IEooc_Methods5_Exercise3
Determining Sector Impact in IO Models
With the Hypothetical Extraction Method –
Hypothetical Extraction Method with Projection Matrices

Goal: *Understand the power series expansion of the A-matrix. Understand how the contribution of an individual industrial sector to the supply chain of a good or service can be identified by filtering out certain paths in the A matrix power series. Understand how the impact of several industrial sectors in the supply chain of a good or service can be determined without double-counting contribution, using projection matrices.*

Background

Environmental footprints such as the carbon footprint are calculated with the Leontief input-output model (Leontief, 1941). The IO-based footprints consider contributions from the entire supply chain of a good or service, and the supply chain is modelled as staggered star-shaped graph of numerous paths (Figure 1). Each path shows how and how much of the different precursor products enters the final product, like electricity use for coal mining for coking for steel production for car manufacturing. Footprints can be broken down by product or by industry sector relatively easily, by diagonalizing either the *y* or the (*Ly*) vectors (see also IEooc_Methods5_Software1).

A more advanced question related to footprints is the following:

How large is the part of the footprint that enters the supply chain via industry X?
or
How much of a product's footprint is associated with the input of product X into the supply chain?

For example:
What are the global impacts of the major material groups, such as biomass, metals, non-metallic minerals and fossil fuels, and how are they linked to each other in the global supply chain (Cabernard et al. 2019)?

How large is the contribution of materials (both direct and supply chain emissions) in the carbon footprint of a nation (Hertwich, 2021)?

In the IO model, information related to industry output is found in the x-vector, where x = $(I-A)^{-1}y$. An intuitive solution to the question above is to extract the material-related industry output (or generally, the focus sectors), from the x vector: $x \rightarrow x_{focus}$, and then calculate the focus sector-related footprint via $(I-A)^{-1} x_{focus}$. This approach leads to serve inconsistencies such as double-counting, however, as

- The industry output x contains not only shipments to other sectors, but also significant consumption by the sector itself. If considered as final demand leaving a subsystem of the economy, all this output will be assigned a footprint, which is over-counting.

- The products of the different focus sectors may be contained in each others' supply chains, for example, aluminium in copper (aluminium bronze), and the impacts of their production would be counted several times.
- For the border case of *all* industry sectors being in the focus, the double counting becomes obvious, as now, the final demand vector is x and not y ($x = x_{focus}$)

Using entries of the x vector directly, as an intermediate step in the supply chain, is thus not feasible.

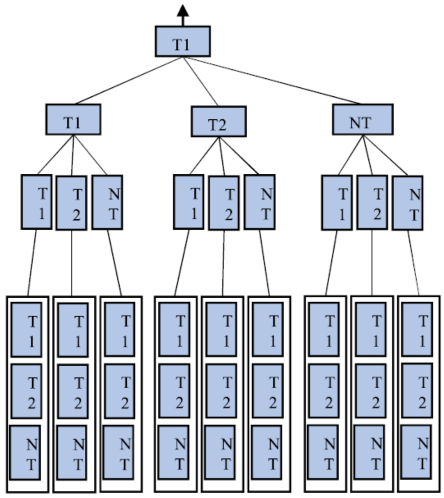

Figure 1: IO system as staggered star-shaped graph. Source: (Dente et al., 2019), © Elsevier.

Hypothetical extraction method: The solution to the identification of the impact of a certain sector in a supply chain is to remove it from the IO table and then compare the results of the two IO models, one that contains and one that does not contain the sector in question. This approach is called the *hypothetical extraction method (HEM)* (Dietzenbacher and Lahr, 2013). HEM is based on the approach that "the effect of certain economic phenomena can be measured by removing them from an input–output (I–O) table and by rebalancing the set of I–O accounts. The difference between the two sets of accounts yields the phenomenon's effect (or importance)" (Dietzenbacher and Lahr, 2013). HEM "can be used to measure the effect of changes in intermediate output, which are otherwise not easily rationalized within a Leontief framework (Dietzenbacher and Lahr, 2013). "It can also be used to estimate the possible effects of the shutdown of a particular establishment or other identifiable segment of an economy" (Dietzenbacher and Lahr, 2013).

Hypothetical extraction method for one or more industry sectors combined

First, the HEM for addressing one target sector is described. We start with the standard Leontief IO model setup.

$$\begin{aligned} x &= (I - A)^{-1} y \\ A &= Z\hat{x}^{-1} \end{aligned} \qquad (1)$$

Here, *Z* is the matrix of inter-industry flows (at scale) and *x* the total industry output vector. *A* is the matrix of technical coefficients (defined by the second row of the equation), and *y* is the exogenous final demand vector.

Now, we identify a single so-called target sector (e.g., the steel industry), and we re-calculate the IO model with all final demand and intermediate consumption of this sector's output removed (hypothetical extraction).

$$x^* = (I - A^*)^{-1} y^* \qquad (2)$$

Here, y^* is the complete exogenous final demand vector with one element (*y_target*) set to zero (the final demand of the target sector output). A^* is the complete matrix of technical coefficients with the row of input of the target sector product to other sectors set to zero. x^* is the resulting total industry output vector, which is also 0 for the single target sector. Note that the dimensions (size) of the vectors and matrices does not change.

The resulting difference in industry output x_0,

$$x_0 = x - x^* \qquad (3)$$

is the difference in amount of industry output between the full system and the almost full system with the target sector removed. It is understood as the impact or the contribution of the target sector and its entire supply chain on the supply chain of the final demand y. The impact on environmental footprints f of this sector can then be estimated by multiplying the stressor matrix S to the industry effect x_0. Again, this illustrates the meaning of the term hypothetical extraction, since x_0 is not an actual output vector but the difference of the output vectors between two IO tables.

$$f = Sx_0 \qquad (4)$$

With this method, the following two tasks are to be addressed:

Task 1) Single-sector HEM: A 10-sector IO table is given in the accompanying excel workbook IEooc_Methods5_Exercise3_HEM_IO_Workbook.xlsx. For this IO table, perform an HEM calculation for sector 6, by defining and calculating the system variables y, y^*, A, A^*, L, L^*, x, x^*, and x_0. (L is the Leontief inverse with $L = (I-A)^{-1}$.)

Task 2) Single-sector HEM: Repeat the task above for sector 4! Compare the results of task 1 and 2.

Interpretation of the single-sector HEM: Since the extracted sector is not considered, its entire output is part of this sector's impact on the supply chain of y. The contribution of the output of the other sectors in x_0 refers to the supply chain of the input into the extracted sector.

One can understand the effect of the HEM by looking at the power series expansion of $(I-A)^{-1}$:

- *y_target = 0:* All paths that link to the final demand for the target sector are set to zero: The HEM-IO model is a one where the entire supply chain of the final demand for the target sector output y_target is not produced.
- *A_target = 0:* Since an entire row of the A matrix is set to zero, all paths that link other industries to the target sector are deleted. The target sector does not produce anything in this model.
- Since the target sector does not produce anything, its own supply chain and precursor product are also not required and not produced, leading to a reduction of the output of the other sectors, which again feeds back into the economy (less precursor products also for these sectors)

Application to more than one sector: The HEM examples above delete a single sector and its entire supply chain. If one wishes to study the impact of more than one sector on the economy, the same approach can be used to identify the combined impact of the removal of several sectors:

Task 3) Single-sector HEM applied to two sectors: Repeat the task above for sectors 4 and 6 simultaneously! That means, perform the HEM as above but this time, with the final and intermediate demand of both sectors 4 and 6 removed at the same time. Compare the results of task 3 with the results of tasks 1 and 2.

Here, two sectors are removed at the same time, so that the difference output vector x_0 contains the full industry output for sectors 4 and 6.

The main finding is that for the remaining sectors the values in x_0 are smaller in the combined case than the sum of the two individual cases. The interpretation is as follows: x_0 contains the avoided production of other sectors, including sector 4 if sector 6 is extracted and vice versa.
When both sectors are extracted at the same time, their output is avoided anyway, because the respective other sector is not part of the IO table anymore anyway.
Hence, the combined extraction leads to a smaller overall reduction in input than the sum of the single sector extractions. Simply adding the effects of the single-sector HEMs would lead to double counting, since x_0 contains the avoided production of other sectors that are also extracted in the combined case.

The HEM described above can be used to study the impact of removing one or more sectors from the economy on total industrial output, carbon footprints, and other indicators calculated with the Leontief supply chain model. When applied to more than one sector at the same time, it allows us to quantify the combined effect of removing these sectors from the economy.
Because of the double-counting problem identified in task 3, it does not allow us to break down the combined effect of removing several sectors into the contributions of the individual sectors.

Here, some more thinking and a more refined approach is needed, which is described below.

Multi-sector HEM with projection matrices to quantify the contribution of different sectors to a supply chain (HEMC)

Cabernard et al. (2019) and Hertwich (2021) give a good example for the need for this method as here, the overall contribution of material production to global consumption-based emissions is estimated and then broken down into different materials.

As we saw above, using the x-vector to estimate sector-level contributions leads to double-counting of impacts. When juxtaposing the removal of different individual sectors with the HEM one by one, double-counting occurs again (see task 3) and a correction is needed to correctly estimate the contribution of individual sectors to overall consumption-based emissions.

For convenience, we call the modified HEM that allows for identifying the contribution of individual sectors *HEMC (c for contribution)*. HEMC is derived by looking at the power series expansion of the Leontief input output model:

$$x = Ly = y + Ay + A^2 y + A^3 y + A^4 y + A^5 y + \dots \quad (5)$$

We ‘walk up’ the supply chain of y, starting from the final demand, and divide all paths in the supply chain into whether they belong to one of the *T* target sectors *t* (whose impact on y we want to study) or the remaining sectors *r*. (*T* is the number of target sectors.) *The underlying accounting principle is that for each inflow i of target sector x’s product into other industries of any given step of the supply chain, the entire upstream inflow Li is accounted for as associated with target sector x.*

We start from the lowest order of the supply chain, the final demand *y*, and continue with the first order, *Ay*, and so on. At each stage $A^n y$, we have a vector of industry output, that can be divided/partitioned into *T* vectors that only have one entry: the output of a target sector at that stage of the supply chain, and one vector with the output of the remaining sectors.
For two target sectors 1 and 2, for example, we can divide the initial industry output x_0, which equals the final demand *y* into the following. (Here, $x_0 = A^0 y$ and is not to be confused with the x_0 in equ. 3.)

$$y = y_1 + y_2 + y_R \qquad (6)$$
$$y_R \equiv y^*$$

Where y_1 only contains the final demand for sector 1 and zeros else, y_2 only contains the final demand for sector 2 and zeros else, and y_R contains the other final demand items and zero for sectors 1 and 2. y_R is identical to y^* defined above.

Introduction of projection matrices: The breakdown of *y* into different sectors is easy to conceptualize, and mathematically, it can be calculated with projection matrices (*https://en.wikipedia.org/wiki/Projection_(linear_algebra)*). Here, we define the projectors as follows:

P_1 is a square matrix (total number of sectors) with only 0s except for a 1 on the diagonal in position of sector 1. Analog for P_2. The projector P_R for the remainder is the unit matrix (only 1s on the diagonal) except for the positions of sectors 1 and 2. Hence, the sum of the three projectors is the unit matrix. With the projectors defined the following relation holds:

$$\begin{aligned} y &= y \\ y &= Iy \\ y &= (P_1 + P_2 + P_R)y \\ y &= y_1 + y_2 + y_R, \quad where \\ y_1 &= P_1 y, \quad y_2 = P_2 y, \quad y_R = P_R y \end{aligned} \qquad (7)$$

Figure 2 shows how the projection matrixes and the resulting projected A matrices and y vectors look like for a 10-sector IO table with two target sectors. For more target sectors, more projectors have to be defined according to the scheme above.

P1	Filter out target sector 1							P1·A								P1·y
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0,222222	0,136364	0,071429	0,186047	0,021739	0,16	0,028571	0,125	17
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P2	Filter out target sector 2							P2·A								P2·y
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	1	0	0	0,296296	0,090909	0	0,093023	0,043478	0	0,114286	0	5
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PR = I - P1 - P2;		Filter out the remaining sectors						PR·A == A*								PR·y == y*
1	0	0	0	0	0	0	0	0,037037	0,090909	0	0,023256	0,043478	0	0,114286	0,3125	12
0	1	0	0	0	0	0	0	0	0,045455	0	0,069767	0,26087	0,04	0	0,0625	4
0	0	1	0	0	0	0	0	0,037037	0,181818	0,214286	0,046512	0	0,04	0,028571	0,125	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0,074074	0	0,071429	0,069767	0	0,08	0,142857	0,1875	30
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	1	0	0,111111	0,090909	0,071429	0,023256	0,065217	0,12	0,028571	0,0625	20
0	0	0	0	0	0	0	1	0	0,045455	0,071429	0,069767	0,195652	0	0,057143	0	0

Figure 2: Example for the use of projectors in the HEMC (screenshot from the sample solution). The target sector rows and the relevant projector matrix entries are highlighted in green.

In a next step, we calculate the entire supply chain of y_1 and y_2 as Ly_1 and Ly_2. These contributions to the total supply chain Ly are clearly attributable to sectors/products 1 and 2, respectively. For the rest y_R, we continue with the power series expansion (Figure 3, step from x_0 to x_1).

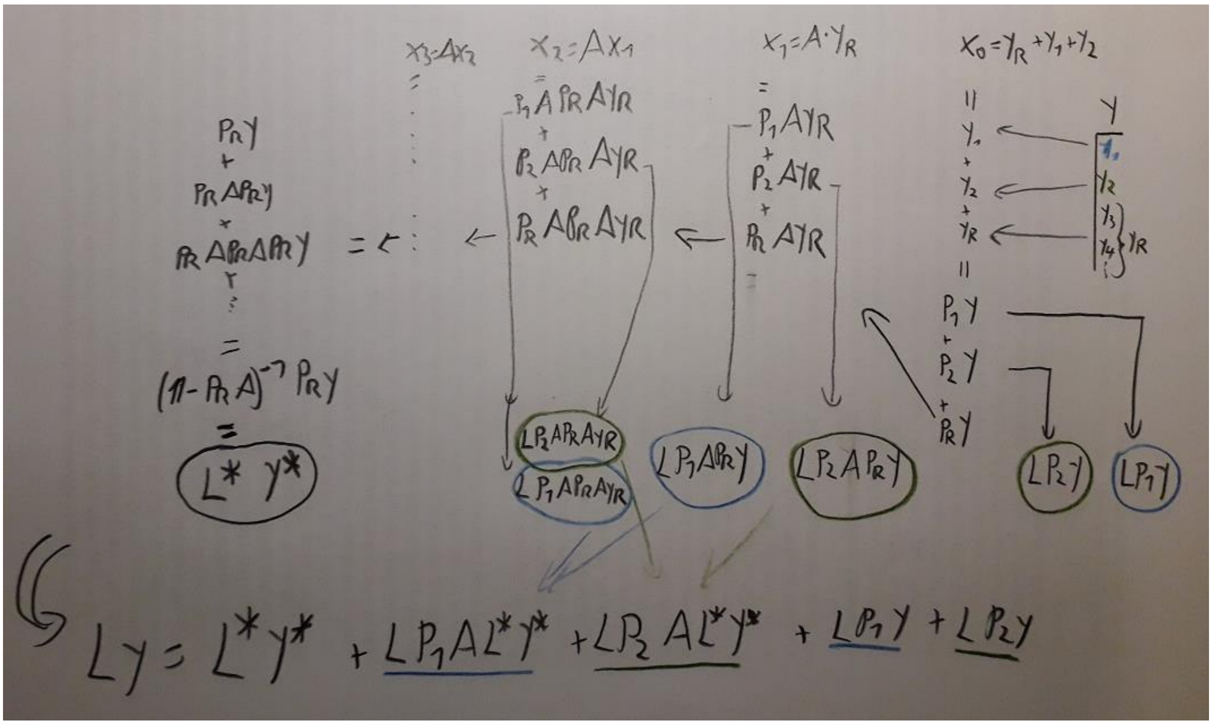

Figure 3: Calculation scheme of the HEMC (c for contribution). Blue: contributions associated with sector 1, green: contributions associated with sector 2.

The next step in the supply chain of the remaining sectors' final demand, Ay_R, is then also divided into the three contributions: the one by target sector 1, the one by target sector 2, and the remaining sectors. This division is done with the same projection matrices as defined above, and Figure 2 shows the effect of multiplying these projectors to the A matrix. The result is the following breakdown of Ay_R (Figure 3, second column from right):

$$Ay_R = IAy_R = (P_1 + P_2 + P_R)Ay_R = P_1Ay_R + P_2Ay_R + P_RAy_R \qquad (8)$$

Here, P_RA is identical to A^* defined above. The two contributions projected into sectors 1 and 2 have an entire supply chain of LP_1Ay_R and LP_2Ay_R, which also is attributed to target sectors 1 and 2, respectively. The remainder, P_RAy_R, is then again multiplied with the A-matrix to quantify the next step of the supply chain (Figure 3, center). Here, equation 8 is applied again, and the contributions of sector 1 and sector 2 are accounted for separately and the remainder, $P_RAP_RAy_R$, is again multiplied with A, and so forth.

Now we can collect the different terms of the series expansion. The remainder, i.e., those supply chains that do not contain any contributions of the target sectors, can be written as follows:

$$(I - P_RA)^{-1}P_Ry = L^*y^* \qquad (9)$$

The contributions of sectors 1 and 2 during the different intermediate steps of the supply chain can also be summarized (Figure 3, green and blue circles) as:

$$LP_1AL^*y^* \quad \text{and} \quad LP_2AL^*y^* \quad (10)$$

So that the total breakdown of the supply chain Ly into the contributions of the different sectors can be written as (Figure 3, bottom):

$$Ly = L^*y^* + LP_1AL^*y^* + LP_2AL^*y^* + LP_1y + LP_2y \quad (11)$$

This is the breakdown of the total supply chain into the contributions by different sectors and the unaffected remainder in the HEMC method. The breakdown can be interpreted as follows:

$x = Ly$	all supply chains
=	
L^*y^*	entire 'rest of the economy'
+	
$LP_1AL^*y^*$	all supply chains of target commodity 1 supplying the rest of the economy
+	
LP_1y	all supply chains of final demand in target sector 1
+	
$LP_2AL^*y^*$	all supply chains of target commodity 2 supplying the rest of the economy
+	
LP_2y	all supply chains of final demand in target sector 2

For more than two sectors, more projection matrices simply have to be added.
Again, the underlying accounting principle is that for each inflow *i* of target sector x's product into other industries of any given step of the supply chain, the entire upstream inflow *Li* is accounted for as associated with target sector x.

The resulting equation for the multi-sector SEM closely resembles the equations used in the literature (Hertwich, 2021) (Cabernard et al., 2019).

Finally, task 3 shall be repeated with the HEMC method, the results compared, and the resulting contributions to environmental footprints estimated.

<u>Task 4) multiple sector HEMC applied to two sectors, with contribution of individual sectors to total industry output and emissions:</u> In the 10 sector IO table, repeat task 3 for the two target sectors 4 and 6 and calculated the breakdown of supply chain emissions for different scopes.

a) Calculate the projected A matrices P_1A, P_2A, P_RA and y vectors P_1y, P_2y, P_Ry.

b) Calculate the different terms in equation 11 and show that this equation holds for the given example!

c) From the emissions inventory *F* (yellow cells), calculate the stressor matrix $S = F\hat{x}^{-1}$, the vector of total emissions *b* (row sum of *F*), the diagonalized vector $\hat{b}$, and the inverse $\hat{b}^{-1}$.

d) Break down the emissions associated with x_0, which is the target sectors the impact or the contribution of the target sectors and their entire supply chain on the supply chain of the final demand *y*. The size of the resulting arrays are indicated below for the different breakdowns a-d.

a. By emitting industry sector. The resulting array, F_{es} (es for emitting sector) has the same shape as F (rows: emissions, columns: emitting sectors).

b. By target sector and its entire supply chain. The resulting array, F_{tsss} (tsss for target sector supply chain) has the same number of rows as F and as many columns as there are target sectors.

c. By first consuming sector, i.e., the industry sectors and final demand that consume the output x of the target sectors. The resulting array, F_{fcs} (fcs for final consuming sector), has as many rows are there are emissions and the number of columns is the number of industry sectors plus 1 (for final demand). For each target sector, a separate F_{fcs} needs to be calculated.

d. By final demand sector. The resulting array, F_{fd} (fd for final demand) has as many rows are there are target sectors and the number of columns is the number of industry sectors plus. To solve this question, we consider F_{es} as exogenous input to the industrial system, and allocate this input to final demand sectors using the resource allocation matrix D_{res}: $D_{res} = \left(\hat{b}^{-1}S\right)\left(L\hat{y}\right)$
(See exercise IEooc_Methods6_Exercise2_Resource_tracing_IO for details.). The final result is then obtained by allocating the target sector-related emissions b_0 to the final demand sectors using D_{res}: $F_{fd} = \hat{b}_0 D_{res}$

e. Check that all breakdowns above add up to the same total for each emission!

The different breakdowns described above are used to address different research questions related to the impact of the target sectors in the economy.

References:

Cabernard, L., Pfister, S., Hellweg, S., 2019. A new method for analyzing sustainability performance of global supply chains and its application to material resources. Sci. Total Environ. 684, 164–177.

Dente, S.M.R., Aoki-suzuki, C., Tanaka, D., Kayo, C., 2019. Resources , Conservation & Recycling Effects of a new supply chain decomposition framework on the material life cycle greenhouse gas emissions — the Japanese case. Resour. Conserv. Recycl. 143, 273–281.

Dietzenbacher, E., Lahr, M.L., 2013. Expanding Extractions. Econ. Syst. Res. 25, 341–360.

Hertwich, E.G., 2021. Increased carbon footprint of materials production driven by rise in investments. Nat. Geosci. 14, 151–155.

Leontief, W.W., 1941. The structure of American economy, 1919-1929; an empirical application of equilibrium analysis. Harvard University Press, Cambridge, Mass.,.