Upstreamness and downstreamness in input-output analysis from local and aggregate information

Silvia Bartolucci
Fabio Caccioli
Francesco Caravelli
Pierpaolo Vivo

Systemic Risk Centre

Discussion Paper Series

Abstract

Ranking sectors and countries within global value chains is of paramount importance to estimate risks and forecast growth in large economies. However, this task is often nontrivial due to the lack of complete and accurate information on the flows of money and goods between sectors and countries, which are encoded in Input-Output (I-O) tables. In this work, we show that an accurate estimation of the role played by sectors and countries in supply chain networks can be achieved without full knowledge of the I-O tables, but only relying on local and aggregate information, e.g., the total intermediate demand per sector. Our method, based on a rank-1 approximation to the I-O table, shows consistently good performance in reconstructing rankings (i.e., upstreamness and downstreamness measures for countries and sectors) when tested on empirical data from the World Input-Output Database. Moreover, we connect the accuracy of our approximate framework with the spectral properties of the I-O tables, which ordinarily exhibit relatively large spectral gaps. Our approach provides a fast and analytical tractable framework to rank constituents of a complex economy without the need of matrix inversions and the knowledge of finer intersectorial details.

Silvia Bartolucci, University College London

Fabio Caccioli, University College London and Systemic Risk Centre, London School of Economics

Francesco Caravelli, Los Alamos National Laboratory

Pierpaolo Vivo, King's College London

Upstreamness and downstreamness in input-output analysis from local and aggregate information

Silvia Bartolucci*, Fabio Caccioli †, Francesco Caravelli‡ Pierpaolo Vivo§

Abstract

Ranking sectors and countries within global value chains is of paramount importance to estimate risks and forecast growth in large economies. However, this task is often non-trivial due to the lack of complete and accurate information on the flows of money and goods between sectors and countries, which are encoded in Input-Output (I-O) tables. In this work, we show that an accurate estimation of the role played by sectors and countries in supply chain networks can be achieved without full knowledge of the I-O tables, but only relying on local and aggregate information, e.g., the total intermediate demand per sector. Our method, based on a rank-1 approximation to the I-O table, shows consistently good performance in reconstructing rankings (i.e., upstreamness and downstreamness measures for countries and sectors) when tested on empirical data from the World Input-Output Database. Moreover, we connect the accuracy of our approximate framework with the spectral properties of the I-O tables, which ordinarily exhibit relatively large spectral gaps. Our approach provides a fast and analytical tractable framework to rank constituents of a complex economy without the need of matrix inversions and the knowledge of finer intersectorial details.

1 Introduction

The introduction of Input-Output (I-O) analysis as a fundamental tool to analyze the inter-relationship between economic sectors of a country was pioneered by W. Leontief, who proposed the construction of the first I-O tables for the United States for the years 1919 and 1929 [1, 2].

---

*Dept. of Computer Science, University College London, 66-72 Gower Street WC1E 6EA London (UK). Centre for Financial Technology, Imperial College Business School, South Kensington SW7 2AZ London (UK).

†Dept. of Computer Science, University College London, 66-72 Gower Street WC1E 6EA London (UK). Systemic Risk Centre, London School of Economics and Political Sciences, WC2A 2AE, London (UK). London Mathematical Laboratory, 8 Margravine Gardens, London WC 8RH (UK).

‡T-Division (Center for Nonlinear Studies and T4), Los Alamos National Laboratory, Los Alamos NM 87545 (USA).

§Dept. of Mathematics, King's College London, Strand WC2R 2LS London (UK).

An I-O table summarizes how the products (outputs) of a given industry or economic sector are used as input to other industries or sectors within the same, or different, economies (for instance, in the case of Import/Export exchanges with other countries) [3]. Understanding the structure and relevance of industrial sectors and countries within the so-called *global value chains* (GVCs), encompassing the different stages of the production process across different countries, is of central importance [4]. To achieve this, a number of indicators and measures have been devised that characterize the relative positioning of industries and economic sectors in the economy. These rely on the calculation of the following technical object,

$$G(A) = (\mathbb{1}_N - A)^{-1} \ , \tag{1}$$

the so-called *Leontief inverse* (or *resolvent*) matrix. Here, $\mathbb{1}_N$ is the $N \times N$ identity matrix (where N is the number of industrial sectors) and A is a sub-stochastic matrix with positive entries, which is related in a simple fashion to the original I-O table. Notably, the upstreamness and downstreamness metrics proposed by Antrás, Chor and collaborators (see Sec. 2 for mathematical definitions) have become widely used and mainstream in recent years [5–7]. One of the main practical challenges of the Input-Output analysis lies in the accurate and reliable compilation of inter-sectorial I-O tables from which the matrix A in formula (1) is derived. This issue is particularly felt at firm-level, where often only aggregate information is available [8].

The main contribution of our paper is to show that up-/downstreamness measures and similar resolvent-like metrics can be approximated with high accuracy even when possessing only *aggregate* and *local* information about the inter-sectorial dependencies encoded within the I-O table. In this case, the required information only amounts the row (or column) sums of the matrix A, representing the total intermediate demand per industry (or the total value of all inputs required by each industry).

More specifically, we propose an approach rooted in complexity science that reconstructs the most likely matrix A derived from I-O tables on the basis of limited/aggregated information, and uses this surrogate information to compute the Leontief inverse and related indicators (e.g., upstreamness and downstreamness). These indicators can be derived from the aggregate information available in a fast – as this procedure does not require to perform a full matrix inversion– and accurate way. Moreover, in this work we connect the accuracy of our approximate framework with the spectral properties of the I-O tables.

1.1 Related literature

There is a vast literature concerning input-output models and how inaccuracies and noise in I-O tables may affect the determination of the relative ranking of industrial sectors and countries within the economy. One strand focuses on the accuracy of the *empirical* input-output matrix denoted by A_{emp} with respect to the *true* matrix A_{true}. The main question is about how errors occurring in the compilation of the Input/Output tables propagate and affect measurements and predictions based on nonlinear functions of $A_{emp} = A_{true} + H$ (for instance, the Leontief Inverse $(\mathbb{1}_N - A_{emp})^{-1}$), where H encodes the stochastic sources of error. Compiling the entries of the matrix A_{emp} is subject to many issues, for instance the difficulty in sampling and surveying firms and flows of goods with great accuracy [9, 10]. This has provided the motivation to study

stochastic models for the input-output analysis.

Evans [11] and Quandt [12] are among the first to look at this problem by constructing random models. Evans [11] assumed that the error matrix H had only one non-zero row and that the errors could be propagated on a row-by-row basis. Quandt [12] assumed that the errors H_{ij} on the matrix elements are independent and normally distributed with mean zero, solved the error propagation problem for a small-size system (e.g. 2×2), and determined the confidence intervals on the expected Leontief Inverse. Later, Simonovits [13] deduced the fundamental inequality $\langle (\mathbb{1}_N - A_{emp})^{-1} \rangle_H \geq (\mathbb{1}_N - \langle A_{emp} \rangle_H)^{-1}$, where the average is taken with respect to independent matrix elements of H. This inequality circumvents the problem of inverting the matrix $\mathbb{1}_N - A_{emp}$, which has so far been one of the major and long-standing theoretical challenges.

One of the first comprehensive theoretical studies of stochastic input-output models is due to West [14]. His starting point is a random matrix H, of which the expected value and the standard error of all the elements are known, with the aim to provide approximating formulas for the expected value and the standard errors of the Leontief Inverse in terms of these known quantities. Some of the assumptions (for instance, that the errors H_{ij} be independent and normally distributed) are however not realistic or plainly incompatible with the sub-stochasticity constraint, and only lead to a closed-form solution for the mean and variances of the deviations from the "true" matrix under very restrictive choices for the variances of the errors in H.

More recently, this approach has been re-evaluated by Kogelschatz [15] - who assumed that the a_{ij} are Beta-distributed and derived estimates for the elements of the Leontief Inverse - and Kozicka [16] - who postulated more realistic distribution for the matrix entries, but provided explicit formulae only for small-size systems.

Within the empirical literature, a number of studies have been also undertaken to characterize the regional inter-sectorial dependence of industries and to discuss the challenges of reconstructing regional data from national accounts and surveys [17].

Given the practical difficulties associated with compiling input-output tables, especially at the regional level, earlier scholars devised "shortcut" methods to estimate the Leontief inverse from incomplete or unreliable information, or even foregoing I-O tables altogether. Katz and Burford [18, 19] derived a formula under the assumption that the matrix A is uniformly drawn from the set of sub-stochastic matrices, and under the rather questionable technical condition that the covariance between the entries of the matrix and the output multipliers be null. Their work hinges on an earlier formula empirically derived by Drake [20]. The general approach based on finding "shortcuts" and foregoing a painstaking compilation of I-O tables was criticized on both technical and conceptual grounds [21–24] before this line of investigation was dropped and even ignored altogether in the subsequent related literature.

The Leontief inverse and the associated indicators have also been looked at through the prism of complexity and network science. Cerina et al. [25] analyzed the properties of the (global and regional) network of industries in different economies reconstructing the monetary goods flows (edges) using the input-output matrix. McNerney et al. [26] used average national output multipliers to predict future economic growth and price changes. In [27], a model for the propagation and amplification of idiosyncratic shocks along the input-output network is provided. In [28], a network analysis of the World Input-Output Data set is undertaken to

analyze the temporal interdependence between countries and industrial sectors.

In recent years the interest in input-output models has grown steadily [29], also in view of a rather compelling connection to models of complexity and networks [28, 30]. Moreover, many of these ideas can in principle be extended to more general sector-product spaces, which saw many uses for the study of the connection between complexity measures, productivity and economic growth [31–34].

In the following section, we will focus on the works by Antràs and Chor [4], Fally et al. [6] and Miller et al. [7], where different incarnations of the so-called upstreamness and downstreamness measures have been first proposed. An early example of a direct application of those measures for the analysis of empirical data on global value chains can be found in [35], now used in multiple contexts [36, 37].

2 Definition of Upstreamness and Downstreamness

Antràs et al. [4] considered a closed economy of N industries. For each industrial sector $i = 1, \ldots, N$ we indicate the value of gross output with Y_i and the total intermediate demand (i.e., the use of the output of an industry as a final good) with F_i. Then the following equality holds in Input-Output tables:

$$Y_i \;\; = F_i + Z_i = F_i + \sum_{j=1}^{N} a_{ij} = \tag{2}$$

$$= F_i + \sum_{j=1}^{N} d_{ij} Y_j \; , \tag{3}$$

with $Z_i = \sum_{j=1}^{N} d_{ij} Y_j$ corresponding to the output of industry i used as intermediate input to other industries (*intermediate demand*) as shown in the scheme in Fig. 1. In [4], $\{d_{ij}\}$ corresponds to the dollar amount of sector i's output used to produce one dollar worth of sector j's output and it is related to the matrix A via the relationship $d_{ij} Y_j = a_{ij}$. The final demand, as detailed in Sec. 4, comprises contributions from different factors including, among others, the final consumption expenditure by households and government, and exports.

Iterating the identity Eq. (2) within Eq. (3), one obtains an infinite sequence of contributions, each representing the use of sector i's output at different levels within the value chain [3]

$$Y_i = F_i + \sum_{j=1}^{N} d_{ij} F_j + \sum_{j=1}^{N} \sum_{k=1}^{N} d_{ik} d_{kj} F_j + \ldots \; . \tag{4}$$

We can finally rewrite Eq. (4) as follows

$$\boldsymbol{Y} = [\mathbb{1}_N - D]^{-1} \boldsymbol{F} \tag{5}$$

using $\sum_{k \geq 0} D^k = [\mathbb{1}_N - D]^{-1}$. In this case, $\mathbb{1}_N$ is the $N \times N$ identity matrix, $D = (d_{ij})$ contains each sector's output in dollar values and $\boldsymbol{F}$ is the vector of final demands. Antràs et al. [4] hence proposed the following measure of upstreamness of the i-th industrial sector

$$U_{1i} = 1 \cdot \frac{F_i}{Y_i} + 2 \cdot \frac{\sum_{j=1}^{N} d_{ij} F_j}{Y_i} + 3 \cdot \frac{\sum_{j,k=1}^{N} d_{ik} d_{kj} F_j}{Y_i} + \ldots = \frac{([\mathbb{1}_N - D]^{-2} \boldsymbol{F})_i}{Y_i} \; , \tag{6}$$

Figure 1: Scheme of the structure of a single-country Input-Output table [3, 38, 39].

where each term contributing to Eq. (4) is weighted by their distance from final use and divided by the output of the sector Y_i. The notation $(\cdot)_i$ is used to indicated the i-th component of the vector. By construction, the terms of the sum that are further upstream in the value chain carry larger weight in the calculation of the upstreamness. Inserting Eq. (4) in Eq. (6), we can rewrite the upstreamness as

$$\boldsymbol{U}_1 = [\mathbb{1}_N - A_U]^{-1}\boldsymbol{1}_N \; , \tag{7}$$

where

$$A_U = Y^{-1}A = \begin{pmatrix} \frac{a_{11}}{Y_1} & \cdots & \frac{a_{1N}}{Y_1} \\ \vdots & \ddots & \vdots \\ \frac{a_{N1}}{Y_N} & \cdots & \frac{a_{NN}}{Y_N} \end{pmatrix} \tag{8}$$

and $Y = \mathrm{diag}(Y_1, \ldots, Y_N)$. The vector $\boldsymbol{1}_N$ is a column vector of N ones. The matrix A_U has non-negative elements, and in this convention it is row-substochastic, i.e., $\sum_j (A_U)_{ij} \leq 1 \; \forall j$. By construction $U_{1i} \geq 1$ and it is precisely equal to 1 if no output of industry i is used as input to other industries but it is only used to satisfy the final demand.

Later, Antràs et al. [5] also established an equivalence between their upstreamness measure and a measure – defined in a recursive fashion – of the "distance" of an industry from the final demand proposed independently by Fally et al. [6]. Fally's upstreamness U_2 is defined as follows:

$$U_{2i} = 1 + \sum_{j=1}^{N} \frac{d_{ij}Y_j}{Y_i} U_{2j} \; . \tag{9}$$

The idea is that $\boldsymbol{U}_2$ aggregates information on the extent to which a sector in a given country produces goods that are sold directly to final consumers or that are sold to other sectors that themselves mainly sell to final consumers. Sectors selling a large share of their output to relatively upstream industries should be therefore considered to be more upstream themselves.

Using the fact that $d_{ij}Y_j = a_{ij}$ we obtain

$$\boldsymbol{U}_2 = [\mathbb{1}_N - A_U]^{-1}\boldsymbol{1}_N \ , \tag{10}$$

where A_U is defined in Eq. (8) as presented in [5].

On the input-side, there exists an analogous accounting identity stating that sector i's total input Y_i is equal to the value of its primary inputs (the so-called value added) V_i plus its intermediate input purchased from all other sectors, namely

$$Y_i = V_i + Z_i = V_i + \sum_{j=1}^{N} a_{ji} = V_i + \sum_{j=1}^{N} d_{ji}Y_j \ , \tag{11}$$

and

$$\boldsymbol{Y} = [\mathbb{1}_N - D^T]^{-1}\boldsymbol{V} \ . \tag{12}$$

Similarly to Antràs et al. (cf. Eq. (6)), Miller and Temurshoev [7] introduced the so-called *downstreamness*, measuring the "average distance between suppliers of primary inputs and sectors as input purchaser along the input demand supply chain" as follows:

$$D_{1i} = 1 \cdot \frac{V_i}{Y_i} + 2 \cdot \frac{\sum_{j=1}^{N} V_j d_{ji}}{Y_i} + 3 \cdot \frac{\sum_{j,k=1}^{N} V_j d_{jk} d_{ki}}{Y_i} + \ldots = \frac{([\mathbb{1}_N - D^T]^{-2}\boldsymbol{V})_i}{Y_i} \ . \tag{13}$$

As before, using Eq. (12), we obtain

$$\boldsymbol{D}_1 = [\mathbb{1}_N - A_D]^{-1}\boldsymbol{1}_N \ , \tag{14}$$

with

$$A_D = (AY^{-1})^T = \begin{pmatrix} \frac{a_{11}}{Y_1} & \cdots & \frac{a_{N1}}{Y_1} \\ \vdots & \ddots & \vdots \\ \frac{a_{1N}}{Y_N} & \cdots & \frac{a_{NN}}{Y_N} \end{pmatrix} \ . \tag{15}$$

The matrix A_D has non-negative elements, and it is row-substochastic, i.e., $\sum_j (A_D)_{ij} \leq 1 \ \forall j$. Finally, as in the upstreamness case, also for the downstreamness, Fally [6] introduced an analogous iterative definition of the form

$$D_{2i} = 1 + \sum_{j=1}^{N} d_{ji} D_{2j} \ , \tag{16}$$

which can be again mapped with simple manipulations onto Eq. (14) using $Y_i d_{ji} = a_{ji}$.

3 Rank-1 approximation with local and aggregate information

In this section, we will discuss how to derive an approximation for the upstreamness and downstreamness metrics discussed in Sec. 2. Let us consider the resolvent $G(A) = (\mathbb{1}_N - A)^{-1}$, where the matrix A stands for A_U or A_D as defined in the previous section. Therefore, A has non-negative entries, and is sub-stochastic. Recall that the vectors of upstreamness and downstreamness are defined as $\boldsymbol{U}_1 = G(A_U)\mathbf{1}_N$ and $\boldsymbol{D}_1 = G(A_D)\mathbf{1}_N$, respectively (cf. Eq. (10), (14)). We are going now to assume that a detailed and accurate knowledge of all the entries of A is *not* available. The only available aggregate information is given by the $2N$ constants $\boldsymbol{r} = (r_1, \ldots, r_N)$ and $\boldsymbol{c} = (c_1, \ldots, c_N)$, namely the *sums* of the N rows and columns of A. This corresponds to knowing only the total intermediate demand per industry and the total value of all inputs required by each industry respectively. This lack of detailed information is actually quite common in supply chain and intrafirm network analysis [8], which in turn leads to the need for inference and reconstruction methods to fill the gaps.

A simple rank-1 approximation $\hat{A}$ for the matrix A is

$$\hat{A} = \frac{1}{N}\boldsymbol{g}\boldsymbol{q}^T = \begin{pmatrix} \frac{g_1 q_1}{N} & \cdots & \frac{g_1 q_N}{N} \\ \vdots & \ddots & \vdots \\ \frac{g_N q_1}{N} & \cdots & \frac{g_N q_N}{N} \end{pmatrix} \ , \tag{17}$$

where the entries of the column vectors $\boldsymbol{g} = (g_1, \ldots, g_N)$ and $\boldsymbol{q} = (q_1, \ldots, q_N)$ are determined imposing the constraint that A and $\hat{A}$ share the same row and column sums

$$r_i = \sum_j A_{ij} \equiv \frac{\sum_k q_k}{N}g_i = \bar{q}\ g_i \ , \tag{18}$$

$$c_j = \sum_i A_{ij} \equiv \frac{\sum_k g_k}{N}q_i = \bar{g}\ q_j \ . \tag{19}$$

This yields eventually the unique matrix

$$\hat{A} = \frac{1}{mN}\boldsymbol{r}\boldsymbol{c}^T \tag{20}$$

with $m = \frac{1}{N}\sum_{ij} A_{ij} = \frac{1}{N}\sum_j c_j = \frac{1}{N}\sum_i r_i$. The rank-1 matrix $\hat{A}$ in (20) is the so-called Maximum Entropy reconstructed matrix (see e.g. [40, 41]) subject to the row and column constraints in (18) and (19) (see also [42–46] for related works).

If the only information we have is about row sums, then the corresponding rank-1 approximation is even simpler

$$\hat{A} = \begin{pmatrix} \frac{r_1}{N} & \cdots & \frac{r_1}{N} \\ \vdots & \ddots & \vdots \\ \frac{r_N}{N} & \cdots & \frac{r_N}{N} \end{pmatrix} \ . \tag{21}$$

Clearly, $\hat{A}$ has a single non-zero, real and positive eigenvalue $\lambda_1 = \frac{1}{mN}\sum_j r_j c_j$ (or $\lambda_1 =$

$\frac{1}{N} \sum_j r_j$ in the case of only-row constraints) due to the Perron-Frobenius theorem, and $N - 1$ zero eigenvalues, therefore we may expect that this approximation will work better the larger the *spectral gap* (or equivalently the smaller the *spectral radius* in the bulk) [1] of the original matrix A is [47, 48]. The empirical I-O matrices A_U, A_D typically show a large spectral gap, suggesting that the rank-1 approximation described in this section should be very effective.

As the empirical I-O matrices A_U, A_D are rather small ($N = 35$), it is more informative to look at their spectral radius. In Sec. 5 we perform a thorough analysis of the spectra of the I-O matrices at the country level and we study how the accuracy of our rank-1 formula is related to the spectral radius. We indeed find that there is a clear negative correlation between the two, i.e. the error made using our approximation increases with Ξ. This said, even in the worst cases, the relative errors remain fairly negligible and the formulae work very well across the entire dataset.

Employing this rank-1 approximation, we can now evaluate the approximate resolvent

$$G(\hat{A}) = (\mathbb{1} - \hat{A})^{-1} = \mathbb{1} + \frac{\hat{A}}{1 - \frac{1}{mN} \sum_j r_j c_j} \ , \tag{22}$$

using the Sherman-Morrison formula [49] for the inverse of a rank-1 matrix, from which it follows that the upstreamness and downstreamness of the i-th industry are respectively approximated by

$$U_{1i} \approx 1 + \frac{r_i}{1 - \frac{1}{mN} \sum_j r_j c_j} \tag{23}$$

$$D_{1i} \approx 1 + \frac{\tilde{r}_i}{1 - \frac{1}{\tilde{m}N} \sum_j \tilde{r}_j \tilde{c}_j} \ , \tag{24}$$

where r_i, c_i and $\tilde{r}_i, \tilde{c}_i$ represent respectively the sum of rows and columns of A_U and A_D. If only the constraint on row is imposed, the formulae above reduce to

$$U_{1i} \approx 1 + \frac{r_i}{1 - \frac{1}{N} \sum_j r_j} \tag{25}$$

$$D_{1i} \approx 1 + \frac{\tilde{r}_i}{1 - \frac{1}{N} \sum_j \tilde{r}_j} \ . \tag{26}$$

The approximate formulae above show that, within our rank-1 approximation, the upstreamness (downstreamness) of sector i is fully determined by the interplay of (i) *local* and *aggregate* information, namely of the total intermediate demand per sector (and/or the total value of all inputs required by a each sector), and (ii) a suitable average of the total intermediate demand (and/or the total value of all inputs) across *all* sectors in the economy.

In spite of the seemingly drastic approximation, which neglects a significant amount of finer intersectorial details, we will show that the aggregate information featuring in our rank-1 formulae is sufficient to determine with high accuracy the relative positioning of countries and

---

[1]The spectral gap is defined as $\Gamma = \lambda_1 - \Xi$, with λ_1 real and < 1 being the Perron-Frobenius eigenvalue. The spectral radius is $\Xi = \max\{|\lambda_2|, \ldots, |\lambda_{N-1}|\}$.

sectors within the global value chains.

In the next sections, we will then calculate upstreamness and downstreamness measures on I-O tables from the NIOT Dataset (see Sec. 4), comparing the results obtained via our approximation with the full calculation using the original formulae, namely Eq. (10) and (14).

4 Dataset

Australia (AUS)	France (FRA)	Netherland (NLD)
Austria (AUT)	Great Britain (GBR)	Poland (POL)
Belgium (BEL)	Greece (GRC)	Portugal (PRT)
Bulgaria (BGR)	Hungary (HUN)	Romania (ROU)
Brazil (BRA)	Indonesia (IDN)	Russia (RUS)
Canada (CAN)	India (IND)	Slovakia (SVK)
China (CHN)	Ireland (IRE)	Slovenia (SVN)
Cyprus (CYP)	Italy (ITA)	Sweden (SWE)
Czech Republic (CZE)	Japan (JPN)	Turkey (TUR)
Germany (DEU)	Korea (KOR)	Taiwan (TWN)
Denmark (DNK)	Lituania (LTU)	United States (USA)
Spain (ESP)	Latvia (LVA)	Luxembourg (LUX)*
Estonia (EST)	Mexico (MEX)	
Finland (FIN)	Malta (MLT)	

Table 1: Countries and their codes in the NIOT database by WIOD [39]. Luxembourg is not included in our analysis as data present inconsistencies across the years.

The empirical I-O matrices used for the experiments have been constructed using the 2013 release of the National Input-Output tables by the World Input-Output Database (WIOD) [39]. The NIOT dataset comprises 39 countries –representing a large fraction of the major world economies – over the years 1995 - 2011. The list of countries and their codes considered in our empirical analysis is presented in Tab. 1. The structure of the input-output table of each country is schematically shown in Fig. 1. The intermediate demand for each country is reported for $N = 35$ economic sectors in terms of the flow (in US million dollars) between sectors. The full list of economic sectors and their codes included in our analysis is summarized in Tab. 2. The final demand is characterized in terms of (i) final consumption expenditure by households, (ii) final consumption expenditure by non-profit organizations serving households (NPISH), (iii) final consumption expenditure by government, (iv) gross fixed capital formation, (v) changes in inventories and valuables and (vi) exports. In the dataset sometimes the change in Inventories and Valuables can be negative, and were assumed to contribute to imports. The entries a_{ij} of each row of the full I-O table are then normalized by the vector outputs Y_j. The normalized intermediate demand sub-matrix is sub-stochastic and represents the matrix A_U. The r_i used in the model are simply the sums over the rows of the matrix A_U (or equivalently if normalized by columns the matrix A_D, respectively in Eq. (8) and (15)) .

Agriculture, Hunting, Forestry and Fishing	AGR
Mining and Quarrying	MIN
Food, Beverages and Tobacco	FOD
Textiles and Textile Products	TXT
Leather, Leather and Footwear	LEA
Wood and Products of Wood and Cork	WOO
Pulp, Paper, Paper , Printing and Publishing	PRT
Coke, Refined Petroleum and Nuclear Fuel	COK
Chemicals and Chemical Products	CHM
Rubber and Plastics	RUB
Other Non-Metallic Mineral	NMM
Basic Metals and Fabricated Metal	MET
Machinery, Nec	MAC
Electrical and Optical Equipment	ELO
Transport Equipment	TRA
Manufacturing, Nec; Recycling	MAN
Electricity, Gas and Water Supply	ELE
Construction	CON
Sale, Maintenance and Repair of Motor Vehicles and Motorcycles; Retail Sale of Fuel	MOT
Wholesale Trade and Commission Trade, Except of Motor Vehicles and Motorcycles	WHO
Retail Trade, Except of Motor Vehicles and Motorcycles; Repair of Household Goods	RET
Hotels and Restaurants	HOT
Inland Transport	ITR
Water Transport	WAT
Air Transport	AIR
Other Supporting and Auxiliary Transport Activities; Activities of Travel Agencies	OTR
Post and Telecommunications	POS
Financial Intermediation	FIN
Real Estate Activities	EST
Renting of MEq and Other Business Activities	REN
Public Admin and Defence; Compulsory Social Security	PUB
Education	EDU
Health and Social Work	HEA
Other Community, Social and Personal Services	SOC
Private Households with Employed Persons	HOU

Table 2: Sectors of the NIOT dataset by WIOD (2013 release) and their sector codes [39].

5 Results

In this section, we compare our approximate formulae for downstreamness and upstreamness with single (Eq. (25) and (26) respectively) and double contraints (Eq. (23) and (24) respectively) with the measures obtained via direct inversion of the empirical I-0 matrix (Eq. (10) and (14) respectively).
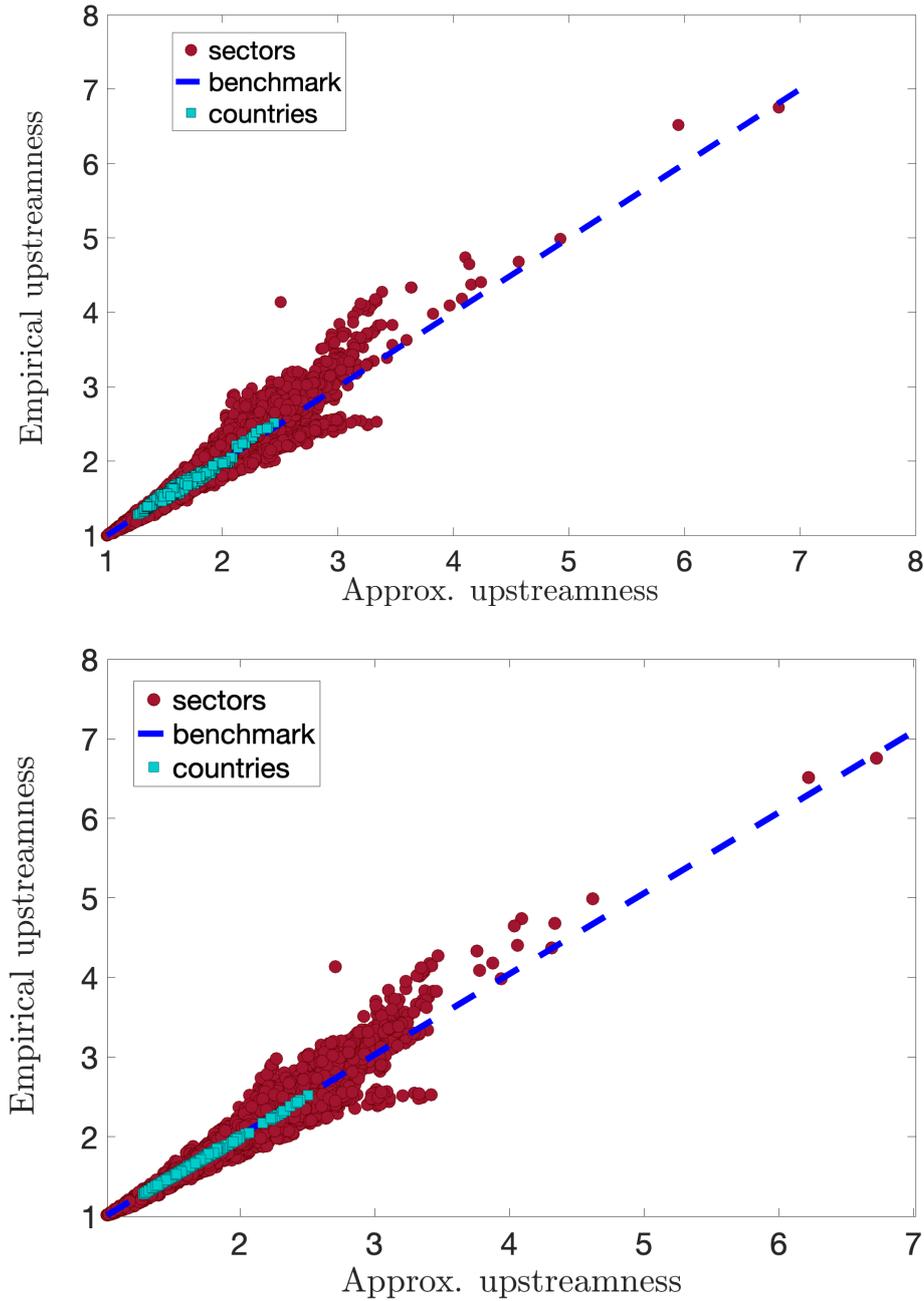
Figure 2: Empirical upstreamness versus approximated upstreamness. Cyan squares represent the upstreamness per country (39 countries) per year (11 year) averaged over 35 industrial sectors from the WIOD dataset (Release 2013). Red full circles represents the upstreamness for all industry sectors in all countries/all years. Top panel: Empirical upstreamness compared with single-constraint approximation in Eq. (25). Bottom panel: Empirical upstreamness compared with double-constraints approximation in Eq. (23).
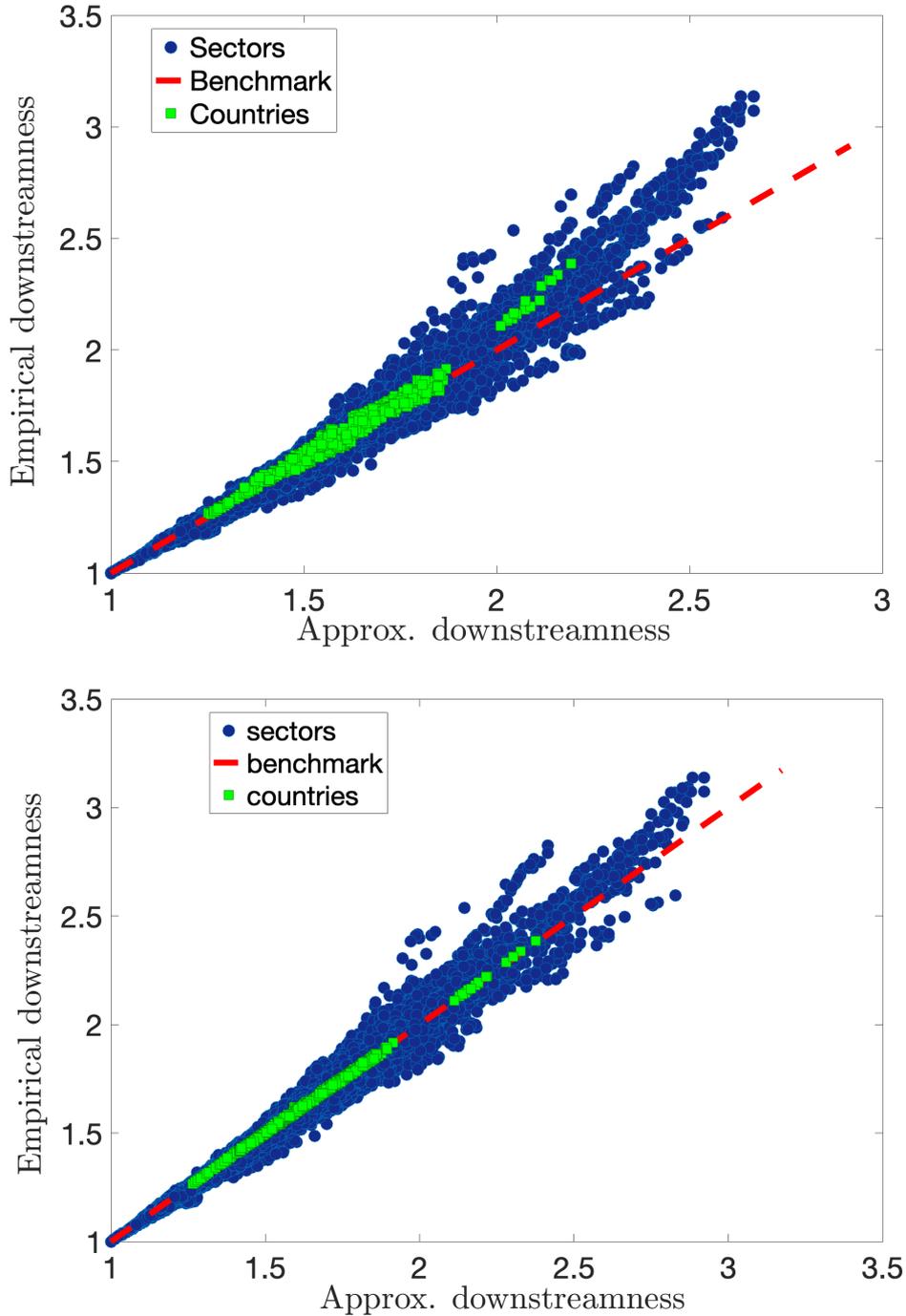
Figure 3: Empirical downstreamness versus approximated downstreamness. Light green squares represent the downstreamness per country (39 countries) per year (11 year) averaged over 35 industrial sectors from the WIOD dataset (Release 2013). Blue full circles represent the down-streamness of all industry sectors in all countries/all years. Top panel: Empirical upstreamness compared with single-constraint approximation in Eq. (26). Bottom panel: Empirical upstream-ness compared with double-constraints approximation in Eq. (24).

In Fig. 2 we plot the empirical average over all sectors (cyan squares) of the *upstreamness*

for 39 countries (listed in Table 1) for all years (1995-2011) versus the approximate value with single (top panel) and double constraints (bottom panel), respectively obtained in Eq. (25) and Eq. (23). We see that the empirical data (663 data points - 39 countries × 17 years) nicely collapse on top of the theoretical benchmark (blue dashed line). In the single constraint case, this implies that the average upstreamness coefficient for a country is determined with high accuracy by the knowledge of a single quantity $\bar{z} = 1 - \frac{1}{N} \sum_j r_j$, corresponding to one minus the average total intermediate demand. We also show the upstreamness values for each sector in each country across the entire period (red full circles) constituting in total $\sim 23k$ data points - 35 sectors × 39 countries × 17 years. At the sector level, we observe a similar good agreement of the empirical exact upstreamness with the approximate values. There are occasional deviation, whose origin can be traced back to a higher degree of heterogeneity in the A matrix with respect to the "flat" rank-1 model introduced in Eq. (21). In the following we will also analyze more closely the relation between the error – discrepancy between the actual values of upstreamness (and downstreamness) calculated via direct inversion and those obtained via our approximate formula – and the spectral properties of the empirical I-O matrix A.

In Fig. 3, we repeat a similar analysis for the *downstreamness*, comparing the values obtained via direct inversion (Eq. (14)) with the approximate values of downstreamness imposing the single or double constraint on the knowledge of row sums, or row and column sums, respectively. Also for this measure, we observe a good agreement between exact and approximate values, both at the sectors (red full circles) and at the aggregate country level (cyan squares). In the following,
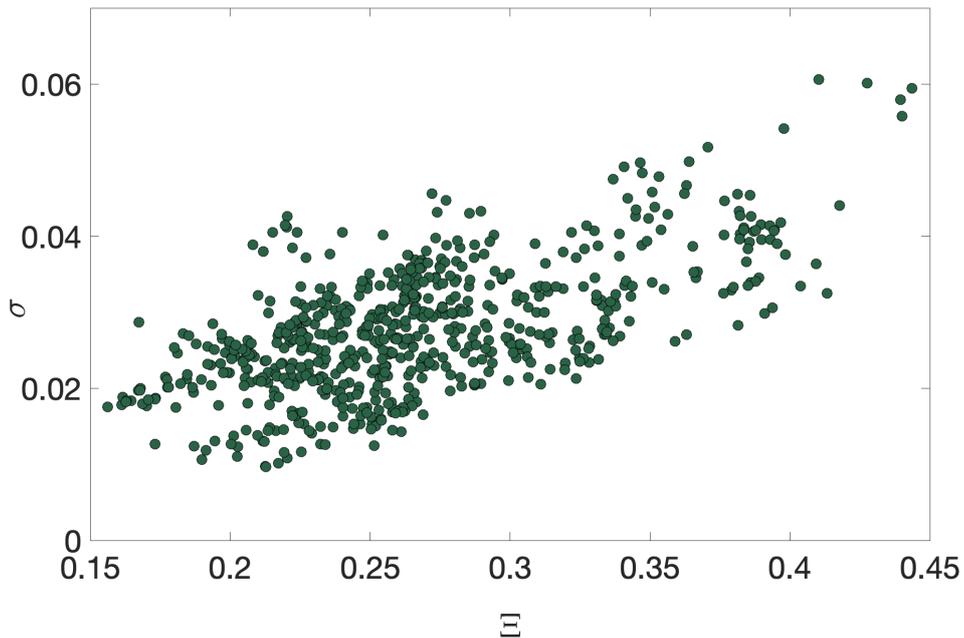

Figure 4: Error σ on approximated vs. exact upstreamness calculated for 39 countries, for the years $1995 - 2011$ year averaged over the sectors as a function of the spectral radius Ξ.

we analyze more closely the error made in the estimation of the upstreamness/downstreamness coefficients via our approximate formulae and link it to spectral properties of the underlying I-O

13

matrix A. In particular, we define the following metric for assessing the error [47]

$$\sigma = \left\langle \left| \frac{\mathcal{R}_i^{(\text{emp})}}{\mathcal{R}_i^{(\text{approx})}} - 1 \right| \right\rangle , \tag{27}$$

where $\mathcal{R}_i$ represents either the upstreamness or the downstreamness values computed via direct inversion ($\mathcal{R}_i^{(\text{emp})}$) and via our approximate formula ($\mathcal{R}_i^{(\text{approx})}$) respectively. The average $\langle \cdots \rangle$ is calculated over all sectors of a given country. Concerning the spectral properties, as shown in [47, 48] the accuracy of the approximation is related to the spectral gap of the matrix A. The matrix A has non-negative entries, therefore it has one real eigenvalue of largest magnitude λ_1 (the Perron-Frobenius eigenvalue), and its spectral gap is defined as $\Gamma = \lambda_1 - \max\{|\lambda_2|, \ldots, |\lambda_{N-1}|\}$. As the empirical I-O matrices are rather small ($N = 35$) it is more informative to look at the spectral radius. We then introduce the spectral radius excluding the Perron-Frobenius λ_1 as

$$\Xi = \max\{|\lambda_2|, \ldots, |\lambda_{N-1}|\} . \tag{28}$$

This definition is consistent with the approach used in the case of Gaussian matrices perturbed with a rank-1 matrix that may force an outlier to split off from the circular bulk [48, 50]. In Fig. 4, we display the error σ made on the approximation for all countries in all years as a function of the spectral radius Ξ of the A_U matrix characterizing each country in each year. As expected, the error grows with the spectral radius, as the rank-1 approximation becomes less accurate in reproducing the underlying intersectorial interactions. In Fig. 5, we show the same relationship labelling the countries for a single year (2011). In the bottom panel, we show the eigenvalue spectrum of two selected countries – namely China and Mexico – displaying respectively among the maximal and minimal errors in the estimation, to highlight spectral differences in the displacement of eigenvalues in the bulk. In this analysis, we find a clear negative correlation between the accuracy of the estimation and the spectral radius, i.e. the error made using our approximation increases (equivalently the accuracy of the approximation decreases) with Ξ. In general though, even in the worst cases, the relative errors remain fairly small ($\sim 5 - 6\%$) and the approximation works very well across the entire sample.
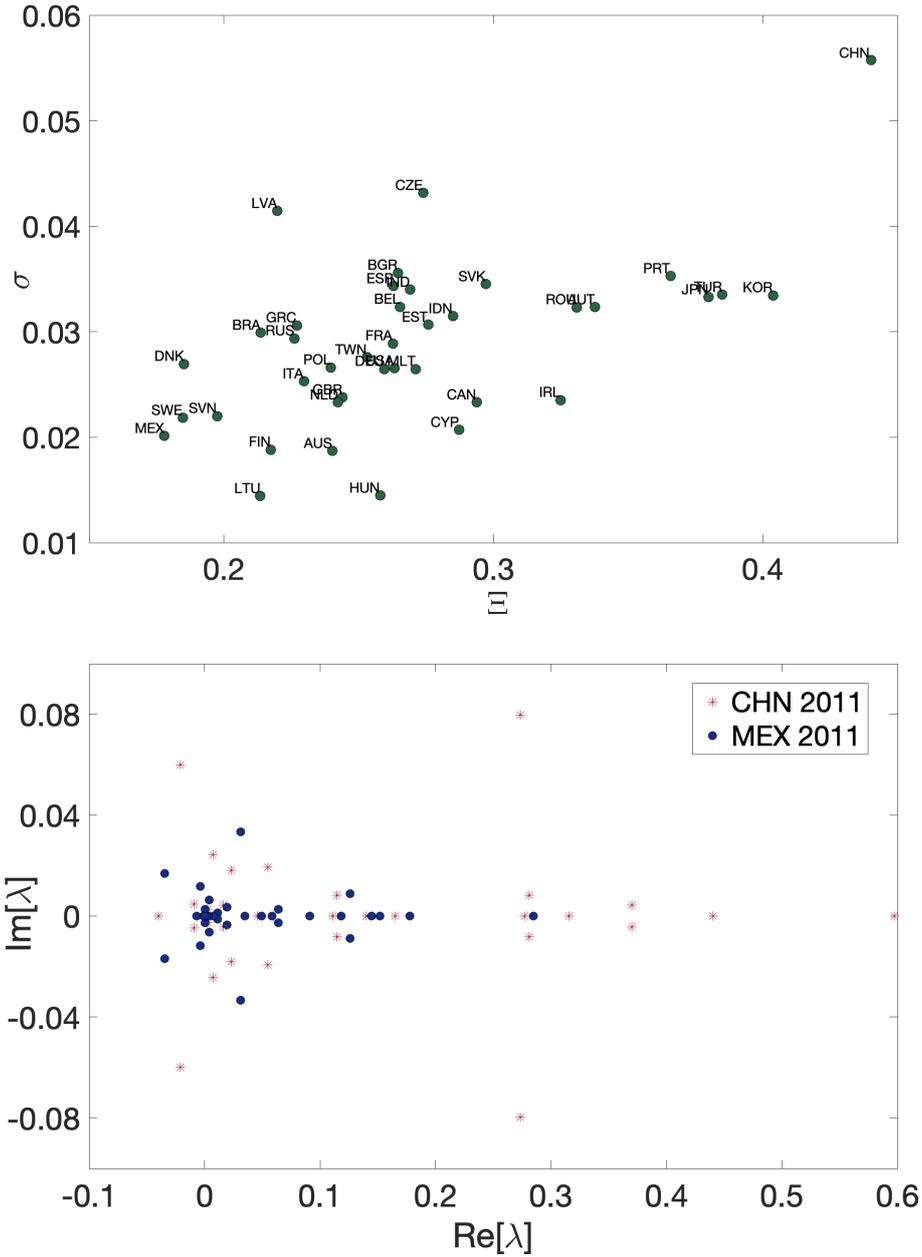
Figure 5: Top panel: Error σ approximated vs. exact upstreamness averaged over the sectors as a function of the spectral radius Ξ of the matrix A_U for all 39 countries in 2011. Bottom panel: Eigenvalue spectrum of the A_U matrix of China (CHN) and Mexico (MEX) in 2011.

6 Upstreamness under aggregation

In this section, we briefly consider how our approximation performs after the Input-Output data matrix has been subject to aggregation (consolidation) of different industrial sectors. The effects of aggregation – i.e. the procedure by which the data are looked at and lumped together at different "granularity" level – have been considered in many works (see [51] for a comprehensive

review). Here we consider the axiomatic formulation of aggregation provided in [52], which is summarized below. Furthermore, our treatment will be confined to the upstreamness, and the row-only rank-1 approximation, as generalizations to the other cases are straightforward.

Consider the definition of upstreamness given in Eq. (7)

$$\boldsymbol{U}_1 = [\mathbb{1}_N - A_U]^{-1}\mathbf{1}_N \ . \tag{29}$$

To make contact with Ref. [52], we rewrite (29) as

$$[\boldsymbol{U}_1{}^T]_N = \mathbf{1}_N^T[\mathbb{1}_N - A_U^T]^{-1} \ , \tag{30}$$

in terms of row vectors $\boldsymbol{U}_1{}^T$ and $\mathbf{1}_N^T$, and a *column*-substochastic $N \times N$ matrix A_U^T. The notation $[\ldots]_N$ indicates that the vector has length N.

Let us assume that we wish to aggregate the N "micro" industrial sectors or commodities into a set of $M < N$ "macro" sectors or commodities. Formally, we can define two matrices, S and T, of size $M \times N$ and $N \times M$ respectively. The $\{0,1\}$ matrix S indicates which micro-sectors should be combined together: $S_{ij} = 1$ if micro-sector j is to be included in macro-sector i. Thus, S is a column stochastic matrix with exactly one 1 in every column, and at least one 1 in every row. The matrix T indicates the proportional weights of each micro-sector within its macro-aggregate. The element $T_{ji} \in (0,1)$ represents the weight w_{ji} that micro-sector j carries within macro-sector i, and therefore is such that $\sum_j T_{ji} = 1$. It follows that T is also column stochastic.

Forming the aggregate $M \times M$ matrix $A_U' = SA_U^TT$ is the most common way used in the literature to create a smaller sub-stochastic matrix from the original matrix A_U, which retains (at a coarser level of detail) some of the information about industrial sectors and commodities provided by A_U. Although other choices of aggregation are possible, it was proven in [52] that the aggregator A_U' is the only one that satisfies three natural axioms of *linearity*, *value added neutrality*, and *partitioning*, therefore in the following we will confine ourselves to this case (the so called *standard aggregator*). It follows from the definition of S and T that $ST = \mathbb{1}_M$ and TS is a column stochastic, idempotent matrix of rank M (see [52] for a proof).

Although in principle any non-negative column-stochastic matrix could play the role of T, in practice it makes most sense to define it as

$$T = \mathrm{diag}(\boldsymbol{w})S^T[\mathrm{diag}(S\boldsymbol{w})]^{-1} \ , \tag{31}$$

where $\boldsymbol{w}$ is a vector of N non-negative numbers, and $\mathrm{diag}(\boldsymbol{w})$ is the diagonal matrix having the vector entries on the diagonal (in their natural order). According to Charnes and Cooper, "The main justification for this mode of consolidation is that it conforms to the way data would be synthesized ab initio if SAT rather than A were the objective" [53]. To better understand how standard aggregation works, consider as an example a 6×6 matrix A_U^T (whose elements we denote α_{ij} for simplicity, so $\alpha_{ij} = a_{ji}/Y_j$). Let

$$S = \begin{pmatrix} 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \ , \tag{32}$$

and $\boldsymbol{w} = (w_1, w_2, w_3, w_4, w_5, w_6)$. Then

$$T = \text{diag}(\boldsymbol{w})S^T[\text{diag}(S\boldsymbol{w})]^{-1} = \begin{pmatrix} 0 & \frac{w_1}{w_1+w_2} & 0 \\ 0 & \frac{w_2}{w_1+w_2} & 0 \\ \frac{w_3}{w_3+w_4} & 0 & 0 \\ \frac{w_4}{w_3+w_4} & 0 & 0 \\ 0 & 0 & \frac{w_5}{w_5+w_6} \\ 0 & 0 & \frac{w_6}{w_5+w_6} \end{pmatrix}, \tag{33}$$

and the aggregator becomes

$$A'_U = SA_U^T T = \begin{pmatrix} \frac{w_3(\alpha_{33}+\alpha_{43})+w_4(\alpha_{34}+\alpha_{44})}{w_3+w_4} & \frac{w_1(\alpha_{31}+\alpha_{41})+w_2(\alpha_{32}+\alpha_{42})}{w_1+w_2} & \frac{w_5(\alpha_{35}+\alpha_{45})+w_6(\alpha_{36}+\alpha_{46})}{w_5+w_6} \\ \frac{w_3(\alpha_{13}+\alpha_{23})+w_4(\alpha_{14}+\alpha_{24})}{w_3+w_4} & \frac{w_1(\alpha_{11}+\alpha_{21})+w_2(\alpha_{12}+\alpha_{22})}{w_1+w_2} & \frac{w_5(\alpha_{15}+\alpha_{25})+w_6(\alpha_{16}+\alpha_{26})}{w_5+w_6} \\ \frac{w_3(\alpha_{53}+\alpha_{63})+w_4(\alpha_{54}+\alpha_{64})}{w_3+w_4} & \frac{w_1(\alpha_{51}+\alpha_{61})+w_2(\alpha_{52}+\alpha_{62})}{w_1+w_2} & \frac{w_5(\alpha_{55}+\alpha_{65})+w_6(\alpha_{56}+\alpha_{66})}{w_5+w_6} \end{pmatrix}. \tag{34}$$

Now, let us assume that the vector of N upstreamness values in (30) can be faithfully approximated by our formula (25), which can be written as

$$[\hat{\boldsymbol{U}}_1{}^T]_N = \mathbf{1}_N^T + \frac{1}{1-\bar{r}_N}\boldsymbol{r}^T, \tag{35}$$

where $\boldsymbol{r}$ is the (column) vector of row sums of the matrix A_U (or the column sums of A_U^T, $r_j = \sum_{i=1}^N \alpha_{ij}$), and $\bar{r}_N$ is their average. Let us further assume that the original data matrix A_U is not known in its entirety (only its row sums are known), but the sectors/commodities in A_U have been aggregated using a *known* pair of matrices S, T – in other words, we are aware of what sectors/commodities have been lumped together (and with which relative weights) and what their aggregate outputs are, but we do not have more detailed information. We ask whether the knowledge of $\boldsymbol{r}, S$ and T is sufficient to determine $[\hat{\boldsymbol{U}}_1{}^T]_M$, namely a faithful approximation for the M upstreamness values of the aggregate model. The answer is affirmative.

First, define

$$[\boldsymbol{U}_1{}^T]_M = \mathbf{1}_M^T[\mathbb{1}_M - A'_U]^{-1} = \mathbf{1}_M^T[\mathbb{1}_M - SA_U^T T]^{-1}, \tag{36}$$

the vector of M upstreamness values, obtained using the aggregate matrix A'_U as a source. The Leontief matrix on the r.h.s. of (36) is equal to the aggregate of the Leontief matrix of the so called *companion matrix* $\bar{A}_U = A_U^T TS$ [52], namely[2]

$$[\mathbb{1}_M - SA_U^T T]^{-1} = S[\mathbb{1}_N - \bar{A}_U]^{-1}T. \tag{37}$$

Imagine now that the true matrix A_U^T appearing on the l.h.s. of (37) is replaced by its best rank-1 approximation, given by $\hat{A}^T$ (see Eq. (21)). It is easy to deduce[3] that in this case the

---

[2]The proof follows by expanding $[\mathbb{1}_M - SA_U^T T]^{-1} = \mathbb{1}_M + SA_U^T T + (SA_U^T T)^2 + \ldots$, and using $(SA_U^T T)^n = S(A_U^T TS)^n T$ and $TST = T$.

[3]This follows from the fact that the rank of the product of two matrices ($\hat{A}$ and TS) is smaller or equal than the smallest rank of the two factors, and that TS is rank-M (and of course none of the matrices involved is a null matrix).

companion matrix will also be rank-1. Applying Sherman-Morrison on the r.h.s. of (37), we get

$$S[\mathbb{1}_N - \hat{A}TS]^{-1}T = \mathbb{1}_M + \frac{1}{1 - \phi(\boldsymbol{r}, S, T)} S(\hat{A}TS)T = \mathbb{1}_M + \frac{1}{1 - \phi(\boldsymbol{r}, S, T)} S\hat{A}T \ , \qquad (38)$$

where we used $S\mathbb{1}_N T = ST = \mathbb{1}_M$, and

$$\phi(\boldsymbol{r}, S, T) = \frac{1}{N} \sum_{i,k=1}^{N} r_i (TS)_{ik} \ . \qquad (39)$$

Eq. (38) shows how to construct a faithful rank-1 approximation for the upstreamness of the aggregate model starting from the knowledge of row sums of the original model, as well as of the matrices T and S implementing the aggregation.

7 Summary and Outlook

In this paper, we have shown that the upstreamness and downstreamness measures introduced in the context of Input-Output analysis at both the inter-sectorial and country level can be faithfully recovered from the knowledge of *aggregate* and *local* information about the I-O table. In other words, the precise determination of the elements of the input-output matrix does not matter much, as long as their distribution does not deviate significantly from the "homogeneous" (flat) model (described in Eq. (21)), and the total intermediate demand per sector is ordinarily sufficient to provide an accurate estimate of the sector's multipliers.

Our rank-1 approximation has been successfully tested on National Input-Output tables obtained from WIOD, where an excellent correlation is obtained between the empirical multipliers and the theoretical formulae (see Fig. 2 and Fig. 3). Small deviations from this remarkably robust regularity are readily attributed to stronger heterogeneity in the empirical sectorial data, which would require refinements to the (single or doubly constrained) rank-1 approximation presented here.

Indeed, sparser or more heterogeneous I-O matrices tend to have a larger spectral radius (or equivalently a smaller spectral gap), as demonstrated in Fig. 4 and Fig. 5. The quality of our rank-1 approximation is very high across the sectors and countries considered, but may be inferior for empirical matrices larger spectral radii – as more eigenvalues besides the largest (Perron-Frobenius) start to play an important role.

In Section 6, we have also shown how our rank-1 approximation is well-behaved with respect to aggregation of sectorial data: knowing what sectors/commodities are lumped together, and what their aggregate outputs are, is sufficient to determine a faithful approximation for the upstreamness values of the aggregate model, as the rank-1 nature of the approximation is preserved upon aggregation.

In a recent paper [54], we further employ the rank-1 approximation as a proxy to investigate the "puzzling" correlations observed between upstreamness and downstreamness at aggregate level [35]. More generally, our approach based on a rank-1 approximation demonstrates that local and aggregate information about I-O tables is ordinarily sufficient to determine the upstreamness and downstreamness at sectorial and country level with high accuracy, while at the same time

providing analytically tractable formulae (Eq. (14), (7)) that avoid matrix inversions altogether. As an outlook for future research, it will be interesting to test the accuracy of our formulae on firm-level data, where data availability and sparsity are greater concerns. In spite of the sparser nature of the data, we would expect our approximation to work well, as recently shown on experiments conducted on synthetic data [47].

Acknowledgements

References

[1] W. Leontief (1986). Input-output economics. Oxford University Press.

[2] W. Leontief (1936). Quantitative Input-Output Relations in the Economic System of the United States. Rev. Econ. Stat. 18, pp. 105–125.

[3] United Nations Department for Economic and Social Affairs Statistics Division (1999). Handbook of Input-Output Table Compilation and Analysis.

[4] P. Antràs and D. Chor (2012). Organizing the global value chain. Econometrica 81(6), 2127-2204.

[5] P. Antràs, D. Chor, T. Fally, and R. Hillberry (2012). Measuring the Upstreamness of Production and Trade Flows. American Economic Review: Papers & Proceedings, 102(3): 412–416.

[6] T. Fally (2012). Production Staging: Measurement and Facts (unpublished).

[7] R. E. Miller and U. Temurshoev (2017). Output upstreamness and input downstreamness of industries/countries in world production. International regional science review 40(5), 443-475.

[8] A. Bacilieri and P. Austudillo-Estevez (2023). Reconstructing firm-level input-output networks from partial information. ArXiv preprint arXiv:2304.00081.

[9] P. S. M. Kop Jansen (1994). Analysis of multipliers in stochastic input-output models. Regional Science and Urban Economics 24, 55-74.

[10] P. Kop Jansen and T. Ten Raa (1990). The choice of model in the construction of Input-Output coefficients matrices. International Economic Review 31 (1), pp. 213- 227.

[11] W. D. Evans (1954). The Effect of Structural Matrix Errors on Interindustry Relations Estimates. Econometrica 22 (4), pp. 461-480.

[12] R. E. Quandt (1958). Probabilistic errors in the Leontief system. Naval Research Logistics Quarterly 5, pp. 155-170.

[13] A. Simonovits (1975). A Note on the Underestimation and Overestimation of the Leontief Inverse. Econometrica 43, pp. 493-498.

[14] G. R. West (1986). A stochastic analysis of an input-output model. Econometrica 54 (2), 363-374.

[15] H. Kogelschatz (2007). On the Solution of Stochastic input-output-Models. Discussion Paper Series n. 447, University of Heidelberg.

[16] M. Kozicka (2019). Novel approach to stochastic Input-Output modeling. RAIRO-Oper. Res. 53, 1155–1169 https://doi.org/10.1051/ro/2018046.

[17] A. L. Sargento (2009). Introducing input-output analysis at the regional level: Basic notions and specific issues. The Regional Economics Application Laboratory (REAL) https://api.semanticscholar.org/CorpusID:158457048.

[18] J. L. Katz and R. L. Burford (1985). Shortcut formulas for output, income and employment multipliers. The Annals of Regional Science 19(2), 61-76.

[19] R. L. Burford (1977). Regional input-output multipliers without a full IO table. The Annals of Regional Science 11(3), 21-38.

[20] R. L. Drake (1976). A short-cut to estimates of regional input-output multipliers: methodology and evaluation. International Regional Science Review 1(2), 1-17.

[21] P. J. Phibbs and A. J. Holsman (1981). An evaluation of the Burford Katz short cut technique for deriving input-output multipliers. The Annals of Regional Science 15(3), 11-19.

[22] R. C. Jensen and G. J. D. Hewings (1985). Shortcut 'Input-Output' Multipliers: The Resurrection Problem (a Reply). Environment and Planning A 17(11), 1551-1552.

[23] R. C. Jensen and G. J. D. Hewings (1985). Shortcut 'input-output' multipliers: A requiem. Environment and Planning A 17(6), 747-759.

[24] R. L. Burford and J. L. Katz (1985). Shortcut 'input-output' multipliers, alive and well: Response to Jensen and Hewings. Environment and Planning A 17(11), 1541-1549.

[25] F. Cerina, Z. Zhu, A. Chessa, and M. Riccaboni (2015). World Input-Output Network. PLoS ONE 10(7): e0134025 https://doi.org/10.1371/journal.pone.0134025.

[26] J. McNerney, C. Savoie, F. Caravelli, and J. D. Farmer (2018). How production networks amplify economic growth. PNAS 119 (1) e2106031118 (2021).

[27] J. Moran and J.-P. Bouchaud (2019). May's instability in large economies. Phys. Rev. E 100, 032307.

[28] R. M. del Rio-Chanona, J. Grujić, and H. J. Jensen (2017). Trends of the World Input and Output Network of Global Trade. PLoS ONE 12(1): e0170817 https://doi.org/10.1371/journal.pone.0170817.

[29] V. M. Carvalho (2012). Input-Output networks: a survey. A report for the European Commission under the CRISIS consortium agreement https://cordis.europa.eu/docs/projects/cnect/1/288501/080/deliverables/001-CRISISD31InputOutput.pdf.

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

[31] C. Hidalgo, K. Bailey, A.-L. Barabási, and R. Hausmann (2007). The product space conditions the development of nations. Science 317, 482-487.

[32] C. Hidalgo and R. Hausmann (2009). The building blocks of economic complexity. PNAS 106, 10570-10575.

[33] A. Tacchella, M. Cristelli, G. Caldarelli, A. Gabrielli, and L. Pietronero (2012). A New Metrics for Countries Fitness and Products complexity. Sci. Reports 2, 723.

[34] G. Caldarelli, M. Cristelli, A. Gabrielli, L. Pietronero, A. Scala, and A. Tacchella (2012). A Network Analysis of Countries' Export Flows: Firm Grounds for the Building Blocks of the Economy. PLoS ONE 7(10): e47278 https://doi.org/10.1371/journal.pone.0047278.

[35] P. Antràs and D. Chor (2018). On the measurement of upstreamness and downstreamness in global value chains. Working Paper 24185 http://www.nber.org/papers/w24185

[36] L. A. López, G. Arce, and P. Osorio (2023). Foreign multinationals affiliates and countries' carbon upstreamness. How could these firms support the fulfilment of emissions reduction targets? Journal of Environmental Management 326, 116714.

[37] P. Caraiani, A. Dutescu, R. Hoinaru, and G. O. Stănilă (2020). Production network structure and the impact of the monetary policy shocks: Evidence from the OECD. Economics Letters 193, 109271.

[38] K. Suganuma (2016). Upstreamness in the global value chain: Manufacturing and services. Monetary and Economic Studies 34, p. 39-66, https://EconPapers.repec.org/RePEc:ime:imemes:v:34:y:2016:p:39-66.

[39] M. P. Timmer, E. Dietzenbacher, B. Los, R. Stehrer, and G. J. de Vries (2015). An Illustrated User Guide to the World Input–Output Database: the Case of Global Automotive Production. Review of International Economics 23, 575–605.

[40] T. Squartini, G. Caldarelli, G. Cimini, A. Gabrielli, and D. Garlaschelli (2018). Reconstruction methods for networks: The case of economic and financial systems. Physics Reports 757, 1-47.

[41] G. Cimini, T. Squartini, D. Garlaschelli, and A. Gabrielli (2015). Systemic Risk Analysis on Reconstructed Economic and Financial Networks. Scientific Reports 5, 15758 (2015).

[42] G. Bianconi (2002). Mean field solution of the Ising model on a Barabási–Albert network. Physics Letters A 303, 166-168.

[43] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes (2008). Critical phenomena in complex networks. Rev. Mod. Phys. 80, 1275 (2008).

[44] J. Park and M. E. J. Newman (2004). The statistical mechanics of networks. Phys. Rev. E 70, 066117.

[45] G. Caldarelli, A. Capocci, P. De Los Rios, and M. A. Muñoz (2002). Scale-Free Networks from Varying Vertex Intrinsic Fitness. Phys. Rev. Lett. 89, 258702.

[46] V. Thibeault, A. Allard, and P. Desrosiers (2024). The low-rank hypothesis of complex systems. Nat. Phys. https://doi.org/10.1038/s41567-023-02303-0.

[47] S. Bartolucci, F. Caccioli, F. Caravelli, and P. Vivo (2023). Ranking influential nodes in networks from aggregate local information. Phys. Rev. Research 5, 033123.

[48] Bartolucci, S., Caccioli, F., Caravelli, F., & Vivo, P. (2021). "Spectrally gapped" random walks on networks: a Mean First Passage Time formula. SciPost Physics, 11(5), 088.

[49] J. Sherman, W. J. Morrison, Adjustment of an Inverse Matrix Corresponding to a Change in One Element of a Given Matrix. *Annals of Mathematical Statistics* **21** (1), 124-127 (1950).

[50] F. Mosam, D. Vidaurre, and E. De Giuli (2021). Breakdown of random matrix universality in Markov models. Physical Review E 104(2), 024305.

[51] K. O. Kymn (1990). Aggregation in input–output models: a comprehensive review, 1946–71. Economic Systems Research 2(1), 65-93.

[52] E. C. Howe and C. R. Johnson (1989). Linear Aggregation of Input-Output Models. SIAM J. Matrix Anal. Appl. 10 (1), 65-79.

[53] A. Charnes and W. W. Cooper (1961). Management Models and Industrial Applications of Linear Programming. Vol I, John Wiley, New York.

[54] S. Bartolucci, F. Caccioli, F. Caravelli, and P. Vivo (2023). Correlation between upstreamness and downstreamness in random global value chains. arXiv preprint arXiv:2303.06603.

Systemic Risk Centre