} & a_{32} = \frac{\text{SALES}_{2\to3}}{\text{SALES}_3} & a_{33} = \frac{\text{SALES}_{3\to3}}{\text{SALES}_3} \end{bmatrix}.$$

Therefore, the following identity is true:

$$A = (WB)^T.$$

The above identity is crucial, since it shows that there is a discrepancy between the theoretical Input-Output matrix and the empirical one. We have to pay lot of attention when working with these data in order not to forget to take the transpose of the empirical matrix, in order to replicate what theory suggests.

At this point, we can finally derive the ultimate table, which is the one available for downloading on the BEA website: the total requirement table industry-by-industry. In the BEA guide they provide computations for obtaining the commodity-by-commodity total requirement table (they do B times W rather than W times B), however at page 24 of chapter 12 they show the formulas they employ to derive the industry-by-industry total requirement table, which we indicate with TR:

$$TR = (I - WB)^{-1},$$

which shows the industry output required per dollar of each industry product delivered to final users.
At this point a little of matrix algebra turns out useful. Consider a $n \times n$ matrix $S = I_n - A$. Since the transpose of an invertible matrix is also invertible, and its inverse is the transpose of the inverse of the original matrix, we can write the following:

$$(S^T)^{-1} = (S^{-1})^T.$$

Morevoer, the following identity holds:

$$S^T = (I - A)^T = I - A^T,$$

then:

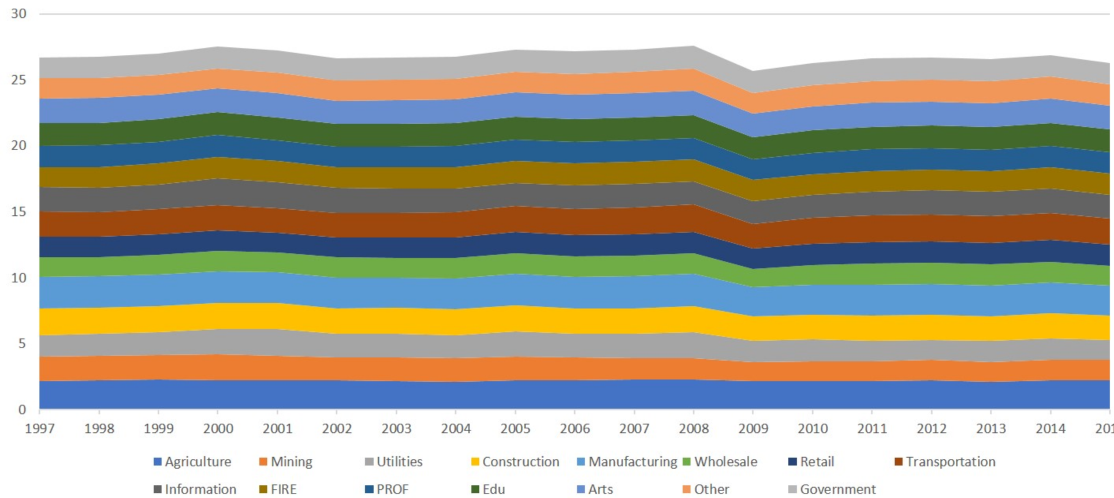
$$\begin{aligned} TR &= (I - WB)^{-1} \\ &= (I - A^T)^{-1} \\ &= (S^T)^{-1} \\ &= (S^{-1})^T \\ &= \left((I - A)^{-1}\right)^T \\ &= H^T \end{aligned}$$

Matrix $H = (I - A)^{-1}$ is exaclty our theoretical Leontief inverse matrix: the industry-by-industry total requirement table available for free-downloading on the BEA website, coincide with the **transposed** Leontief inverse matrix. At this point, we are able to define a transformation which allows us to pass from the row data to the empirical counterpart of the theoretical I-O matrix A:

$$A = f(TR_{BEA}) = I_n - \left[(TR_{BEA})^T\right]^{-1}, \tag{1}$$

where, A is therefore a function of the row data TR_{BEA}, the 15 industry-by-industry total requirement table. By taking the transpose of the original table the empirical counterpart of matrix A is obtained.

The last issue to discuss is the following: the industry-by-industry total requirement table's spreadsheet, contains 19 tables, one per each year since 1997 to 2015 (estimates are yearly updated). Which one to use?
We choose the midsample table of table of 1997.
to check for the potentialrelevance of this choice we computed total requirement for every industry and year, by simply summing up the columns of the total requirement tables, and then we looked at the evolution of these values over time. Results are shown in the figure below:



The values are substantially stable over time pointing to stable connections among industries over time.

1.3 Spatial variables and the Cobb-Douglas production function

Apatial variables are constructed consistently with the Cobb-Douglas production functions.
In Section 4 we defined the spatial varaibles in this way:

$$\Delta y_{i,t}^{up} = \sum_{j\neq i}^{n} \hat{a}_{ji} \cdot \Delta y_{j,t}$$

$$\Delta y_{it}^{down} = \sum_{j\neq i}^{n} a_{ij} \cdot \Delta y_{j,t}.$$

Where $\Delta y_{j,t}$ accounts for the percent growth rate of real value added of industry i recorded in year t.
For example the global variables of industry 1 are:

$$\Delta y_{1,t}^{up} = \frac{Sales_{1\to 2} \cdot \Delta y_{2,t} + \cdots + Sales_{1\to n} \cdot \Delta y_{n,t}}{Sales_1}$$

$$\Delta y_{1,t}^{down} = \frac{Sales_{2\to 1} \cdot \Delta y_{2,t} + \cdots + Sales_{n\to 1} \cdot \Delta y_{n,t}}{Sales_1}$$

They can be interpreted as a weighted average of the real value added of other industries, with weights given by the relative importance of every industry as a customer (Y^{up}) or a supplier (Y^{down}) to industry i.
Why not expressing the global variables as the percent change of a linear combination rather than a linear combination of percentages? We could have expressed the global variables in this alternative way:

$$\Delta y_{i,t}^{up} = \Delta\Big(\sum_{j\neq i}^{n} \hat{a}_{ji} \cdot Y_{jt}\Big)$$

$$\Delta y_{i,t}^{down} = \Delta\Big(\sum_{j\neq i}^{n} a_{ij} \cdot Y_{jt}\Big).$$

At this point we might link every spatial variable to its corresponding dependent variable:

$$\Delta y_{i,t} = \beta^{up} \cdot \Delta y_{i,t}^{up} + \beta^{down} \cdot \Delta y_{i,t}^{down},$$

taking a logarithmic approximation we have:

$$\frac{\partial}{\partial y_{i,t}} \ln y_{i,t} = \beta^{up} \cdot \frac{\partial}{\partial y_{i,t}} \ln y_{i,t}^{up} + \beta^{down} \cdot \frac{\partial}{\partial y_{i,t}} \ln y_{i,t}^{down},$$

by integrating both terms we have:

$$\begin{aligned}
\ln y_{i,t} &= \beta^{up} \cdot \ln \Big(\sum_{j \neq i}^{n} \hat{a}_{ji} \cdot Y_{jt} \Big) + \beta^{down} \cdot \ln \Big(\sum_{j \neq i}^{n} a_{ij} \cdot Y_{jt} \Big) \\
&= \ln \left(\Big(\sum_{j \neq i}^{n} \hat{a}_{ji} \cdot Y_{jt} \Big)^{\beta^{up}} \cdot \Big(\sum_{j \neq i}^{n} a_{ij} \cdot Y_{jt} \Big)^{\beta^{down}} \right),
\end{aligned}$$

by elevating both sides, we finally obtain:

$$y_{i,t} = \Big(\sum_{j \neq i}^{n} \hat{a}_{ji} \cdot Y_{jt} \Big)^{\beta^{up}} \cdot \Big(\sum_{j \neq i}^{n} a_{ij} \cdot Y_{jt} \Big)^{\beta^{down}}.$$

The last expression is telling us that the industries are linked among themselves thorugh a relationship which has nothing to do with a Cobb-Douglas production function.

Consider now the definition we employ in the paper:

$$\begin{aligned}
\Delta y_{i,t} &= \beta^{up} \cdot \delta y_{i,t}^{up} + \beta^{down} \cdot \delta y_{i,t}^{down} \\
&= \beta^{up} \cdot \Big(\sum_{j \neq i}^{n} \hat{a}_{ji} \Delta y_{j,t} \Big) + \beta^{down} \cdot \Big(\sum_{j \neq i}^{n} a_{ij} \Delta y_{j,t} \Big),
\end{aligned}$$

substituting again the percent changes with the logarithmic approximation, we have:

$$\frac{\partial}{\partial y_{i,t}} \ln y_{it} = \beta^{up} \cdot \Big(\sum_{j \neq i}^{n} \hat{a}_{ji} \cdot \frac{\partial}{\partial y_{j,t}} \ln y_{jt} \Big) + \beta^{down} \cdot \Big(\sum_{j \neq i}^{n} a_{ij} \cdot \frac{\partial}{\partial y_{j,t}} \ln y_{j,t} \Big),$$

by integrating both terms we have:

$$\begin{aligned}
\ln y_{it} &= \beta^{up} \cdot \Big(\sum_{j \neq i}^{n} \hat{a}_{ji} \cdot \ln y_{jt} \Big) + \beta^{down} \cdot \Big(\sum_{j \neq i}^{n} a_{ij} \cdot \ln y_{j,t} \Big) \\
&= \sum_{j \neq i}^{n} \underbrace{(\beta^{up} \cdot \hat{a}_{ji} + \beta^{down} \cdot a_{ij})}_{\phi_{ij}} \cdot \ln y_{j,t} \\
&= \ln \left(\prod_{j \neq i}^{n} y_{jt}^{\phi_{ij}} \right),
\end{aligned}$$

by removing the natural logarithm from both sides, we have:

$$y_{i,t} = \prod_{j \neq i}^{n} y_{j,t}^{\phi_{ij}}$$

The above expression closely reflects a Cobb-Douglas production function, thus providing a theoretical justification for the way we construct spatial variables.

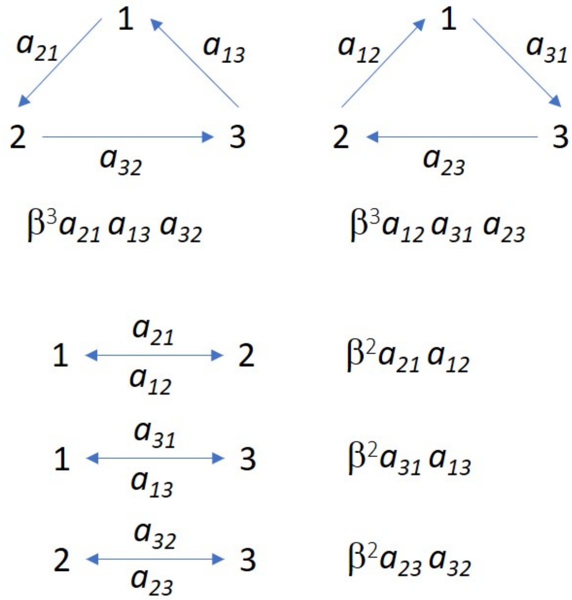
1.4 Interpretation of the total, direct and indirect effects

In this section we explain in details what the determinant d of matix $(1 - \beta \cdot A_0)$ represents from an economic point of view.[1]
First of all, notice that the reciprocal of the determinant can be interpreted as the convergence point of a geometric summation:

$$\frac{1}{d} = \sum_{i=1}^{\infty} \Big(\beta^2 \cdot \big(a_{12}a_{21} + a_{13}a_{31} + a_{32}a_{23}\big) + \beta^3 \cdot \big(a_{12}a_{23}a_{31} + a_{21}a_{13}a_{32}\big)\Big)^i = \sum_{i=1}^{\infty} K^i.$$

Where $K = \beta^2 \cdot (a_{12}a_{21} + a_{13}a_{31} + a_{32}a_{23}\big) + \beta^3 \cdot \big(a_{12}a_{23}a_{31} + a_{21}a_{13}a_{32})$.
Basically, suppose that a little change in a sector's output occurs, then, such a change triggers a cascade effect of other changes in other sectors and then come back to it via the input-output network. If we imagined a kind of temporal-sequentiality in transferring such a shock, we would see that at every step the change that occurs is exactly K: a unit change occurs, then this change is transferred to other sectors and come back, by generating a total change of K, but then this change is transferred again and comes back to its initial point thus generating a further change of K^2 and so on and so forth. The limit point of such a series is exactly the reciprocal of the determinant of matrix $I_3 - \mathbf{T}$. Infact, notice that every element of K represents a particular sectors' relationship, as it is represented in the figure below:

[1] To ease notation we write simply β instead of β^{down}.

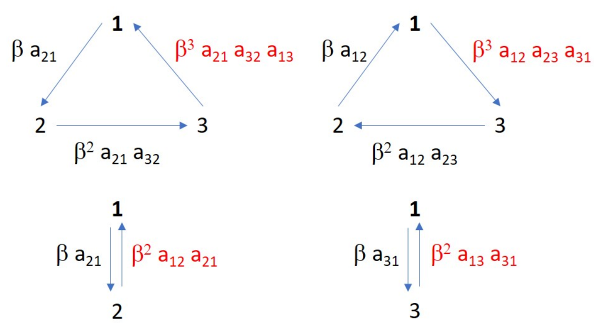


Now, all these changes actually occurs simultaneously, therefore, every variation is subject to an amplification due to the sectors' interconnections. Such an amplification is exactly expressed by the reciprocal of d, which can be seen as a "simultaneity multiplier".
At this point, we can examine more carefully all the elements of $\mathbf{S} = (I - \beta \cdot A_0)^{-1}$. In particular we can make the following distinction:

1. **Direct effect**: the direct effect to sector i of a macro shock δ ($DE_{\delta,i}$), is the change in sector i growth rate as if it would be the only sector in the economy subject to that shock:
$$DE_{\delta,i} = \frac{1}{d} \cdot \mathbf{S}_{ii} \cdot \delta.$$
Basically, the direct effects are collected on the main diagonal of the previous matrix. For instance, the direct effect of a shock to sector 1 is:
$$DE_{\delta,1} = \delta + \frac{\beta^2 \cdot (a_{31}a_{13} + a_{21}a_{12}) + \beta^3 \cdot (a_{12}a_{23}a_{31} + a_{21}a_{13}a_{32})}{d}\delta.$$
Notice that the direct effect could be decomposed into two parts: the shock itself (δ) plus the network effect that such a shock triggers, amplified by the simultaneity multiplier. Notice infact that the instantaneous effect of a direct shock to only sector 1, is transferred on sector 2 and then goes back to sector

1, for a total change of $\beta^2 \cdot a_{21}a_{12}$. The same is true for sector 3 for a total change of $\beta^2 \cdot a_{31}a_{13}$. Moreover, the shock, once transferred upon sectors 2 and 3, it can get back to sector 1 inderectly via the connections between sector 2 and sector 3, for a total change of $\beta^3 a_{21}a_{32}a_{13}$ when the shock flows down to sector 2 and $\beta^3 a_{31}a_{23}a_{12}$ when the shock flows down to sector 3. This scheme is represented also in the figure below:

Downstream propagation of a shock to sector 1



2. **Indirect effect**: The indirect effect to sector i of a macro shock δ, $(IE_{\delta,i})$, is the summation of all the effects of shocks which hit other sectors and then are transferred to sector i:

$$IE_{\delta,i} = \frac{1}{d} \cdot \left(\sum_{j \neq i}^{n} \mathbf{S}_{ij} \right) \cdot \delta.$$

For instance, the indirect effect of a macro shock on sector 1 is:

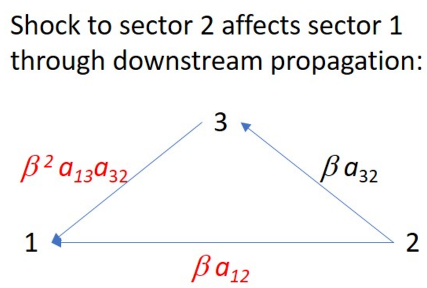
$$IE_{\delta,1} = \frac{1}{d} \cdot \Big[(\beta \cdot a_{12} + \beta^2 \cdot a_{32}a_{13}) + (\beta \cdot a_{13} + \beta^2 \cdot a_{12}a_{23}) \Big] \cdot \delta$$

$$= \frac{\beta \cdot (a_{12} + a_{13}) + \beta^2 \cdot (a_{32}a_{13} + a_{12}a_{23})}{d} \cdot \delta$$

Consider now the generic element of matrix $\mathbf{S}$, which we call $\mathbf{S}_{ij}$. Such an element provides the specific impact of a shock which hits directly sector j and then is transferred to sector i. Such an effect could be interpreted as an instantaneous effect which is then amplified via the simultaneity multiplier.
To better understand this point, consider for instance the following element:

$$\mathbf{S}_{12} = \beta \cdot a_{12} + \beta^2 \cdot a_{32}a_{13}.$$

This element provides the "instantaneous" effect of a shock which hits sector 2 but then is transferred upon sector 1 (for this reason we call it "indirect"). Such a transfer occurs through the direct linkage between sector 2 and 1 ($\beta \cdot a_{12}$, sales of sector 2 to sector 1) and through the indirect connections via sector 3 ($\beta^2 \cdot a_{32}a_{13}$, that is, the sales of sector 2 to sector 3 and then the sales of sector 3 to sector 1). This propagation is shown in the figure below:



At this point we can show what the total, direct and indirect effects are defined:

1. The direct effect of a shock is not equal for every sector, since it depends on the network interactions specific for that sector. For this reason, we computed the effect on the economy, which is a weighted average of the single industries' effects, where weights are collected into the $n \times 1$ vector W (every industry weight is the relative size of the sector with resepct to GDP):

$$DE_\delta = \frac{\delta}{d} \cdot diag(\mathbf{S}) \cdot W.$$

2. The total effect of a macro shock is by consequence the summation of the direct and indirect effect. Formally:

$$TE_{\delta,i} = \frac{\delta}{d} \cdot \Big(\sum_{j=1}^{n} \mathbf{S}_{ij} \Big).$$

Again, we report the formula for the total effect on the economy:

$$TE_\delta = \frac{\delta}{d} \cdot \mathbf{S} \cdot W.$$

3. The average indirect effect will be computed indirectly by taking the difference between the total and the direct effect:

$$IE_\delta = TE_\delta - DE_\delta.$$

4. The same things will be replicated identically for the expenditure shock ($EB_t = 1$):
$$ADE_\gamma = \frac{\gamma}{|I_n - \Gamma|} \cdot \frac{1}{n} \cdot Tr(S(\mathbf{\Gamma})).$$

$$ATE_\gamma = \frac{\gamma}{|I_n - \Gamma|} \cdot \frac{1}{n} \cdot \mathbf{1}' \cdot S(\mathbf{\Gamma}) \cdot \mathbf{1}.$$

$$AIE_\gamma = ATE_\gamma - ADE_\gamma$$

1.5 Details on Database of Exogenous Fiscal Adjustment Plans

For the purpose of this project we follow Alesina et al.(2015) and use the annual fiscal adjustment plans data. Our focus is US. We adopt the same data employed by Alesina et Al.(2015), however we do several modifications. The reason for such a discrepancy can be explained by the fact that we deal in a different way with long-run adjustments. In fact, Romer and Romer(2010) distinguish between deficit driven and long-run growth driven adjustments. Long-run driven adjustments can be both positive and negative. In order to take them into account we follow the rule: sum up positive and negative components of long-run growth driven adjustments together with deficit driven adjustments and include the sum into the database if and only if it is non negative[2].

Moreover, it is worth to notice two years of Reagan presidency. The rule described above leads to drop from the sample the deficit-driven adjustment implemented in the US in years 1983-84 because it was smaller than the contemporaneous negative long-run growth-driven adjustment.

Slight modifications in years 1980 and 1981 are due to the same logic. In 1980 we include the positive long-run growth tax increases[3]. In 1981 following the rule we consider the sum of the deficit - driven tax hike and long-run growth driven tax decrease.

Other slight modifications, consistent with the previous reasoning, are in years 1985, 1986, 1990. We record initial announcement of the Social Security Amendment 1983 as in Alesina et al.(2015) in the announced part of the plan, however additionally we record revisions to already announced adjustments for the years 1985, 1986, 1990 as a surprise component[4]. Revisions result in further austerity.

Overall, the differences between Alesina et al.(2015) dataset and our dataset are minimal. Importantly, following Alesina et al.(2015), we scale all the measures by GDP on the year prior to the consolidation in order to avoid potential endogeneity issues[5].

To illustrate the procedure of fiscal plan construction consider the case of 1990

[2] Modifications are light because positive long-run driven adjustments, that is tax increase due to long-run growth reasons, are very uncommon.

[3] In 1980 the Crude Oil Windfall Profit Tax Act was signed. It is scheduled as a series of tax increases. However, such reforms were not due to deficit driven reason but for long-run growth reasons.

[4] Budgets 1985, 1987, 1989, 1991 provide revision estimates.

[5] Romer and Romer 2010 scale their fiscal shocks by the nominal GDP in the year at the time of the change

OBRA (Omnibus Budget Reconciliation Act) - 1990, which is considered as exclusively motivated by a deficit reduction motive and therefore exogenous for the estimation of the output effect of fiscal corrections[6].

Insert Table I here

Table I illustrates how the plan is reclassified by DeVries et al and R&R using different sources. OBRA - 1990 plans fiscal adjustment both on revenue and expenditure side over the period 1991-1995. R&R concentrate only on the revenue adjustment and lump in the first quarter of 1991 all the relevant adjustment (that therefore adds up adjustment to be implemented in 1991 and 1992), the post 1992 adjustment are not included because of their small size. "...almost all the revenue provisions were effective January 1, 1991. Thus the first full fiscal year the changes were scheduled to be in effect was fiscal 1992. We therefore use the estimated revenue effect from the Budget for that year as our revenue estimate. That is, we estimate that there was a tax increase of $35.2 billion in 1991Q1..." Devries et al. 2013 after the reclassification from fiscal to calendar year, use the implementation rather than the announcement as a criterion to attribute shocks to each period[7].

Table II illustrates reclassification of shocks in to the fiscal adjustment plans that identifies separately the announced and implemented shocks.

Insert Table II here

[6]Difference in the table relative to Devries et al. 2013 is due to two facts. First the scaling is done using the GDP of the year prior to consolidation. Second, to be consistent with Alesina et al. only for the revenue part we use the CBO 1998 document Projecting Federal Tax Revenues and the Effects of Changes in the tax Law, p.31 (the difference is very small and does not influence main results of the paper)

[7]R&R propose several measures of the tax adjustments, generated respectively by including or not the retroactive components of the measures. There are no cases of retroactive components in deficit driven adjustments, and the retroactive components of a long run do not affect our measure of revenue adjustments.

1.6 Estimation and Simulation Procedure in Details

In this Appendix section we explain in details the procedure adopted to estimate via Maximum Likelihood the model

$$\Delta y_{i,t} = c_i + \left(\beta^{down} \cdot \Delta y^d_{i,t} + \delta^u \cdot e^u_t + \delta^a \cdot e^a_{t,t} + \delta^f \cdot e^f_t\right) \cdot TB_t + \\ + \left(\beta^{up} \cdot \Delta y^u_{i,t} + \gamma^u \cdot e^u_{i,t} + \gamma^a \cdot e^a_{i,t,t} + \gamma^f \cdot e^f_{i,t}\right) \cdot EB_T.$$

First of all, by subsuming the fixed effects and the fiscal shocks in a matrix called X and doing the same for their coefficients (we group them into a vector named β), we can use the following compact representation of our model:

$$\begin{aligned}
&(H_t)^{-1} \cdot \underset{n\times 1}{\Delta y_t} = \underset{n\times(n+6)}{X_t} \cdot \beta + \varepsilon_t \\
&(H_t)^{-1} = I_n - (\beta^{down} \cdot A_0 \cdot TB_t + \beta^{up} \cdot \hat{A}_0' \cdot EB_t) \\
&\varepsilon_t \sim \mathcal{N}(0, \Omega), \forall t \in \{1, ..., T\} \\
&\Omega = diag(\sigma_1^2, ..., \sigma_n^2) \\
&\varepsilon_t \perp \varepsilon_{t+i}, \quad \forall t \in \{1, ..., T\}, \forall i \in \mathcal{Z}
\end{aligned}$$

LeSage&Pace (2009) shows how to implement the calculation of the Maximum Likelihood Estimator for such a model. However, our model specification differs slightly from theirs (the standard SAR framework). In particular, we have a panel dataset and the network is activated in different years according with some dummy variables: TB_t and EB_t.

In order to derive the log-likelihood of our model at time t let's start off by setting $(H_t)^{-1} \cdot \Delta y_t = Z_t$, we have that:

$$Z_t = (H_t)^{-1} \cdot \Delta y_t \sim \mathcal{N}(X_t\beta, \Omega),$$

Therefore we have

$$\Delta y_t \sim \mathcal{N}(H_t X_t \beta, H_t \Omega H_t')$$

The density function of the random vector Δy_t is:

$$f(\underset{n\times 1}{\Delta y_t} | X_t, \rho, \beta, \Omega) = \frac{1}{\sqrt{(2\pi)^n \cdot |H_t \Omega H_t'|}} \exp\left\{-\frac{1}{2} \cdot (\Delta y_t - H_t X_t \beta)' \cdot (H_t \Omega H_t')^{-1} \cdot (\Delta y_t - H_t X_t \beta)\right\},$$

with $\rho = (\beta^{down} \quad \beta^{up})$.
Given that:

$$(H_t \Omega H_t')^{-1} = (H_t')^{-1} \cdot \Omega^{-1} \cdot H_t^{-1}$$

and

$$|H_t \Omega H_t'| = |H_t|^2 \cdot |\Omega|$$

we have:

$$\begin{aligned}
f(\Delta y_t|\cdot) &= (2\pi)^{-n/2} \cdot |H_t|^{-1} \cdot |\Omega|^{-1/2} \cdot \exp\left\{ -\frac{1}{2}(Z_t - X_t\beta)' \cdot H_t' \cdot (H_t')^{-1} \cdot \Omega^{-1} \cdot H_t^{-1} \cdot H_t \cdot (Z_t - X_t\beta)\right\} \\
&= (2\pi)^{-n/2} \cdot |(I_n - \beta^{down} A_0 TB_t - \beta^{up} \hat{A}_0' EB_t)^{-1}|^{-1} \cdot |\Omega|^{-1/2} \exp\left\{ -\frac{1}{2}\varepsilon_t' \Omega^{-1} \varepsilon_t \right\} \\
&= (2\pi)^{-n/2} \cdot |I_n - \rho_1 \cdot W_1 \cdot TB_t - \rho_2 \cdot W_2 \cdot EB_t| \cdot |\Omega|^{-1/2} \exp\left\{ -\frac{1}{2}\varepsilon_t' \Omega^{-1} \varepsilon_t \right\},
\end{aligned}$$

with $\rho_1 = \beta^{down}$, $\rho_2 = \beta^{up}$, $A_0 = W_1$ and $\hat{A}_0' = W_2$ (to ease notation).
At this point we need to find the likelihood of the random vector $\boldsymbol{\Delta y_t}$:

$$\boldsymbol{\Delta y_t} = \begin{bmatrix} \Delta y_1 & \dots & \Delta y_T \end{bmatrix}'.$$

Since our model is static and we have assumed

$$cov(\varepsilon_t, \varepsilon_{t-k}) = \underset{n \times n}{\mathbf{0}},$$

we consider our variables Δy_t, to be *iid*. By consequence, the following holds:

$$\begin{aligned}
f(\underset{nT \times 1}{\boldsymbol{\Delta y_t}}|X_1, \dots, X_T, \rho, \beta, \Omega) = \prod_{t=1}^{T} f(\underset{n \times 1}{\Delta y_t}|X_t, \rho, \beta, \Omega) = \big((2\pi)^n |\Omega|\big)^{-T/2} \cdot \\
\cdot \prod_{t=1}^{T} |I_n - \rho_1 \cdot W_1 \cdot TB_t - \rho_2 \cdot W_2 \cdot EB_t| \exp\big\{ -\frac{1}{2} \cdot \sum_{t=1}^{T} \varepsilon_t' \Omega^{-1} \varepsilon_t \big\}.
\end{aligned}$$

Now we divide the time series of length T in three different subperiods. In doing so, consider the following new parameters:

- t_1: set of years when a tax based fiscal adjustment occurs. Formally:

$$t_1 := \{1, ..., t, ..., T_1 | t \text{ such that } TB_t = 1\}$$

We set:

$$H_t | t \in t_1 = (I_n - \rho_1 \cdot W_1)^{-1} = H_\tau$$

- t_2: set of years when an expenditure tax based fiscal adjustment occurs. Formally:
$$t_2 := \{1, ..., t, ..., T_2 | t \text{ such that } EB_t = 1\}$$
We set:
$$H_t | t \in t_2 = (I_n - \rho_2 \cdot W_2)^{-1} = H_\gamma$$
- t_3: set of years when neither a tax based fiscal adjustment nor an expediture based fiscal adjustment occurs. Formally:
$$t_3 := \{1, ..., t, ..., T_3 | t \text{ such that } TB_t = 0 \land EB_t = 0\}$$
We set:
$$H_t | t \in t_3 = (I_n)^{-1} = I_n$$

Therefore, we have that t_1, t_2 and t_3 account for a partition of the whole time series and $T = T_1 + T_2 + T_3$. By consequence we have:

$$\begin{aligned}
\prod_{t=1}^{T} |I_n - \rho_1 W_1 TB_t - \rho_2 W_2 EB_t| &= \prod_{t=1}^{T} |H_t^{-1}| \\
&= \prod_{t=1}^{T} \frac{1}{|H_t|} \\
&= \prod_{t \in t_1}^{T_1} \frac{1}{|H_t|} \cdot \prod_{t \in t_2}^{T_2} \frac{1}{|H_t|} \cdot \prod_{t \in t_3}^{T_3} \frac{1}{|H_t|} \\
&= |H_\tau|^{-T_1} \cdot |H_\gamma|^{-T_2} \cdot |I_n|^{-T_3} \\
&= |I_n - \rho_1 \cdot W_1|^{T_1} \cdot |I_n - \rho_2 W_2|^{T_2}
\end{aligned}$$

At this point, we rewrite the probability density function of our dependent variable as:

$$\begin{aligned}
f(\boldsymbol{\Delta y_t} | X_1, \ldots, X_T, \rho, \beta, \Omega) &= (2\pi)^{-nT/2} \cdot |\Omega|^{-T/2} \cdot \\
&\cdot |I_n - \rho_1 \cdot W_1|^{T_1} \cdot |I_n - \rho_2 W_2|^{T_2} \cdot \exp\left\{ -\frac{1}{2} \cdot \sum_{t=1}^{T} \varepsilon_t' \cdot \Omega^{-1} \cdot \varepsilon_t \right\}.
\end{aligned}$$

Eventually, we express the log-likelihood of our dataset:

$$\begin{aligned}
\log \mathcal{L}(\rho, \beta, \Omega | \Delta y_1, \ldots, \Delta y_T, X_1, \ldots, X_T) &= -\frac{nT}{2} \ln(2\pi) - \frac{T}{2} \cdot \ln(|\Omega|) + \\
&+ T_1 \cdot \ln(|I_n - \rho_1 \cdot W_1|) + T_2 \cdot \ln(|I_n - \rho_2 W_2|) - \frac{1}{2} \cdot \sum_{t=1}^{T} \varepsilon_t' \cdot \Omega^{-1} \cdot \varepsilon_t.
\end{aligned}$$

with:

$$\varepsilon_t = Z_t - X_t \cdot \beta = H_t^{-1} \cdot \Delta y_t - X_t\beta = (I_n - \rho_1 W_1 TB_t - \rho_2 W_2 EB_t) \cdot \Delta y_t - X_t \cdot \beta.$$

Furthermore, we impose the condition $\lambda_{\min}^{-1} < \hat{\rho}_1 < \lambda_{\max}^{-1}$ and $\mu_{\min}^{-1} < \hat{\rho}_2 < \mu_{\max}^{-1}$, where λ and μ are the eigenvalues of the spatial matrices W_1 and W_2 respectively.[8] Such a condition guarantees that the Variance-Covariance Matrix of the ML estimator is positive definite.

At this point we concentrate the log-likelihood by computing the partial derivatives of it. Let's start with deriving the concentrated estimator of β. In our model β contains the $n = 15$ fixed effects plus the 6 coefficients in front of the fiscal shocks: unexpected, announced and future for both taxes and expenditures.

$$\frac{\partial \log \mathcal{L}(\rho, \beta, \Omega|\cdot)}{\partial \beta} = -\frac{1}{2} \cdot \frac{\partial (\sum_{t=1}^{T} \varepsilon_t' \cdot \Omega^{-1} \varepsilon_t)}{\partial \beta}.$$

Note that:

$$\sum_{t=1}^{T} \varepsilon_t' \cdot \Omega^{-1} \varepsilon_t = \begin{bmatrix} \varepsilon_1' & \cdots & \varepsilon_T' \end{bmatrix} \cdot \Sigma^{-1} \cdot \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_T \end{bmatrix} = \varepsilon' \cdot \Sigma^{-1} \cdot \varepsilon,$$

where:

$$\underset{nt \times nT}{\Sigma} = \begin{bmatrix} \Omega & \underset{n\times n}{\mathbf{0}} & \cdots & \underset{n\times n}{\mathbf{0}} \\ \underset{n\times n}{\mathbf{0}} & \Omega & \cdots & \underset{n\times n}{\mathbf{0}} \\ \vdots & \vdots & \ddots & \vdots \\ \underset{n\times n}{\mathbf{0}} & \underset{n\times n}{\mathbf{0}} & \cdots & \Omega \end{bmatrix}$$

Also:

$$\Sigma^{-1} = \begin{bmatrix} \Omega & \underset{n\times n}{\mathbf{0}} & \cdots & \underset{n\times n}{\mathbf{0}} \\ \underset{n\times n}{\mathbf{0}} & \Omega & \cdots & \underset{n\times n}{\mathbf{0}} \\ \vdots & \vdots & \ddots & \vdots \\ \underset{n\times n}{\mathbf{0}} & \underset{n\times n}{\mathbf{0}} & \cdots & \Omega \end{bmatrix}^{-1} = \begin{bmatrix} \Omega^{-1} & \underset{n\times n}{\mathbf{0}} & \cdots & \underset{n\times n}{\mathbf{0}} \\ \underset{n\times n}{\mathbf{0}} & \Omega^{-1} & \cdots & \underset{n\times n}{\mathbf{0}} \\ \vdots & \vdots & \ddots & \vdots \\ \underset{n\times n}{\mathbf{0}} & \underset{n\times n}{\mathbf{0}} & \cdots & \Omega^{-1} \end{bmatrix}.$$

Moreover:

$$\begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_T \end{bmatrix} = \varepsilon = Z - X \cdot \beta = \begin{bmatrix} Z_1 \\ \vdots \\ Z_T \end{bmatrix} - \begin{bmatrix} X_1 \\ \vdots \\ X_T \end{bmatrix} \cdot \beta,$$

[8]See Ord (1975)

therefore,:

$$\sum_{t=1}^{T} \varepsilon_t' \cdot \Omega^{-1} \varepsilon_t = (Z - X \cdot \beta)' \cdot \Sigma^{-1} \cdot (Z - X \cdot \beta) =$$
$$= Z' \cdot \Sigma^{-1} \cdot Z - 2 \cdot Z' \cdot \Sigma^{-1} \cdot X \cdot \beta + \beta' \cdot X \cdot \Sigma^{-1} \cdot X \cdot \beta$$

At this point it can be verified that:

$$\text{(FOC)} \quad \frac{\partial \log \mathcal{L}(\rho, \beta, \Omega|\cdot)}{\partial \beta} = X' \cdot \Sigma^{-1} \cdot Z - X' \cdot \Sigma^{-1} \cdot X \cdot \beta = 0$$

$$\beta = (X'\Sigma^{-1}X)^{-1} X' \Sigma^{-1} Z.$$

The above estimator is the GLS estimator. The result is not surprising, since we have simply solved a standard squared deviation minimization problem.
Furthermore, we need to estimate the variance of the model to fully concentrate the likelihood in order to simply solve a two variable maximization problem.

$$\frac{\partial \log \mathcal{L}(\rho, \beta, \Omega|\cdot)}{\partial \Omega} = -\frac{T}{2} \cdot \frac{\partial(\ln(|\Omega|))}{\partial \Omega} - \frac{1}{2} \cdot \sum_{t=1}^{T} \frac{\partial(\varepsilon_t' \cdot \Omega^{-1} \cdot \varepsilon_t)}{\partial \Omega} =$$
$$= -\frac{T}{2} \cdot (\Omega')^{-1} - \frac{1}{2} \cdot \sum_{t=1}^{T} (-\Omega^{-1} \cdot \varepsilon_t \cdot \varepsilon_t' \cdot \Omega^{-1})$$
$$= \frac{1}{2} \cdot \Omega^{-1} \cdot \big[(\sum_{t=1}^{T} \varepsilon_t \cdot \varepsilon_t') \cdot \Omega^{-1} - T \big] = 0. \quad \text{(FOC)}$$

From the FOC it follows that:

$$\Omega = \frac{\sum_{t=1}^{T} \varepsilon_t \cdot \varepsilon_t'}{T} = \frac{1}{T} \cdot \begin{bmatrix} \sum_{t=1}^{T} \varepsilon_{1,t}^2 & \sum_{t=1}^{T} \varepsilon_{1,t} \cdot \varepsilon_{2,t} & \cdots & \sum_{t=1}^{T} \varepsilon_{1,t} \cdot \varepsilon_{n,t} \\ & \sum_{t=1}^{T} \varepsilon_{2,t}^2 & \cdots & \sum_{t=1}^{T} \varepsilon_{2,t} \cdot \varepsilon_{n,t} \\ & & \ddots & \vdots \\ & & & \sum_{t=1}^{T} \varepsilon_{n,t}^2 \end{bmatrix}$$

Since we assume Ω to be diagonal, we are only interested in the variances of the sectors:

$$\Omega = diag(\sigma_1^2, \ldots, \sigma_n^2)$$

$$\sigma_i^2 = \frac{1}{T} \cdot \sum_{t=1}^{T} \varepsilon_{i,t}^2$$

In order to pass to Feasible GLS estimator, we need to use the OLS residuals:

$$\hat{\varepsilon}_{i,t} = Z_{i,t} - X_{i,t} \cdot \underbrace{(X'X)^{-1}X'Z}_{\beta_{OLS}}.$$

Since every equation has 7 parameters (the industry fixed effect plus the 6 fiscal shocks coefficients) we have:

$$\hat{\sigma}_i^2 = \frac{1}{T-7} \cdot \sum_{t=1}^{T} \hat{\varepsilon}_{i,t}^2,$$

and finally we have:

$$\hat{\Omega} = diag(\hat{\sigma}_1^2, \ldots, \hat{\sigma}_n^2)$$
$$\hat{\beta} = (X'\hat{\Sigma}^{-1}X)^{-1}X'\hat{\Sigma}^{-1}Z$$

At this point it is simply a matter of solving the following problem:

$$\max_{\rho_1, \rho_2} \quad \log \mathcal{L}(\rho_1, \rho_2, \hat{\Omega}, \hat{\beta} \mid \cdot \ ) \quad \text{s.t. } \rho_1 \in (\lambda_{\max}^{-1}, \lambda_{\max}^{-1}) \text{ and } \rho_2 \in (\mu_{\min}^{-1}, \mu_{\max}^{-1})$$

Once estimated the coefficients of model

$$\Delta y_{i,t} = c_i + \left(\beta^{down} \cdot \Delta y_{i,t}^d + \delta^u \cdot e_t^u + \delta^a \cdot e_{t,t}^a + \delta^f \cdot e_t^f\right) \cdot TB_t +$$
$$+ \left(\beta^{up} \cdot \Delta y_{i,t}^u + \gamma^u \cdot e_{i,t}^u + \gamma^a \cdot e_{i,t,t}^a + \gamma^f \cdot e_{i,t}^f\right) \cdot EB_T.$$

we proceeded with computing the standard errors of the estimates. In order to do that, we computed analitically the elements of the Fisher Information Matrix ($\mathcal{I}$). In fact recall that:

$$\sqrt{n} \cdot (\hat{\theta}_0 - \theta) \xrightarrow{d} \mathcal{N}(0, \mathcal{I}^{-1})$$

In order to derive the Fisher Information Matrix we firstly need to obtain the total gradient of the log-likelihood function. Let's start with the spatial coefficient ρ_1:

$$\frac{\partial \log \mathcal{L}(\theta|\Delta y, X)}{\partial \rho_1} = T_1 \frac{1}{|I_n - \rho_1 W_1|} \frac{\partial |I_n - \rho_1 W_1|}{\partial \rho_1} - \frac{1}{2} \sum_{t=1}^{T} \frac{\partial (Z_t' \Omega^{-1} Z_t)}{\partial \rho_1} - 2 \frac{\partial (Z_t' \Omega^{-1} X_t \beta)}{\partial \rho_1}.$$

By some matrix algebra, it is possible to show that:

$$\frac{\partial(Z_t'\Omega^{-1}Z_t)}{\partial\rho_1} = -TB_t \cdot \Delta y_t' \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t - TB_t \cdot \Delta y_t' \cdot W_1'\Omega^{-1} \cdot \Delta y_t$$
$$+ 2\rho_1 \cdot TB_t^2 \cdot \Delta y_t' \cdot W_1 \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t' + 2\rho_2 \cdot TB_t \cdot EB_t \cdot \Delta y_t' \cdot W_1 \cdot \Omega^{-1} \cdot W_2 \cdot \Delta y_t'$$

Since our fiscal adjustment plans are mutually exclusive, we have that $TB_t \cdot EB_t = 0$ for all t. Moreover, by rearranging the above expression, we get:

$$\frac{\partial(Z_t'\Omega^{-1}Z_t)}{\partial\rho_1} = -2 \cdot TB_t \cdot \Delta y_t' \cdot (I_n - \rho_1 \cdot W_1') \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t$$

After other matrix algebra, we get:

$$-2 \cdot \frac{\partial(Z_t \cdot \Omega^{-1}X_t\beta)}{\partial\rho_1} = 2 \cdot TB_t \cdot \Delta y_t' \cdot W_1' \cdot \Omega^{-1} \cdot X_t \cdot \beta$$

Wrapping up all together, and employing the notation introduced earlier: $(I_n - \rho_1 W_1)^{-1} = H_\tau$, we have:

$$\begin{aligned}
\frac{\partial \log \mathcal{L}(\theta|\Delta y, X)}{\partial\rho_1} &= T_1 \frac{1}{|I_n - \rho_1 W_1|} \frac{\partial|I_n - \rho_1 W_1|}{\partial\rho_1} + \\
&+ \sum_{t\in t_1}^{T_1} \Big[\Delta y_t' \cdot (I_n - \rho_1 \cdot W_1') \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t - \Delta y_t' \cdot W_1' \cdot \Omega^{-1} \cdot X_t \cdot \beta\Big] = \\
&= T_1 \frac{1}{|I_n - \rho_1 W_1|} \cdot |I_n - \rho_1 W_1| \cdot Tr\Big((I_n - \rho_1 W_1)^{-1} \cdot (-W_1)\Big) + \\
&+ \sum_{t\in t_1}^{T_1} \Big[\big((I_n - \rho_1 \cdot W_1) \cdot \Delta y_t\big)' \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t - \beta' \cdot X_t' \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t\Big] \\
&= -T_1 \cdot Tr\Big(H_\tau \cdot W_1\Big) + \sum_{t\in t_1}^{T_1} \Big[\big(Z_t - X_t\beta\big)' \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t\Big] \\
&= \sum_{t\in t_1}^{T_1} \big(\varepsilon_t' \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t\big) - T_1 \cdot Tr\big(H_\tau \cdot W_1\big).
\end{aligned}$$

By simmetry we have that:

$$\frac{\partial \log \mathcal{L}(\theta|\Delta y, X)}{\partial\rho_2} = \sum_{t\in t_2}^{T_2} \big(\varepsilon_t' \cdot \Omega^{-1} \cdot W_2 \cdot \Delta y_t\big) - T_2 \cdot Tr\big(H_\gamma \cdot W_2\big),$$

with $H_\gamma = (I_n - \rho_2 W_2)^{-1}$, from the previous notation.
As far as concern the derivative with respect to β, we have already seen when concentrating the log-likelihood that:

$$\begin{aligned}
\frac{\partial \log \mathcal{L}(\theta|\Delta y, X)}{\partial \beta} &= X' \cdot \Sigma^{-1} \cdot Z - X' \cdot \Sigma^{-1} \cdot X \cdot \beta \\
&= X' \cdot \Sigma^{-1} \cdot (Z - X \cdot \beta) = \\
&= X' \cdot \Sigma^{-1} \cdot \varepsilon = \\
&= \sum_{t=1}^{T} X_t' \cdot \Omega^{-1} \cdot \varepsilon_t.
\end{aligned}$$

Concerning the derivatives with respect to σ_i^2, we need firstly to acknwoledge that:

$$\sum_{t=1}^{T} \varepsilon_t' \cdot \Omega^{-1} \cdot \varepsilon_t = \sum_{t=1}^{T} \sum_{i=1}^{n} \frac{\varepsilon_{i,t}^2}{\sigma_i^2} = \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \sum_{t=1}^{T} \varepsilon_{i,t}^2,$$

and that:

$$\ln(|\Omega|) = \ln(\prod_{i=1}^{n} \sigma_i^2) = \sum_{i=1}^{n} \ln(\sigma_i^2).$$

Therefore, we have that:

$$\begin{aligned}
\frac{\partial \log \mathcal{L}(\theta|\Delta y, X)}{\partial \sigma_i^2} &= -\frac{T}{2} \frac{\partial \ln(|\Omega|)}{\partial \sigma_i^2} - \frac{1}{2} \cdot \frac{\partial}{\partial \sigma_i^2} \sum_{t=1}^{T} \varepsilon_t' \cdot \Omega^{-1} \cdot \varepsilon_t \\
&= -\frac{T}{2 \cdot \sigma_i^2} + \frac{1}{2 \cdot \sigma_i^4} \cdot \sum_{t=1}^{T} \varepsilon_{i,t}^2.
\end{aligned}$$

We now have all the elements to write down the gradient of the log-likelihood:

$$\nabla \log \mathcal{L}(\theta|\Delta y, X) = \underset{38\times 1}{\begin{bmatrix} \dfrac{\partial \log \mathcal{L}(\theta|\Delta y, X)}{\partial \rho_1} \\[2ex] \dfrac{\partial \log \mathcal{L}(\theta|\Delta y, X)}{\partial \rho_2} \\[2ex] \dfrac{\partial \log \mathcal{L}(\theta|\Delta y, X)}{\partial \beta} \\[2ex] \dfrac{\partial \log \mathcal{L}(\theta|\Delta y, X)}{\partial \sigma_1^2} \\[2ex] \vdots \\[2ex] \dfrac{\partial \log \mathcal{L}(\theta|\Delta y, X)}{\partial \sigma_n^2} \end{bmatrix}} = \begin{bmatrix} \sum_{t\in t_1}^{T_1} \big(\varepsilon_t' \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t\big) - T_1 \cdot Tr\big(H_\tau \cdot W_1\big) \\[2ex] \sum_{t\in t_2}^{T_2} \big(\varepsilon_t' \cdot \Omega^{-1} \cdot W_2 \cdot \Delta y_t\big) - T_2 \cdot Tr\big(H_\gamma \cdot W_2\big) \\[2ex] \sum_{t=1}^{T} X_t' \cdot \Omega^{-1} \cdot \varepsilon_t \\[2ex] -\dfrac{T}{2\cdot\sigma_1^2} + \dfrac{1}{2\cdot\sigma_1^4} \cdot \sum_{t=1}^{T} \varepsilon_{1,t}^2 \\[2ex] \vdots \\[2ex] -\dfrac{T}{2\cdot\sigma_n^2} + \dfrac{1}{2\cdot\sigma_n^4} \cdot \sum_{t=1}^{T} \varepsilon_{n,t}^2 \end{bmatrix}$$

The gradient contains overall 38 elements, that is, we need to estimate 38 parameters. By consequence, the Fisher Information Matrix will be a 38×38 array.
Let's start with the first row of the matrix: all the derivatives of $\dfrac{\partial \log \mathcal{L}(\theta|\Delta y, X)}{\partial \rho_1}$ with resepct to all the parameters. To simplify notation I will refer with $\mathcal{H}_{ij}$ to the element of row i and column j of the Hessian matrix.

$$\begin{aligned}
\mathcal{H}_{1,1} = \frac{\partial^2 \log \mathcal{L}(\theta|\Delta y, X)}{\partial \rho_1^2} &= \sum_{t\in t_1}^{T_1} \big(\frac{\partial \varepsilon_t'}{\partial \rho_1} \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t\big) - T_1 \cdot \frac{\partial Tr\big(H_\tau \cdot W_1\big)}{\partial \rho_1} \\
&= \sum_{t\in t_1}^{T_1} \big((-\Delta y_t' \cdot W_1') \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t\big) - T_1 \cdot Tr\big(\frac{\partial H_\tau}{\partial \rho_1} \cdot W_1\big) = \\
&= -\sum_{t\in t_1}^{T_1} \big(\Delta y_t' \cdot W_1' \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t\big) - T_1 \cdot Tr\big((-H_\tau \cdot (-W_1) \cdot H_\tau) \cdot W_1\big) = \\
&= -T_1 \cdot Tr\big(W_1 \cdot H_\tau \cdot W_1 \cdot H_\tau\big) - \sum_{t\in t_1}^{T_1} \big(\Delta y_t' \cdot W_1' \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t\big)
\end{aligned}$$

Simmetrically we have:

$$\mathcal{H}_{2,2} = \frac{\partial^2 \log \mathcal{L}(\theta|\Delta y, X)}{\partial \rho_2^2} =$$
$$= -T_2 \cdot Tr\big(W_2 \cdot H_\gamma \cdot W_2 \cdot H_\gamma\big) - \sum_{t\in t_2}^{T_2} \big(\Delta y_t' \cdot W_2' \cdot \Omega^{-1} \cdot W_2 \cdot \Delta y_t\big)$$

Going back to the first row, we now calculate the cross derivative with respect to $\rho 2$. Before doing so, recall that, being the log-likelihood a continuously differentiable function, the Schwarz's theorem applies and the Hessian matrix is symmetric.

$$\mathcal{H}_{1,2} = \mathcal{H}_{2,1} = \frac{\partial^2 \log \mathcal{L}(\theta|\Delta y, X)}{\partial \rho_1 \partial \rho_2} = 0.$$

Going on with the calculation we have:

$$\begin{aligned}\mathcal{H}_{1,3:1,23} = \frac{\partial^2 \log \mathcal{L}(\theta|\Delta y, X)}{\partial \rho_1 \partial \beta} &= \sum_{t\in t_1}^{T_1} \big(\frac{\partial \varepsilon_t'}{\partial \beta} \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t\big) \\ &= -\sum_{t\in t_1}^{T_1} X_t' \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t \\ &= -X_\tau' \cdot \underbrace{(I_{T_1} \otimes \Omega^{-1})}_{\Sigma_\tau^{-1}} \cdot (I_{T_1} \otimes W_1) \cdot \Delta y_\tau\end{aligned}$$

where $\mathcal{H}_{1,3:1,23}$ means all the elements of the first row, from column 3 up to column 23. X_τ and Δy_τ represent X and Δy but for the only years when a tax based fiscal adjustment occur:

$$X_\tau = \underset{T_1 n\times k}{\begin{bmatrix} X_1 \\ \vdots \\ X_t \\ \vdots \\ X_{T_1}\end{bmatrix}} \quad \text{and} \quad \Delta y_\tau = \underset{T_1 n\times k}{\begin{bmatrix} \Delta y_1 \\ \vdots \\ \Delta y_t \\ \vdots \\ \Delta y_{T_1}\end{bmatrix}} \quad \text{with } t \in t_1,$$

Symmetrically:

$$\begin{aligned}
\mathcal{H}_{2,3:2,23} = \frac{\partial^2 \log \mathcal{L}(\theta|\Delta y, X)}{\partial \rho_2 \partial \beta} &= \sum_{t \in t_2}^{T_2} \big(\frac{\partial \varepsilon_t'}{\partial \beta} \cdot \Omega^{-1} \cdot W_2 \cdot \Delta y_t\big) \\
&= -\sum_{t \in t_2}^{T_2} X_t' \cdot \Omega^{-1} \cdot W_2 \cdot \Delta y_t \\
&= -X_\gamma' \cdot \underbrace{(I_{T_2} \otimes \Omega^{-1})}_{\Sigma_\gamma^{-1}} \cdot (I_{T_2} \otimes W_2) \cdot \Delta y_\gamma,
\end{aligned}$$

with:

$$X_\gamma = \underset{T_2 n \times k}{\begin{bmatrix} X_1 \\ \vdots \\ X_t \\ \vdots \\ X_{T_2} \end{bmatrix}} \quad \text{and} \quad \Delta y_\gamma = \underset{T_2 n \times k}{\begin{bmatrix} \Delta y_1 \\ \vdots \\ \Delta y_t \\ \vdots \\ \Delta y_{T_2} \end{bmatrix}} \quad \text{with } t \in t_2,$$

$$\begin{aligned}
\mathcal{H}_{3,3:23,23} = \frac{\partial^2 \log \mathcal{L}(\theta|\Delta y, X)}{\partial \beta^2} &= \frac{\partial}{\partial \beta^2}\Bigg(\sum_{t=1}^{T} X_t' \cdot \Omega^{-1} \cdot \varepsilon_t\Bigg) \\
&= \sum_{t=1}^{T} X_t' \cdot \Omega^{-1} \cdot \frac{\partial (Z_t - X_t \cdot \beta)}{\partial \beta^2} \\
&= \sum_{t=1}^{T} X_t' \cdot \Omega^{-1} \cdot X_t \\
&= -X' \cdot \Sigma^{-1} \cdot X.
\end{aligned}$$

$$\mathcal{H}_{3,24:23,38} = \frac{\partial^2 \log \mathcal{L}(\theta|\Delta y, X)}{\partial \beta \partial \sigma^2} = \sum_{t=1}^{T} X_t' \cdot \frac{\partial \Omega^{-1}}{\partial \sigma^2} \cdot \varepsilon_t$$

The generic element of the above matrix is a $k \times 1$ vector:

$$-\sigma_1^{-4} \cdot \sum_{t=1}^{T} X_{1,t}' \cdot \varepsilon_{i,t}.$$

Going on with the calculation:

$$\mathcal{H}_{i,i|i\in[24,38]} = \frac{\partial^2 \log \mathcal{L}(\theta|\Delta y, X)}{\partial(\sigma_i^2)^2} = \frac{T}{2} \cdot \frac{1}{\sigma_i^4} \cdot \left(1 - \frac{2}{T \cdot \sigma_i^2} \cdot \sum_{t=1}^{T} \varepsilon_{i,t}^2\right).$$

$$\mathcal{H}_{23+i,23+j|i,j\in[1,n]} = \frac{\partial^2 \log \mathcal{L}(\theta|\Delta y, X)}{\partial\sigma_i^2 \partial\sigma_j^2} = 0 \quad \forall i \neq j.$$

$$\begin{aligned}
\mathcal{H}_{1,24:1,38} = \frac{\partial^2 \log \mathcal{L}(\theta|\Delta y, X)}{\partial\rho_1 \partial\sigma_i^2} &= \frac{\partial}{\partial\sigma_i^2}\Big(\sum_{t\in t_1}^{T_1} \varepsilon_t' \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t\Big) \\
&= \frac{\partial}{\partial\sigma_i^2}\Big(\sum_{t\in t_1}^{T_1} Tr\big(\varepsilon_t' \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t\big)\Big) \\
&= \frac{\partial}{\partial\sigma_i^2}\Big(Tr\big((\sum_{t\in t_1}^{T_1} \Delta y_t \cdot \varepsilon_t') \cdot \Omega^{-1} \cdot W_1\big)\Big) \\
&= Tr\Big((\sum_{t\in t_1}^{T_1} \Delta y_t \cdot \varepsilon_t') \cdot \frac{\partial\Omega^{-1}}{\partial\sigma_i^2} \cdot W_1\Big)
\end{aligned}$$

Note that

$$\frac{\partial\Omega^{-1}}{\partial\sigma_i^2} = \begin{bmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & -\sigma_i^{-4} & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{bmatrix} = diag(0, \cdots, 0, -\sigma_i^{-4}, 0, \cdots, 0)$$

Simmetrically:

$$\mathcal{H}_{2,24:2,38} = \frac{\partial^2 \log \mathcal{L}(\theta|\Delta y, X)}{\partial\rho_2 \partial\sigma_i^2} = Tr\Big((\sum_{t\in t_2}^{T_2} \Delta y_t \cdot \varepsilon_t') \cdot \frac{\partial\Omega^{-1}}{\partial\sigma_i^2} \cdot W_2\Big)$$

At this point we have all the elements to construct the Hessian matrix of the log-likelihood.
To sum up, first row:

- $\mathcal{H}_{1,1} = -T_1 \cdot Tr\big(W_1 \cdot H_\tau \cdot W_1 \cdot H_\tau\big) - \sum_{t\in t_1}^{T_1} \big(\Delta y_t' \cdot W_1' \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t\big)$

- $\mathcal{H}_{1,2} = 0$

- $\mathcal{H}_{1,3:1,23} = -\sum_{t \in t_1}^{T_1} X_t' \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t$

- $\mathcal{H}_{1,24:1,38} = Tr\Big(\big(\sum_{t \in t_1}^{T_1} \Delta y_t \cdot \varepsilon_t'\big) \cdot \dfrac{\partial \Omega^{-1}}{\partial \sigma_i^2} \cdot W_1\Big).$

Second row:

- $\mathcal{H}_{2,1} = 0$

- $\mathcal{H}_{2,2} = -T_2 \cdot Tr\big(W_2 \cdot H_\gamma \cdot W_2 \cdot H_\gamma\big) - \sum_{t \in t_2}^{T_2} \big(\Delta y_t' \cdot W_2' \cdot \Omega^{-1} \cdot W_2 \cdot \Delta y_t\big)$

- $\mathcal{H}_{2,3:2,23} = -\sum_{t \in t_2}^{T_2} X_t' \cdot \Omega^{-1} \cdot W_2 \cdot \Delta y_t$

- $\mathcal{H}_{2,24:2,38} = Tr\Big(\big(\sum_{t \in t_2}^{T_2} \Delta y_t \cdot \varepsilon_t'\big) \cdot \dfrac{\partial \Omega^{-1}}{\partial \sigma_i^2} \cdot W_2\Big).$

From row 3 to row 23:

- $\mathcal{H}_{3,1:23,1} = \mathcal{H}_{1,3:1,23}'$

- $\mathcal{H}_{3,2:23,2} = \mathcal{H}_{2,3:2,23}'$

- $\mathcal{H}_{3,3:23,23} = \sum_{t=1}^{T} X_t' \cdot \Omega^{-1} \cdot X_t$

- $\mathcal{H}_{3,24:23,38} = \sum_{t=1}^{T} X_t' \cdot \dfrac{\partial \Omega^{-1}}{\partial \sigma^2} \cdot \varepsilon_t$

From row 24 to the last row (number 38):

- $\mathcal{H}_{24,1:38,1} = \mathcal{H}_{1,24:1,38}'$

- $\mathcal{H}_{24,2:38,2} = \mathcal{H}_{2,24:2,38}'$

- $\mathcal{H}_{24,3:38,23} = \mathcal{H}'_{3,24:23,38}$

- $$\mathcal{H}_{23+i,23+j|i,j\in[1,n]} = \begin{cases} \dfrac{T}{2} \cdot \dfrac{1}{\sigma_i^4} \cdot \left(1 - \dfrac{2}{T \cdot \sigma_i^2} \cdot \displaystyle\sum_{t=1}^{T} \varepsilon_{i,t}^2\right) & \forall i = j \in [1,n] \\ \\ 0 \quad \forall i \neq j \end{cases}$$

The last step we have to make to finally obtain the Fisher Information Matrix is taking expectations of every element.

$$
\begin{aligned}
E[\mathcal{H}_{1,1}] &= -T_1 \cdot Tr\big(W_1 \cdot H_\tau \cdot W_1 \cdot H_\tau\big) - \sum_{t \in t_1}^{T_1} E\big[\Delta y_t' \cdot W_1' \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t\big] = \\
&= -T_1 \cdot Tr\big(W_1 \cdot H_\tau \cdot W_1 \cdot H_\tau\big) - \sum_{t \in t_1}^{T_1} E\big[Tr\big(W_1 \cdot \Delta y_t \cdot \Delta y_t' \cdot W_1' \cdot \Omega^{-1}\big)\big] = \\
&= -T_1 \cdot Tr\big(W_1 \cdot H_\tau \cdot W_1 \cdot H_\tau\big) - \sum_{t \in t_1}^{T_1} Tr\big(W_1 \cdot E\big[\Delta y_t \cdot \Delta y_t'\big] \cdot W_1' \cdot \Omega^{-1}\big) = \\
&= -T_1 \cdot Tr\big(W_1 \cdot H_\tau \cdot W_1 \cdot H_\tau\big) - \sum_{t \in t_1}^{T_1} Tr\big(W_1 \cdot E\big[H_\tau \cdot X_t \cdot \beta \cdot \varepsilon_t' \cdot H_\tau' + \\
&+ H_\tau \cdot X_t \cdot \beta \cdot \beta' \cdot X_t' \cdot H_\tau' + H_\tau \cdot \varepsilon_t \cdot \varepsilon_t' \cdot H_\tau' \cdot \varepsilon_t \cdot \beta' \cdot X_t' \cdot H_\tau'\big] \cdot W_1' \cdot \Omega^{-1}\big) = \\
&= -T_1 \cdot Tr\big(W_1 \cdot H_\tau \cdot W_1 \cdot H_\tau\big) - \sum_{t \in t_1}^{T_1} Tr\Big(W_1 \cdot \Big[H_\tau \cdot X_t \cdot \beta \cdot E[\varepsilon_t'] \cdot H_\tau' + \\
&+ H_\tau \cdot X_t \cdot \beta \cdot \beta' \cdot X_t' \cdot H_\tau' + H_\tau \cdot E[\varepsilon_t \cdot \varepsilon_t'] \cdot H_\tau' + E[\varepsilon_t] \cdot \beta' \cdot X_t' \cdot H_\tau'\Big] \cdot W_1' \cdot \Omega^{-1}\Big) = \\
&= -T_1 \cdot Tr\big(W_1 \cdot H_\tau \cdot W_1 \cdot H_\tau\big) - \\
&- \sum_{t \in t_1}^{T_1} Tr\Big(W_1 \cdot \Big[H_\tau \cdot X_t \cdot \beta \cdot \beta' \cdot X_t' \cdot H_\tau' + H_\tau \cdot \Omega \cdot H_\tau'\Big] \cdot W_1' \cdot \Omega^{-1}\Big) = \\
&= -T_1 \cdot Tr\big(W_1 \cdot H_\tau \cdot W_1 \cdot H_\tau\big) - \\
&- \sum_{t \in t_1}^{T_1} Tr\big(W_1 \cdot H_\tau \cdot X_t \cdot \beta \cdot \beta' \cdot X_t' \cdot H_\tau' \cdot W_1' \cdot \Omega^{-1} + W_1 \cdot H_\tau \cdot \Omega \cdot H_\tau' \cdot W_1' \cdot \Omega^{-1}\big) = \\
&= -T_1 \cdot Tr\big(W_1 \cdot H_\tau \cdot W_1 \cdot H_\tau + H_\tau' \cdot W_1' \cdot \Omega^{-1} \cdot W_1 \cdot H_\tau \cdot \Omega\big) - \\
&- \sum_{t \in t_1}^{T_1} Tr\big(\beta' \cdot X_t' \cdot H_\tau' \cdot W_1' \cdot \Omega^{-1} \cdot W_1 \cdot H_\tau \cdot X_t \cdot \beta\big) =
\end{aligned}
$$

Setting $M_1^\tau = H'_\tau \cdot W'_1 \cdot \Omega^{-1} \cdot W_1 \cdot H_\tau$ we can rewrite the above identity as:

$$
\begin{aligned}
E[\mathcal{H}_{1,1}] &= -T_1 \cdot Tr\big(W_1 \cdot H_\tau \cdot W_1 \cdot H_\tau + M_1^\tau \cdot \Omega\big) - \sum_{t \in t_1}^{T_1} \beta' \cdot X'_t \cdot M_1^\tau \cdot X_t \cdot \beta = \\
&= -T_1 \cdot Tr\big(W_1 \cdot H_\tau \cdot W_1 \cdot H_\tau + M_1^\tau \cdot \Omega\big) - \beta' \cdot X'_\tau \cdot \big(I_{T_1} \otimes M_1^\tau\big) \cdot X_\tau \cdot \beta.
\end{aligned}
$$

Simmetrically:

$$
E[\mathcal{H}_{2,2}] = -T_2 \cdot Tr\big(W_2 \cdot H_\gamma \cdot W_2 \cdot H_\gamma + M_1^\gamma \cdot \Omega\big) - \beta' \cdot X'_\gamma \cdot \big(I_{T_2} \otimes M_1^\gamma\big) \cdot X_\gamma \cdot \beta.
$$

with $M_1^\gamma = H'_\gamma \cdot W'_2 \cdot \Omega^{-1} \cdot W_2 \cdot H_\gamma$.

Going on with the calculation:

$$
\begin{aligned}
E[\mathcal{H}_{1,3:1,23}] &= E\Big[-\sum_{t \in t_1}^{T_1} X'_t \cdot \Omega^{-1} \cdot W_1 \cdot \Delta y_t\Big] = \\
&= -\sum_{t \in t_1}^{T_1} X'_t \cdot \Omega^{-1} \cdot W_1 \cdot E\Big[H_\tau \cdot X_t \cdot \beta + H_\tau \cdot \varepsilon_t\Big] = \\
&= -\sum_{t \in t_1}^{T_1} X'_t \cdot \Omega^{-1} \cdot W_1 \cdot H_\tau \cdot X_t \cdot \beta \\
&= X'_\tau \cdot (I_{T_1} \otimes M_2^\tau) \cdot X_\tau \cdot \beta
\end{aligned}
$$

with $M_2^\tau = \Omega^{-1} \cdot W_1 \cdot H_\tau$.

Simmetrically:

$$
E[\mathcal{H}_{2,3:2,23}] = X'_\gamma \cdot (I_{T_2} \otimes M_2^\gamma) \cdot X_\gamma \cdot \beta
$$

with $M_2^\gamma = \Omega^{-1} \cdot W_2 \cdot H_\gamma$.

Next step:

$$
\begin{aligned}
E[\mathcal{H}_{1,24:1,38}] &= Tr\Big(\big(\sum_{t\in t_1}^{T_1} E\big[\Delta y_t\cdot\varepsilon_t'\big]\big)\cdot\frac{\partial\Omega^{-1}}{\partial\sigma_i^2}\cdot W_1\Big)=\\
&= Tr\Big(\big(\sum_{t\in t_1}^{T_1} E\big[\Delta y_t\cdot\varepsilon_t'\big]\big)\cdot\frac{\partial\Omega^{-1}}{\partial\sigma_i^2}\cdot W_1\Big)=\\
&= Tr\Big(\big(\sum_{t\in t_1}^{T_1} H_\tau\cdot E\big[\varepsilon_t\cdot\varepsilon_t'\big]\big)\cdot\frac{\partial\Omega^{-1}}{\partial\sigma_i^2}\cdot W_1\Big)=\\
&= T_1\cdot Tr\Big(H_\tau\cdot\Omega\cdot\frac{\partial\Omega^{-1}}{\partial\sigma_i^2}\cdot W_1\Big)=\\
&= T_1\cdot Tr\Big(\Omega\cdot\frac{\partial\Omega^{-1}}{\partial\sigma_i^2}\cdot W_1\cdot H_\tau\Big),
\end{aligned}
$$

Notice that

$$\Omega\cdot\frac{\partial\Omega^{-1}}{\partial\sigma_i^{-2}}=-\sigma_i^2\cdot I_{ii}$$

where the generic element of matrix I_{ii} is given by

$$\omega_{s,t}=\begin{cases}1 & s=i,j=i\\ 0 & \text{otherwise}\end{cases}$$

Therefore

$$
\begin{aligned}
E[\mathcal{H}_{1,23+i}] &= T_1\cdot\sigma_i^{-2}\cdot Tr\Big(I_{ii}\cdot W_1\cdot H_\tau\Big)=\\
&= T_1\cdot\sigma_i^{-2}\cdot\Big(W_1\cdot H_\tau\Big)_{ii}
\end{aligned}
$$

Finally we have that:

$$E[\mathcal{H}_{1,24:1:38}]=T_1\cdot diag\Big(\Omega^{-1}\cdot W_1\cdot H_\tau\Big)=T_1\cdot diag(M_2^\tau).$$

Simmetrically:

$$E[\mathcal{H}_{2,24:2:38}]=T_2\cdot diag\Big(\Omega^{-1}\cdot W_2\cdot H_\gamma\Big)=T_2\cdot diag(M_2^\gamma).$$

Going on:

$$E[\mathcal{H}_{3,3:23,23}] = E[\sum_{t=1}^{T} X_t' \cdot \Omega^{-1} \cdot X_t] = \sum_{t=1}^{T} X_t' \cdot \Omega^{-1} \cdot X_t = X' \cdot \Sigma^{-1} \cdot X$$

$$\begin{aligned} E[\mathcal{H}_{3,24:23,38}] &= E[\sum_{t=1}^{T} X_t' \cdot \frac{\partial \Omega^{-1}}{\partial \sigma^2} \cdot \varepsilon_t] \\ &= \sum_{t=1}^{T} X_t' \cdot \frac{\partial \Omega^{-1}}{\partial \sigma^2} \cdot E[\varepsilon_t] \\ &= \underset{k \times n}{\mathbf{0}} \end{aligned}$$

$$E[\mathcal{H}_{23+i,23+j|i,j\in[1,n]}] = \begin{cases} \dfrac{T}{2} \cdot \dfrac{1}{\sigma_i^4} \cdot \left(1 - \dfrac{2}{T \cdot \sigma_i^2} \cdot \displaystyle\sum_{t=1}^{T} E[\varepsilon_{i,t}^2]\right) & \forall i = j \in [1,n] \\ \\ 0 \quad \forall i \neq j \end{cases}$$

$$\begin{aligned} &= \begin{cases} -\dfrac{T}{2} \cdot \dfrac{1}{\sigma_i^4} & \forall i = j \in [1,n] \\ 0 \quad \forall i \neq j \end{cases} \\ &= -\frac{T}{2} \cdot \begin{bmatrix} \sigma_1^{-4} & 0 & \cdots & 0 \\ 0 & \sigma_2^{-4} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ o & 0 & \cdots & \sigma_n^{-4} \end{bmatrix} = -\frac{T}{2} \cdot V \end{aligned}$$

We finally have all the elements of the Fisher Information Matrix for our panel (with dummy variables) spatial model:

$$\mathcal{I} =$$

$$
\begin{bmatrix}
\begin{array}{c} -T_1 \cdot Tr\left(W_1 \cdot H_\tau \cdot W_1 \cdot H_\tau + M_1^\tau \cdot \Omega\right) - \\ - \beta' \cdot X_\tau' \cdot \left(I_{T_1} \otimes M_1^\tau\right) \cdot X_\tau \cdot \beta \end{array} & 0 & \left(X_\tau' \cdot (I_{T_1} \otimes M_2^\tau) \cdot X_\tau \cdot \beta\right)' & T_1 \cdot diag(M_2^\tau)' \\
0 & \begin{array}{c} -T_2 \cdot Tr\left(W_2 \cdot H_\gamma \cdot W_2 \cdot H_\gamma + M_1^\gamma \cdot \Omega\right) - \\ - \beta' \cdot X_\gamma' \cdot \left(I_{T_2} \otimes M_1^\gamma\right) \cdot X_\gamma \cdot \beta \end{array} & \left(X_\gamma' \cdot (I_{T_2} \otimes M_2^\gamma) \cdot X_\gamma \cdot \beta\right)' & T_2 \cdot diag(M_2^\gamma)' \\
X_\tau' \cdot (I_{T_1} \otimes M_2^\tau) \cdot X_\tau \cdot \beta & X_\gamma' \cdot (I_{T_2} \otimes M_2^\gamma) \cdot X_\gamma \cdot \beta & X' \cdot \Sigma^{-1} \cdot X & \underset{k \times n}{\mathbf{0}} \\
T_1 \cdot diag(M_2^\tau) & T_2 \cdot diag(M_2^\gamma) & \underset{n \times k}{\mathbf{0}} & -\frac{T}{2} \cdot V
\end{bmatrix}
$$

We now have all the elements to introduce our simulation procedure: since the ML estimator is asymptotically normally distributed around the true parameters, we ran a Monte-Carlo experiment by drawing the coefficients (indicated with $\tilde{\theta}$) from the following distribution:

$$\tilde{\theta} \sim \mathcal{N}(\hat{\theta}_{MLE}, \mathcal{I}(\hat{\theta}_{MLE})).$$

The term $\mathcal{I}(\hat{\theta}_{MLE})$ represents the estimated analytical Fisher Information Matrix.

At every draw we calculated the effect of tax and expenditure shock. Iterating this procedure 10,000 times allowed us to obtain a distribution of a tax and expenditure shock, thus closely mimicing the procedure adopted by Romer and Romer(2010), when they construct the confidence bands of their impulse response functions.

1.6.1 Bayesian MCMC

The model's parameters have also been estimated by the Bayesian MCMC, introduced by LeSage(1997) to provide a heteroscedastic robust estimator of the parameters of the SAR models. A Bayesian framework has been introduced since a Maximum Likelihood Heterscedasticity robust estimator was not possible to derive, because of the single-dimensional nature of the data usually employed in spatial econometrics problems. The reason to adopt such a methodology in our problem is twofold: first, we provide an alternative estimation procedure (robustness); second, we seek to improve the efficiency of the MLE estimates. In fact, LeSage(1997) shows through an experiment that the Bayesian MCMC delivers slightly more significant estimates than the ML estimator, within the homoscedastic framework. We therefore verify this fact, in our heteroscedastic, panel case:

$$
\begin{aligned}
&A_t \cdot \underset{n\times 1}{\Delta y_t} = \underset{n\times(n+6)}{X_t} \cdot \beta + \varepsilon_t \\
&A_t = I_n - \beta^{down} \cdot A_0 \cdot TB_t + \beta^{up} \cdot \hat{A}_0' \cdot EB_t \\
&\varepsilon_t \sim \mathcal{N}(0, \Omega), \forall t \in \{1, ..., T\} \\
&\Omega = \sigma^2 \cdot V \quad \text{with } V = diag(v_1, ..., v_n) \\
&\varepsilon_t \perp \varepsilon_{t+i}, \quad \forall t \in \{1, ..., T\}, \quad \forall i \in \mathcal{Z} \\
&\pi(\beta) \propto constant \\
&\pi(\sigma^2) \propto \frac{1}{\sigma^2} \\
&\pi(\frac{r}{v_i}) \overset{iid}{\sim} \chi^2_{(r)}, \quad \forall i \in \{1, ..., n\} \\
&\beta^{down} \sim Beta(d, d) \\
&\beta^{up} \sim Beta(d, d).
\end{aligned}
$$

Notice that we add prior information on the spatial coefficients: rather than letting them vary from λ_{min}^{-1} to 1, we draw it from a Beta whose support is (0,1). Such a prior was introduced by LeSage and Parent(2007); basically we rule out the possibility to have negative spatial coefficients, which is a reasonable assumption in our case, where we expect the network to be positively correlated with the dependent variable. Moreover, setting the parameter d to be close to 1, makes the Beta prior to resemble a uniform(0,1) distribution, with the advantage of putting less density on the boundaries: recall that when the spatial coefficients approach 1 (which coincide with λ_{max}^{-1}), matrix A_t becomes not invertible, and the model becomes unstable, which we believe it is a very unlikely result.
Secondly, we model the heteroscedastic terms as done in LeSage(1997); however, unlike him, we set the hyperparameter r to be equal to 3 rather than 4, as he suggests. This is because reducing the magnitude of r implies more confidence on heteroscedasticity, which is in line with our prior belief.
The Bayesian MCMC is developed independently, thus avoiding the “griddy Gibbs” procedure adopted by LeSage and Pace to overcome the huge dimension problem of standard spatial econometrics. Deriving all the formulae analytically allows to obtain more precise results, as LeSage and Pace point out. We use the standard “Metropolis within Gibbs” algorithm, and we obtain an approximation of the posterior densities for every parameter of the model. Eventually, we draw from the posteriors the parameters, like a usual MonteCarlo simulation and we use them to

construct the empirical distribution of the average total, direct and indirect effects of a tax and expenditure shock, as done before for the Maximum Likelihood Estimation.
In particular, all the steps of the procedures are:

1. Set up initial values for the parameters: $\beta_{(0)}, \sigma^2_{(0)}, V_{(0)}, \rho_{1,(0)}, \rho_{2,(0)}$.

2. Draw $\beta_{(1)}$ from the conditional posterior distribution:

$$P(\beta_{(0)}|\mathcal{D}, \sigma^2_{(0)}, V_{(0)}, \rho_{1,(0)}, \rho_{2,(0)}) = \mathcal{N}(c^*, L^*) \propto \mathcal{L}(\theta|\mathcal{D}) \cdot \mathcal{N}(c, L)$$
$$c^* = \frac{1}{T} \cdot (\sum_{t=1}^{T} X'_t \cdot V^{-1}_{(0)} \cdot X_t + \frac{\sigma^2_{(0)}}{T} \cdot L^{-1})^{-1} \cdot (\frac{1}{T} \cdot \sum_{t=1}^{T} X'_t \cdot V^{-1}_{(0)} \cdot H_t \cdot \Delta y_t + \frac{\sigma^2_{(0)}}{T} \cdot L^{-1} \cdot c)$$
$$L^* = \frac{\sigma^2_{(0)}}{T} \cdot (\sum_{t=1}^{T} X'_t \cdot V^{-1}_{(0)} \cdot X_t + \frac{\sigma^2_{(0)}}{T} \cdot L^{-1})^{-1}$$

   Notice that, setting the diagonal elements of matrix L (the prior varcov matrix) to tend to infinity (we set them up equal to 1 billion), it is like assuming a non informative prior distribution on parameter β. Notice in fact, that the parameters of the distribution tend to be equal to the FGLS estimator and its variance.

3. Draw $\sigma^2_{(1)}$ from the conditional posterior distribution:

$$P(\sigma^2_{(1)}|\mathcal{D}, \beta_{(1)}, V_{(0)}, \rho_{1,(0)}, \rho_{2,(0)}) = \Gamma^{-1}(\frac{\theta_1}{2}, \frac{\theta_2}{2}) \propto \mathcal{L}(\theta|\mathcal{D}) \cdot \Gamma^{-1}(a, b)$$
$$\theta_1 = nT + 2a \quad \theta_2 = \sum_{t=1}^{T} \varepsilon'_t \cdot V^{-1}_{(0)} \cdot \varepsilon_t + 2b$$

   In practice we draw $\sigma^2_{(1)}$ from

$$\frac{\theta_2}{\chi_{\theta_1}}$$

   Notice that, setting a and b (the prior parameters) equal to 0, is like putting a Jefferey's prior on σ^2, which is exaclty what we do.

4. Draw $v_{i,(1)}$ from the following conditional posterior distribution:

$$P(v_{i,(1)}|\mathcal{D},\sigma^2_{(1)},\beta_{(1)},\rho_{1,(0)},\rho_{2,(0)}) = \Gamma^{-1}(\frac{q_1}{2},\frac{q_2}{2}) \propto \mathcal{L}(\theta|\mathcal{D}) \cdot \Gamma^{-1}(\frac{r}{2},\frac{r}{2})$$

$$q_1 = r + T \quad q_2 = \frac{1}{\sigma^2_{(1)}} \cdot \sum_{t=1}^{T} \varepsilon^2_{i,t} + r$$

In practice we draw $v_{i,(1)}$ from:

$$\frac{q_2}{\chi_{q_1}}$$

As anticipated above in the paper, since we are confident on the heteroskedastic behavior of industry value added, we set our prior hyperparameter r to be equal to 3 rather than 4, as done by LeSage&Pace (2009).
Replicating this procedure n times, we get a first simulation of matrix $V_{(1)}$.

5. We now need to draw the spatial coefficients. However we cannot apply simple Gibbs Sampling, since the conditional posterior distribution is not defined for them. Therefore, we apply the Metropolis Hastings algorithm:

   (a) Draw ρ_1^c (where the c superscript stands for "candidate") from the proposal distribution:

   $$\rho_1^c = \rho_{1,(0)} + c_1 \cdot \mathcal{N}(0,1)$$

   (b) Run a bernoulli experiment to determine the updated value of ρ_1:

   $$\rho_{1,(1)} = \begin{cases} \rho_1^c & \pi \quad \text{(accept)} \\ \rho_{1,(0)} & 1-\pi \quad \text{(reject)} \end{cases}$$

   Where π is equal to

   $$\pi = \min\{1, \psi_{MH_1}\}$$

   Setting: $A_\tau(\rho_1) = I_n - \rho_1 \cdot W_1$ we have:

   $$\psi_{MH_1} = \frac{|A_\tau(\rho_1^c)|}{|A_\tau(\rho_{1,(0)})|} \cdot \exp\Big\{ - \frac{1}{2\sigma^2_{(1)}} \cdot \sum_{t\in t_1}^{T_1} \Big[\Delta y_t' \cdot \big(A_\tau(\rho_1^c)' \cdot V_{(1)}^{-1} \cdot A_\tau(\rho_1^c) -$$
   $$- A_\tau(\rho_{1,(0)})' \cdot V_{(1)}^{-1} \cdot A_\tau(\rho_{1,(0)})\big) \cdot \Delta y_t -$$
   $$- 2\beta' \cdot X_t' \cdot V_{(1)}^{-1}\big(A_\tau(\rho_1^c) - A_\tau(\rho_{1,(0)})\big) \cdot \Delta y_t\Big]\Big\} \cdot$$
   $$\cdot \left[\frac{\rho_1^c \cdot (1-\rho_1^c)}{\rho_{1,(0)} \cdot (1-\rho_{1,(0)})}\right]^{d-1} \cdot \mathbf{1}\big(0 \le \rho_1^c \le 1\big)$$

Basically, we compute the probability to accept the candidate value from the proposal distribution, and then we update the value of ρ_1 by running the bernoulli experiment with such a probability of success. Notice that if we draw a value of ρ_1 outside the support of the beta prior, $\psi_{MH_1} = 0$ and then $\pi = 0$ and we clearly reject the candidate value. Eventually, notice that d is the parameter of the beta prior that we set equal to 1.1, on both ρ_1 and ρ_2; this is to resemble a uniform(0,1) but with less density on its boundary values.

(c) Once updated ρ_1, we replicate the procedure for ρ_2. Setting $A_\gamma(\rho_2) = I_n - \rho_2 \cdot W_2$ we have:

$$\psi_{MH_2} = \frac{|A_\gamma(\rho_2^c)|}{|A_\gamma(\rho_{2,(0)})|} \cdot \exp\Bigg\{ - \frac{1}{2\sigma^2_{(1)}} \cdot \sum_{t \in t_2}^{T_2} \Big[\Delta y_t' \cdot \big(A_\gamma(\rho_2^c)' \cdot V_{(1)}^{-1} \cdot A_\gamma(\rho_2^c) -$$
$$- A_\gamma(\rho_{2,(0)})' \cdot V_{(1)}^{-1} \cdot A_\gamma(\rho_{2,(0)}) \big) \cdot \Delta y_t -$$
$$- 2\beta' \cdot X_t' \cdot V_{(1)}^{-1} \big(A_\gamma(\rho_2^c) - A_\gamma(\rho_{2,(0)}) \big) \cdot \Delta y_t \Big] \Bigg\} \cdot$$
$$\cdot \left[\frac{\rho_2^c \cdot (1 - \rho_2^c)}{\rho_{2,(0)} \cdot (1 - \rho_{2,(0)})} \right]^{d-1} \cdot \mathbf{1}\big(0 \leq \rho_2^c \leq 1\big)$$

6. At this point we need to update the variance of the proposal distributions: if the acceptance rate (number of acceptances over number of iterations of the Markov Chain) of the first parameter ρ_1 falls below 40% we need to reduce the value of c_1, the so called tuning parameter, which regulate the variance of the proposal distribution. We reduce the variance in this way:

$$c_1' = \frac{c_1}{1.1}.$$

In this way, we are able to draw values closer to the current value of ρ_1, and therefore, we expect to increase the acceptance rate. On the contrary, if the acceptance rate rises above 60%, we need to increase the tuning parameter, in order to draw values far from the current value, in this way we increase the chance to explore low density parts of the distribution, thus reducing the probability of accepting the candidate value and, by consequence, the acceptance rate:

$$c_1' = 1.1 \cdot c_1.$$

Clearly we replicate this procedure also for ρ_2.

7. Once updated all the values, we replicate procedure 2-6, 35,000 times.

8. We drop (*Burn-in phase*) the first 10% of the iterations, thus obtaining a vector of 31,500 observations for each of the parameters, which account for the simulated posterior distributions.

In order to obtain the distributions of the simulated fiscal plans, we simply draw the value of the parameters from the posteriors, and we calculate the effects of a simulated fiscal plan, as described in the paper.

1.7 Placebo

A reader might wonder whether the network effect we measure, is really created by an existent network or, it could be a byproduct of our particular SAR model: basically an undesired consequence of regressing the dependent variable on a convex combination of itself, which always leads to at least a weak correlation between them. If the spatial matrix we employ, which comes from real data, is really depiticing an existing network, we might expect that running our same model but, by employing a simulated spatial matrix, should deliver weaker results than what we measure by employing the original spatial matrix.
Therefore, we conducted several "Placebo" Tests, where we simulated the spatial matrix in different ways:

1. *Row Shuffling*: we permutated the elements within the rows of the original spatial matrix.

2. *Column Shuffling*: we permutated the elements within the columns of the original spatial matrix.

3. *Total Shuffling*: we shuffled all the elements of the original matrix.

4. *Half Randomization*: we constructed an artificial spatial matrix by drawing its elements from a Uniform distribution 0-0.4; since most of the elements of the original matrix are containd in such an interval. However, matrix $\hat{A}$ has been constructed by adopting the original data transformation, starting from the artificial matrix A.

5. *Full Randomization*: same as half randomization, but in this, case, we simulated also matrix $\hat{A}$.

In our procedure, we conducted 100 simulations for each Placebo Test. At every iteration we stored in a 3D array, the mean and the asymptotic t-statistics (mean over standard deviation) for each component of the Average Effect of a fiscal shock (total, direct and indirect for both taxes and expenditure shocks).
In a second stance, we plotted the results in a graph which has on the horizontal axe the mean of the Average Effect, while on the vertical axe, the asymptotic t-stats. Therefore, every siulation is summarized by a couple: the mean and asyptotic t-stats of the average effect, which in our 2D graph, represents a point. The graphs we obtain are therefore 6 scatter plots (each for shock type and component - e.g. Average Direct Effect for and Expenditure Shock), where 100 points (shown in small

blue dots) represent the 100 simulations, plus the original result (shown with a big red dot).
Since the average effects are negative, we expect to see the big red dot in the bottom left part of the graph, while, we expect the see a blue cloud of small dots shifted up and rightward: which means a weaker and less statistically significant shock effect. From the figures below we got exactly this result in all the Placebo Tests we conducted.

Figure 1: Row Shuffle

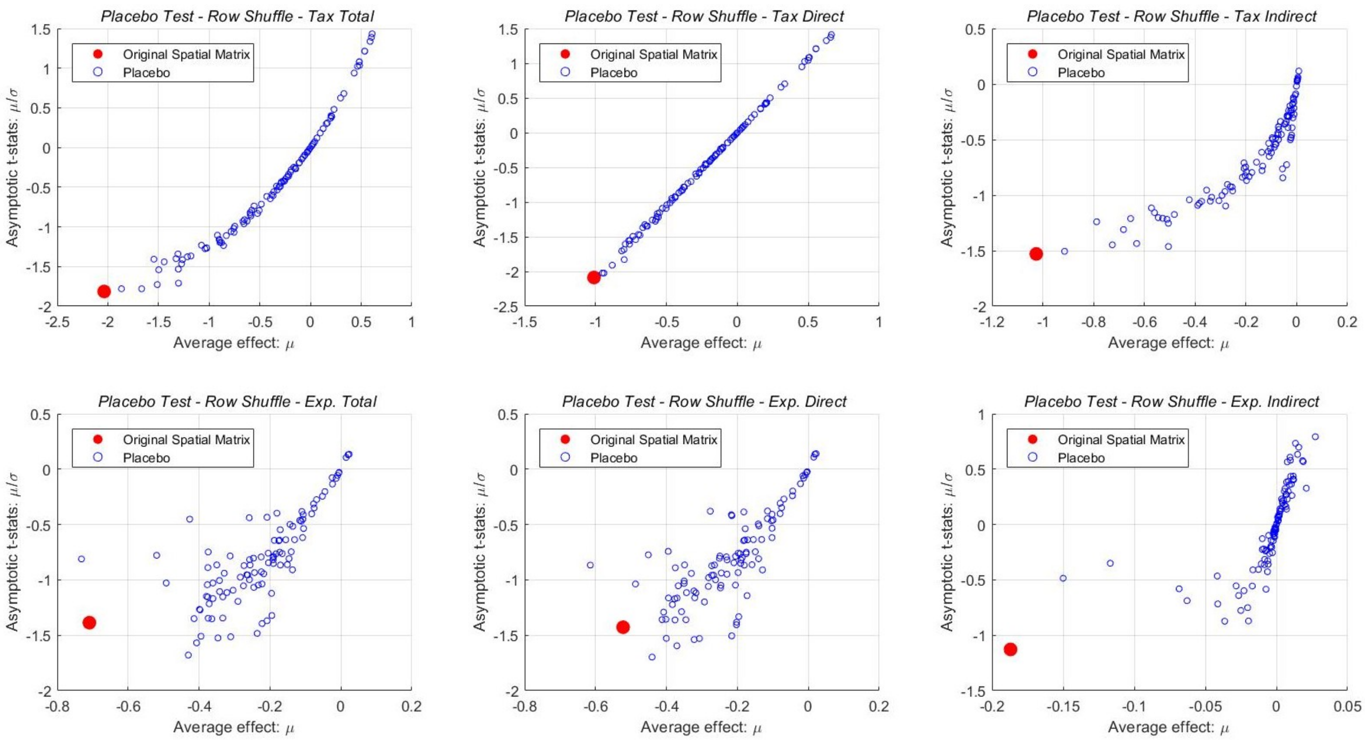


Figure 2: Column Shuffle

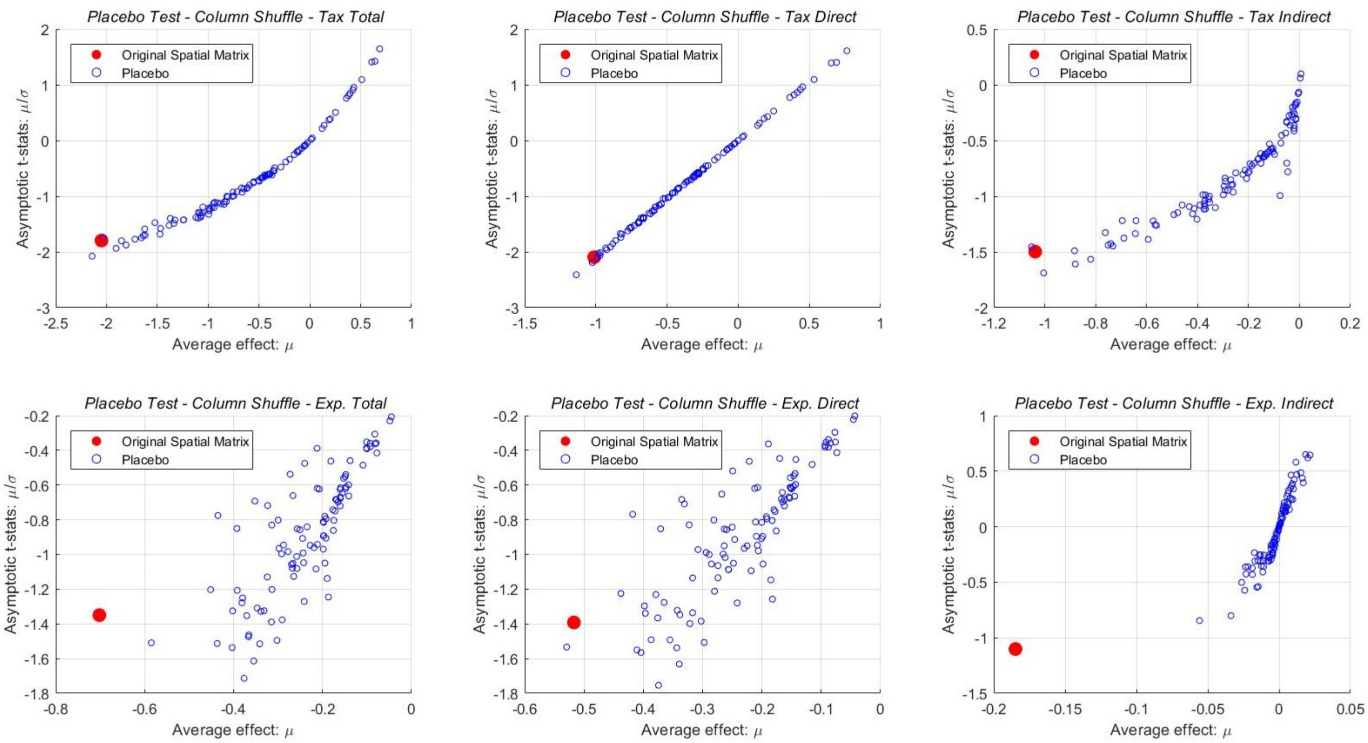


Figure 3: Total Shuffle

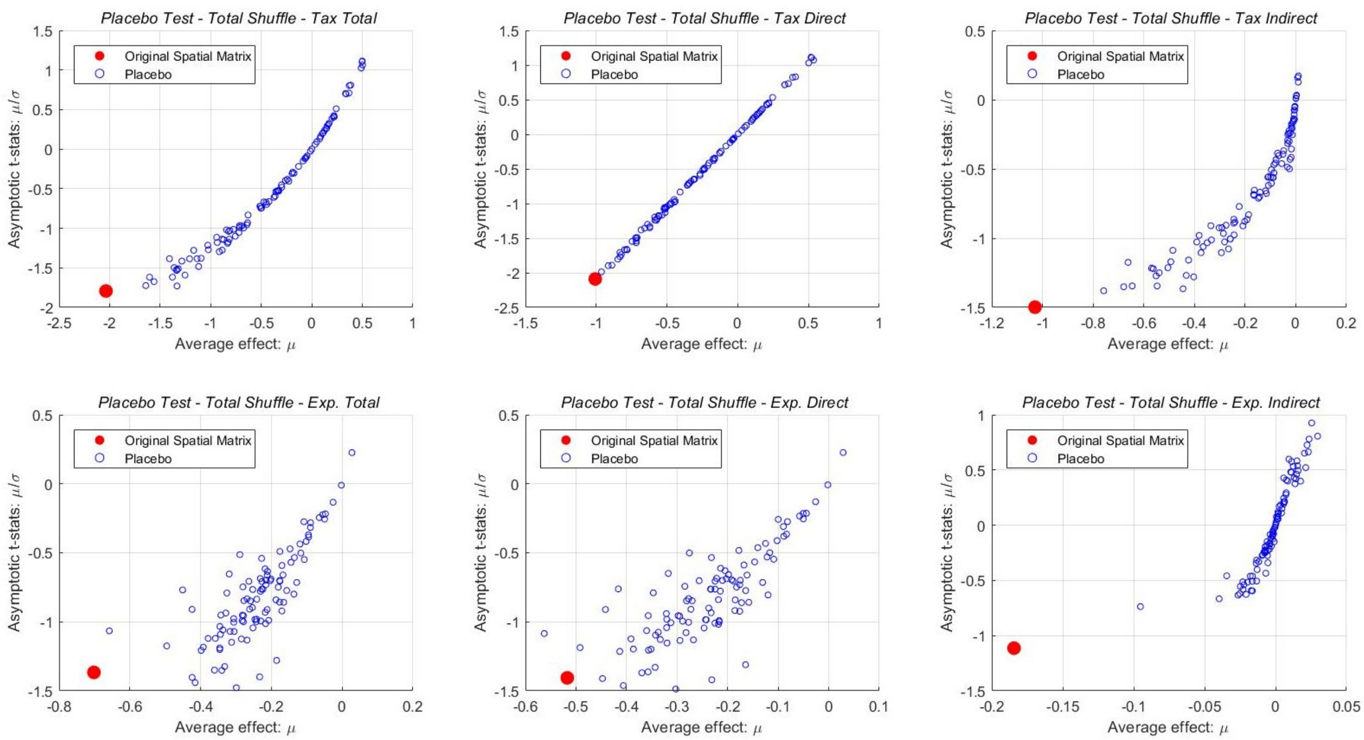


Figure 4: Half Randomization

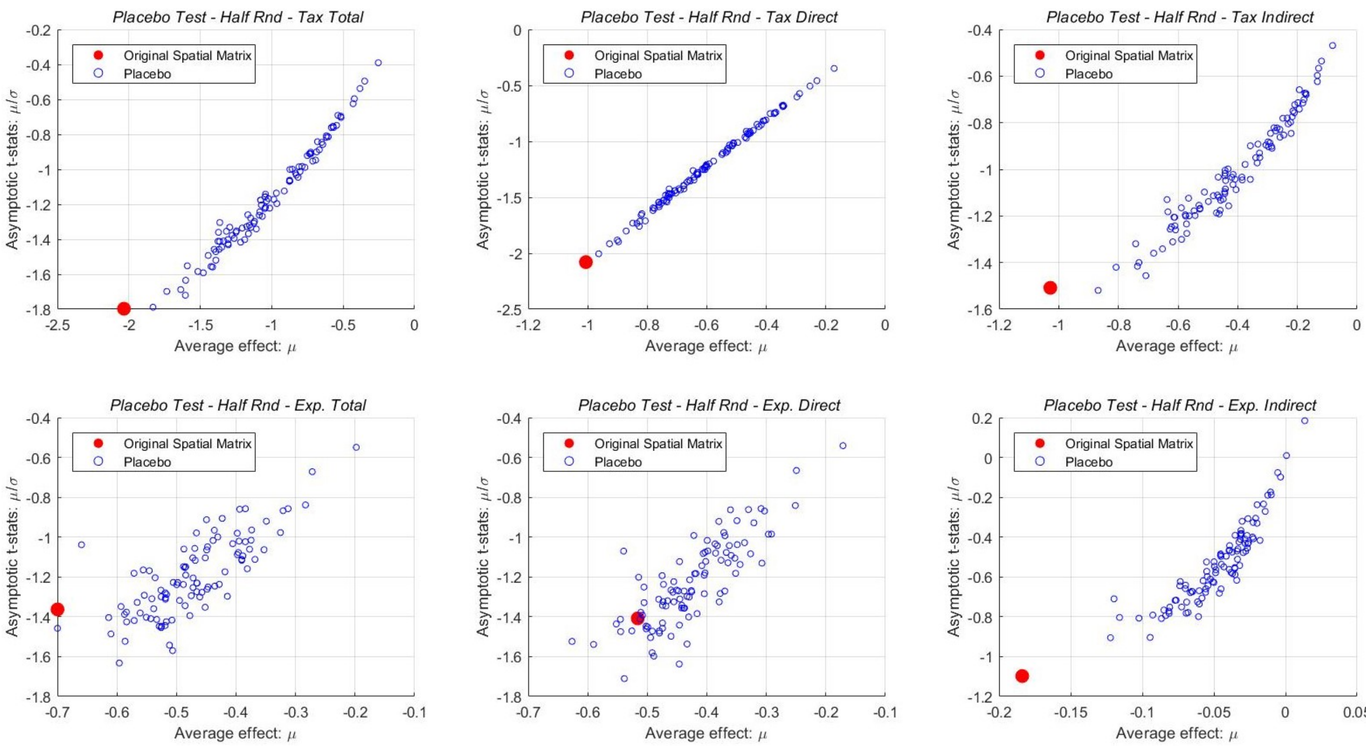
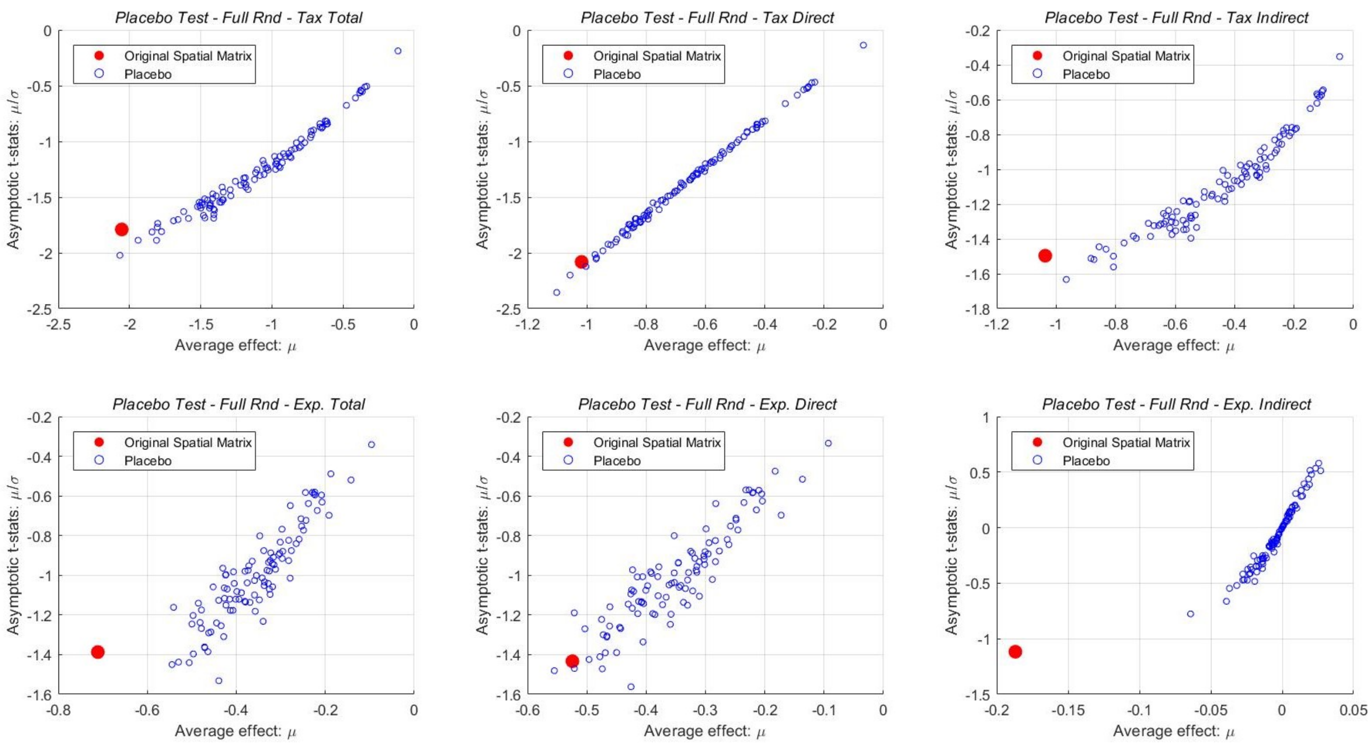


Figure 5: Full Randomization



Notice the following facts:

- The red dots always stand in the bottom left part of the graphs, as expected.
- The indirect effect, which is the one which reflects the network, always comes out weakened when using simulated data.
- The higher volatility of the expenditure shock Placebo Test, is due to the fact that the spending shocks' weights are simulated too, and therefore increase the degree of uncertainty with respect to tax shocks.

1.8 Non-Row Normalized Data

In this section we show that applying row-normalization - which is a standard practice in spatial econometric frameworks - does not affect at all our results. First of all, notice that the row-normalization changes the interpretation of the elements of the weight matrix, A (and by consequence, also of matrix $\hat{A}$). In particular, replicating what done in Section 4.2.1 of the paper, we can show the economic interpretation for each element of the weight matrix, which we denote here with $\tilde{A}$ (the row-normalized version of the I-O matrix, A). We have:

$$\tilde{A} = \begin{bmatrix} \tilde{a}_{11} = \dfrac{\text{SALES}_{1\to1}}{\text{INPUT}_1} & \tilde{a}_{12} = \dfrac{\text{SALES}_{2\to1}}{\text{INPUT}_1} & \tilde{a}_{13} = \dfrac{\text{SALES}_{3\to1}}{\text{INPUT}_1} \\ \tilde{a}_{21} = \dfrac{\text{SALES}_{1\to2}}{\text{INPUT}_2} & \tilde{a}_{22} = \dfrac{\text{SALES}_{2\to2}}{\text{INPUT}_2} & \tilde{a}_{23} = \dfrac{\text{SALES}_{3\to2}}{\text{INPUT}_2} \\ \tilde{a}_{31} = \dfrac{\text{SALES}_{1\to3}}{\text{INPUT}_3} & \tilde{a}_{32} = \dfrac{\text{SALES}_{2\to3}}{\text{INPUT}_3} & \tilde{a}_{33} = \dfrac{\text{SALES}_{3\to3}}{\text{INPUT}_3} \end{bmatrix}$$

where

$$\text{INPUT}_i = \text{SALES}_{1\to i} + \text{SALES}_{2\to i} + \text{SALES}_{3\to i}.$$

Since the coefficients change, also the interpretation of the network coefficient β^{down} (β^{up} for $\hat{A}$). However, working with non-row normalized spatial matrices, does not affect our results.
We report in V the results of the Monte Carlo simulation (starting from ML estimates), conducted with non-row-normalized spatial matrices:

Insert Table V here.

At the same time, running the inverted model using a non-row-normalized weight matrix, delivers the following results, summarized in VI:

Insert Table VI here.

1.9 Additional Tables

Table I: Table of exogenous fiscal shocks

United States: Budgetary Impact of OBRA-90[9]

Original Data	CBO:The 1990 Budget Agreement						1992 Budget of the US Government					
	1991	1992	1993	1994	1995	1991-1995	1991	1992	1993	1994	1995	1991-1995
CUMULATIVE CHANGE												
Tax	18	33	31	36	38	156	22.5	35.2	32.7	37.5	38.6	166.5
Spending	17	35	49	79	97	277						
CHANGES												
Tax	18	15	-2	5	2	38	22.5	12.7	0	0	0	35.2
Spending	17	18	14	30	18	97						

Reclassification by Calendar Year (January-December)

		DeVries et al.							Romer&Romer						
	1989	1990	1991	1992	1993	1994	1995	1990-1995	1990	1991Q1	1992	1993	1994	1995	1990-1995
CHANGES															
Tax		4.5	17.25	10.75	−0.25	4.25	1.5	38		35.2					
Spending		4.25	17.25	17	18	27	13.5	97							
Change in percent of GDP															
Tax		0.080	0.305	0.190	−0.004	0.075	0.027	0.67		0.59					
Spending		0.075	0.305	0.300	0.318	0.477	0.239	1.71							
		0.155	0.610	0.490	0.314	0.552	0.266	2.3							
Nominal GDP	5657.7	5979.6	6174.1	6539.3	6878.7	7308.8	7664.1			5888					

*: Billions of U.S. dollars, by Fiscal Year (October-September)

Table II: US OBRA-90 in percent of GDP

	Revenue adjustments									Expenditure adjustments							
		IMF	τ_t^u	$\tau_{t,0}^a$	$\tau_{t,1}^a$	$\tau_{t,2}^a$	$\tau_{t,5}^a$	IMF	g_t^u	$g_{t,0}^a$	$g_{t,1}^a$	$g_{t,2}^a$	$g_{t,3}^a$	$g_{t,4}^a$	$g_{t,5}^a$		
1990	0.00	0.080	0.080	0.00	0.305	0.190	0.00	0.075	0.027	0.075	0.075	0.00	0.305	0.300	0.318	0.477	0.239
1991	0.590	0.305	0.00	0.305	0.190	0.00	0.075	0.027	0.00	0.305	0.00	0.305	0.300	0.318	0.477	0.239	0.00
1992	0.00	0.190	0.00	0.190	0.00	0.075	0.027	0.00	0.00	0.300	0.00	0.300	0.318	0.477	0.239	0.00	0.00

Table III: descriptive statistics of estimated fixed effects for the **baseline** model. In the left panel of the table, you can see estimates for industry fixed effects, while on the right panel, you can find the estimated variances for each of the 15 sectors.

Fixed Effects					*Variances*				
	MLE	Std Dev	t-stat	p-value		MLE	Std Dev	t-stat	p-value
c_1	1.260	2.327	0.542	0.294	σ_1^2	200.129*	46.529	4.301	0.000
c_2	4.522	2.903	1.558	0.060	σ_2^2	311.043*	72.316	4.301	0.000
c_3	1.535	0.845	1.817	0.035	σ_3^2	25.436	5.914	4.301	0.000
c_4	2.336	1.099	2.125	0.017	σ_4^2	44.005	10.231	4.301	0.000
c_5	0.802	0.619	1.296	0.097	σ_5^2	13.415	3.119	4.301	0.000
c_6	2.188	0.599	3.651	0.000	σ_6^2	12.803	2.977	4.301	0.000
c_7	1.464	0.519	2.822	0.002	σ_7^2	9.530	2.216	4.301	0.000
c_8	1.906	0.672	2.835	0.002	σ_8^2	15.600	3.627	4.301	0.000
c_9	2.783	0.673	4.136	0.000	σ_9^2	15.274	3.551	4.301	0.000
c_{10}	3.092	0.421	7.352	0.000	σ_{10}^2	6.039	1.404	4.301	0.000
c_{11}	4.648	0.587	7.924	0.000	σ_{11}^2	11.614	2.700	4.301	0.000
c_{12}	4.078	0.381	10.711	0.000	σ_{12}^2	4.961	1.153	4.301	0.000
c_{13}	3.141	0.464	6.770	0.000	σ_{13}^2	7.424	1.726	4.301	0.000
c_{14}	2.343	0.609	3.845	0.000	σ_{14}^2	12.604	2.930	4.301	0.000
c_{15}	1.923	0.308	6.253	0.000	σ_{15}^2	3.041	0.707	4.301	0.000

*: The first two sectors (Agriculture and Mining) have very high variances. This is consistent with the extreme volatile nature of output in those two sectors.

Table IV: descriptive statistics of estimated fixed effects and variances for the **inverted** model.

	Fixed Effects					*Variances*			
	MLE	Std Dev	t-stat	p-value		MLE	Std Dev	t-stat	p-value
c_1	1.323	2.319	0.571	0.284	σ_1^2	198.653	46.187	4.301	0.000
c_2	4.553	2.905	1.567	0.059	σ_2^2	311.541	72.432	4.301	0.000
c_3	1.591	0.842	1.890	0.029	σ_3^2	25.228	5.865	4.301	0.000
c_4	2.220	1.098	2.021	0.022	σ_4^2	43.972	10.224	4.301	0.000
c_5	0.776	0.609	1.274	0.101	σ_5^2	12.968	3.016	4.299	0.000
c_6	2.324	0.593	3.918	0.000	σ_6^2	12.446	2.894	4.301	0.000
c_7	1.603	0.539	2.972	0.001	σ_7^2	10.216	2.375	4.301	0.000
c_8	1.890	0.674	2.804	0.003	σ_8^2	15.690	3.648	4.301	0.000
c_9	2.752	0.672	4.095	0.000	σ_9^2	15.226	3.540	4.301	0.000
c_{10}	3.201	0.423	7.560	0.000	σ_{10}^2	6.101	1.419	4.301	0.000
c_{11}	4.626	0.591	7.824	0.000	σ_{11}^2	11.815	2.747	4.301	0.000
c_{12}	4.175	0.390	10.699	0.000	σ_{12}^2	5.197	1.208	4.301	0.000
c_{13}	3.080	0.477	6.459	0.000	σ_{13}^2	7.997	1.859	4.301	0.000
c_{14}	2.279	0.609	3.743	0.000	σ_{14}^2	12.583	2.926	4.301	0.000
c_{15}	1.990	0.301	6.621	0.000	σ_{15}^2	2.908	0.676	4.300	0.000

Table V: TB and EB adjustments average effects - Baseline model - Non-Row Normalized weight matrices

	Tax Tot	Tax Dir	Tax Ind	Exp Tot	Exp Dir	Exp Ind
Point Estim.	-1.776	-0.962	-0.815	-0.886	-0.568	-0.317
Mean	-1.856	-0.965	-0.891	-0.919	-0.573	-0.346
Std Dev	1.023	0.466	0.586	0.621	0.367	0.278
$Pr(x < 0)$	98.22%	98.22%	98.22%	94.13%	94.13%	94.12%
1%	-4.706	-2.094	-2.768	-2.500	-1.426	-1.194
5%	-3.678	-1.738	-2.006	-2.000	-1.178	-0.868
10%	-3.198	-1.562	-1.678	-1.719	-1.040	-0.706
16%	-2.817	-1.419	-1.414	-1.523	-0.937	-0.598
50%	-1.777	-0.962	-0.798	-0.887	-0.572	-0.300
84%	-0.870	-0.501	-0.351	-0.319	-0.208	-0.094
90%	-0.631	-0.366	-0.243	-0.157	-0.104	-0.045
95%	-0.344	-0.204	-0.135	0.049	0.033	0.013
99%	0.173	0.103	0.065	0.399	0.264	0.128

Table VI: T Band EB adjustments average effects - Inverted model - Non-row Normalized weight matrices

	Tax Tot	Tax Dir	Tax Ind	Exp Tot	Exp Dir	Exp Ind
Point Estim.	-0.528	-0.272	-0.256	-0.475	-0.413	-0.062
Mean	-0.602	-0.298	-0.303	-0.484	-0.411	-0.073
Std Dev	0.945	0.483	0.482	0.444	0.370	0.099
$Pr(x < 0)$	73.08%	73.08%	73.08%	87.00%	87.00%	81.07%
1%	-3.189	-1.423	-1.876	-1.599	-1.275	-0.427
5%	-2.257	-1.097	-1.183	-1.243	-1.022	-0.265
10%	-1.815	-0.919	-0.918	-1.059	-0.888	-0.199
16%	-1.501	-0.777	-0.718	-0.914	-0.775	-0.156
50%	-0.523	-0.295	-0.217	-0.467	-0.409	-0.046
84%	0.300	0.186	0.111	-0.052	-0.047	0.003
90%	0.539	0.325	0.196	0.070	0.063	0.013
95%	0.810	0.506	0.303	0.223	0.197	0.031
99%	1.269	0.811	0.519	0.492	0.450	0.084

2 Additional Figures

Figure 6: Average Effects - Baseline Model

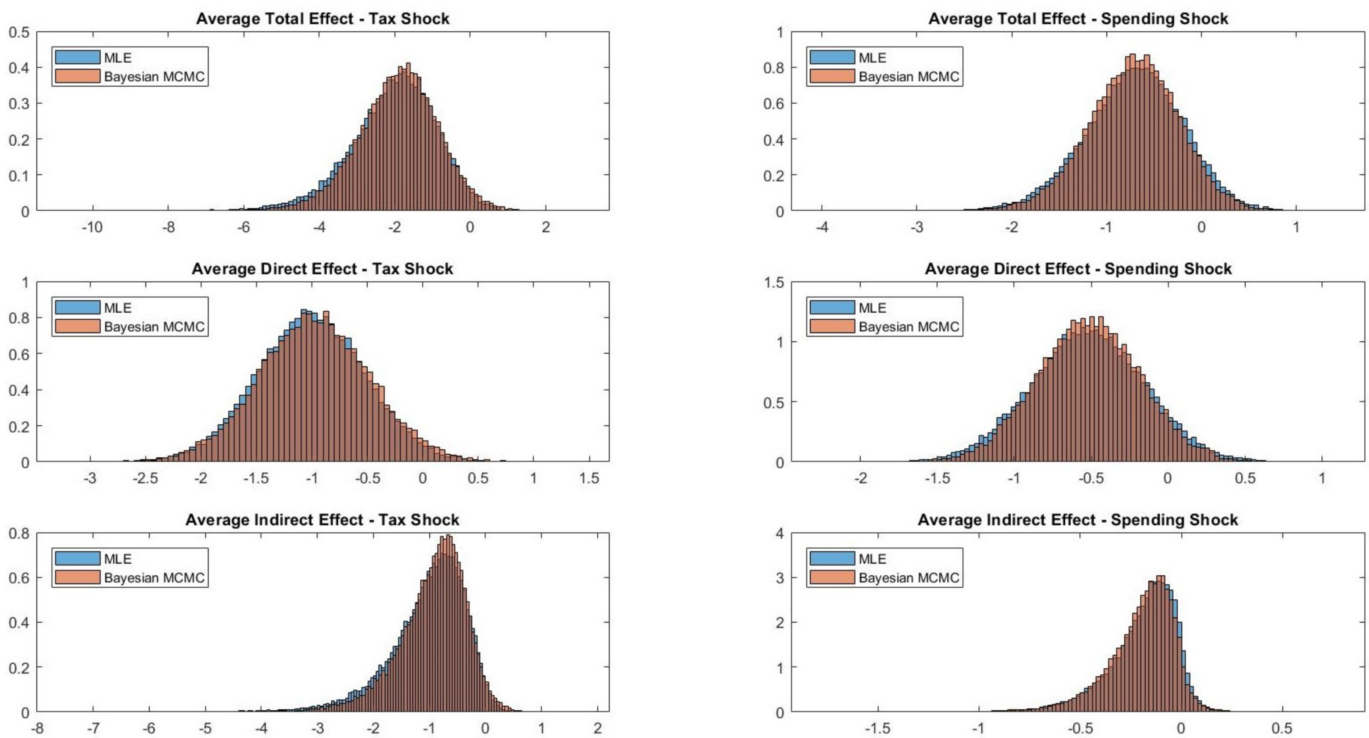


Figure 7: Average Effects - Inverted Model