The Oxford Martin Programme on

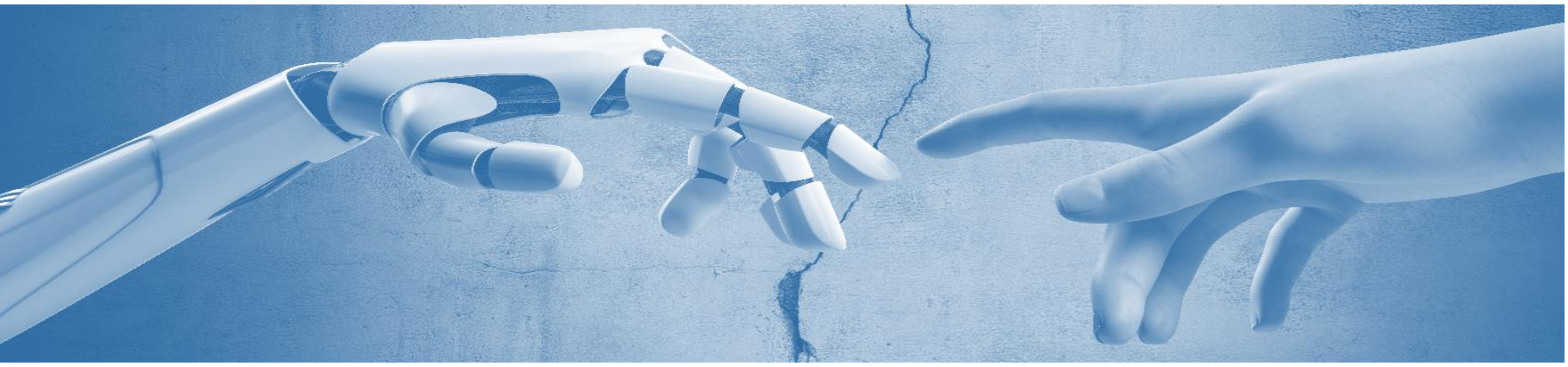
Technological and Economic Change

The Persistent Statistical Structure of the US Input-Output Coefficient Matrices: 1963-2007

Luis Daniel Torres Gonzalez and Jangho Yang

April 2018

OMPTEC Working Paper No. 2018-3



The Persistent Statistical Structure of the US Input-Output Coefficient Matrices: 1963-2007

Luis Daniel Torres Gonzalez* and Jangho Yang† ‡

Abstract

The paper finds evidence for the existence of a statistical structure in the US input-output (I-O) coefficient matrices for 1963-2007 and characterizes the identified statistical regularities. For various aspects of **A** matrices we find smooth and unimodal empirical distributions (ED) with a remarkable stability in their functional form for most of the samples. The ED of all entries, diagonal entries, row sums, and the entries of the (left- and right-hand) Perron-Frobenius eige nvectors are well described by fat-tailed distributions while the ED of column sums and eigenvalues' moduli are explained by the Normal and the Beta distribution, respectively. The paper provides several economic interpretations of these statistical results as well as some implications and potential uses for the structural and stochastic I-O analysis.

Keywords: Statistical Structure, Stochastic Input-Output Analysis, Structural Analysis, Structural Change, US Input-Output Accounts, Bayesian estimation.
JEL Class: C11, C46, C51, C52, C67, D57, O51.

*Department of Economics, New School for Social Research, 6 East 16th Street, New York, NY, USA 10011-8872, +52 1 5529208389, torrl352@newschool.edu.

†Institute for New Economic Thinking at the Oxford Martin School, University of Oxford, Oxford OX2 6ED, UK, +44 7432 379960, jangho.yang@oxfordmartin.ox.ac.uk

‡Oxford Martin Programme on Technological and Economic Change, Oxford Martin School, University of Oxford, Oxford, Oxford, OX1 3BD

1 Introduction

In his groundbreaking book *The Structure of the American Economy*, Leontief (1951) is interested in studying the *structure* of the economy, that is, the study of the general interdependence of the different elements of the economic system. To do so, Leontief develops the modern Input-Output (I-O) model and constructs the statistical measurements needed for its application.

The fundamental element of the I-O model is the I-O flow matrix $\underset{n\times n}{\mathbf{Z}} = \{z_{ij}\}$: the matrix that records the inter-industry money flows and tracks the intermediate commodity inputs which a given industry needs from other industries for the production of its own output. In the money flow z_{ij}, the i-th industry sells an amount of output valuated at z_{ij} to the j-th industry. From the perspective of the i-th industry, this sale represents production for intermediate consumption, whereas to the j-th industry, the amount z_{ij} represents the purchase of intermediate commodities as inputs for its production process. The output of all industries is written as the vector $\underset{1\times n}{\mathbf{x}} = \{x_j\}$, which can be represented as a diagonal matrix $\underset{n\times n}{\hat{\mathbf{x}}} = \text{diag}\{x_1, \dots, x_n\}$. Assuming a technology with fixed proportions between inputs i and outputs j, that is $z_{ij} = a_{ij} \cdot x_j$, we can construct the I-O coefficient matrix $\underset{n\times n}{\mathbf{A}} = \mathbf{Z}\hat{\mathbf{x}}^{-1} = \left\{a_{ij}\right\}$, where each a_{ij} represents the value of i-th industry's production required to produce one unit of value of industry j's output.

When $\mathbf{Z}$ and $\mathbf{A}$ are constructed from actual inter-industry data, they provide the *description* of the complex network of industries' interrelations (Leontief, 1986); therefore, matrix $\mathbf{A}$ and its extensions, such as Leontief's inverse $\underset{n\times n}{\mathbf{L}} = (\mathbf{I} - \mathbf{A})^{-1} = \{L_{ij}\}$, have been used as a centerpiece of various multi-sectoral models to study prices, production, and employment, among others (see Pasinetti, 1977, Miller and Blair, 2009, and ten Raa, 2005 for an extensive exposition of the applications of I-O modeling and linear multisectoral models). The importance of $\mathbf{Z}$ and $\mathbf{A}$ matrices has attracted attention of researchers in I-O analysis. Some examples of methods developed and applied by *structural analysis* include: sensitivity analysis, fundamental economic structure, fields of influence, qualitative I-O techniques, network I-O analysis, and structural decomposition analysis (see Jensen et al., 1988, 1991; Tarancón, 2003; and Miller and Blair, 2009).

Despite the substantial interest in the study of structural properties of I-O matrices, more focus is needed on *statistical analysis* of observed $\mathbf{Z}$ and $\mathbf{A}$ matrices and on the identification of *statistical regularities*. This gap in the literature is rather puzzling considering the importance of 'stylized facts' of empirical evidence in the development of proper theories. There are some reasons for this gap. Probabilistic and statistical analysis of I-O matrices in general have only been considered marginally in the field (West, 1986). Most of the I-O studies that do adopt statistical and probabilistic approaches only pertain to the precision of I-O modeling and the estimation/update of the I-O entries (see Jackson, 1989; Jackson and West, 1989; Temurshoev, 2017). In these studies, the probabilistic structures of entries z_{ij} and a_{ij} are often assumed to emerge from 'measurement errors' that involve data collection, confidentiality, reporting, and sampling, among others (ten Raa, 2005; Rueda et al., 2013, Temurshoev, 2017). This understanding of the probabilistic structure in the data —solely as a result of human errors— often leads to the normality assumption based on a sampling theory and discourages a more empirically grounded characterization of the observed regularities that might come from the actual economic processes.

This paper addresses the statistical characteristics of 10 US' **A** matrices for 1963-2007 (a period of significant technological change) and establishes, for the first time, stylized facts for a series of variables which represent different aspects of the matrices. First, we evaluate if *there is* a persistent statistical structure in **A** matrices. Next, we *characterize* the identified structure for each variable based on several representative probabilistic models. From the models, we provide a preliminary economic interpretation of these statistical regularities. Finally, we explore some implications and potential uses of our results for existing I-O analyses.

The methodology taken in this study is consistent with the holistic approach advocated in Jensen et al. (1988) and Jensen et al. (1991).[1] This is because we understand any subset of the $n \times n$ cells in relation to the rest of the elements of a matrix by using an empirically grounded probabilistic model and try to connect the observed regularities to some fundamental economic processes that govern the interdependence of production in an economy..

For every considered variable we find smooth and unimodal empirical distributions (ED). Moreover, we find that the functional form of the ED for all the variables is persistent over extended periods of time. As an example of our results, we show that the distribution of column sums of an **A** matrix is well explained by a symmetric distribution such as the Normal distribution, while the distribution of row sums is explained by a rapidly decaying distribution with a fat tail such as Log-Normal and Weibull distributions. Since the column sum represents the ratio of the intermediate commodities inputs to output (whose inverse is the productivity of intermediate commodities), the symmetric and unimodal pattern of its distribution would require some theoretical model to explain why there is a cross-sectional convergence of productivities. In contrast, the row sums of an **A** matrix represent the importance of industry i's product relative to the production of the rest of the industries. The fact that we observe a fat-tailed pattern in the row sum distribution calls for an explanation of the endogenous concentration of supply chains on a small number of industries.

Section 2 reviews the statistical regularities of actual **A** matrices already identified in the literature. Following that, (in Section 3) empirical evidence is presented on the existence of a statistical structure in the observed **A** matrices and statistical results on its characterization. We study 10 **A** matrices covering a 44-year period from 1963-2007 and with the number of sectors ranging from 351 to 478. In this section we construct the ED of the following elements of an **A** matrix: the pooled and diagonal elements, the column and row sums, the absolute value of the eigenvalues, and the elements of the Perron-Frobenius (PF) eigenvectors —the left- and right-hand eigenvectors of an **A** matrix associated

[1] 'The holistic approach views the cells of the input-output table not simply as observations of individual categories of economic activity, but as elements of an entity, which collectively present a "portrait" of an economy in terms of general structural characteristics. This view suggests new approaches to analysis and description of economic structure, both in terms of whole-economy concepts and in terms of the contribution of parts to the whole as a micro-macro interface.' (Jensen et al., 1991, p. 229)

with the PF eigenvalue (the maximum eigenvalue). Next, we fit probabilistic models to the obseved ED of target variables to understand the statistical structure of the matrices. Last, we provide a preliminary economic interpretation of the results. In Section 4 we provide some implications and potential uses of our results for the existing I-O analysis. Section 5 concludes the paper by summarizing the results and suggesting future lines of research.

2 A Review of the Identified Statistical Regularities of I-O Matrices

Three major lines of research on probabilistic and statistical aspects of I-O matrices and tables were identified: 1) the stochastic I-O analysis literature, 2) the Bródy's Conjecture debate, and 3) recent prices of production models. The objective of this section is not to give an exhaustive review of all their probabilistic and statistical exercises, but instead to focus on the statistical regularities found in their studies.

2.1 The Stochastic I-O Analysis

The most substantial literature involving random aspects in multisectoral models is the 'stochastic I-O analysis,' whose key motivation is to study the precision of I-O modeling and to estimate/update the I-O entries. In most studies, the justification for the stochastic approach to I-O modeling comes from the potential *measurement errors* generated while constructing $\mathbf{Z}$ matrices or the commodity-by-industry Use $\mathbf{U} = \{u_{kj}\}$ and Make $\mathbf{V} = \{v_{jk}\}$ tables, for $k = 1, \ldots, m$, where m is the number of commodities. The sources of randomness in entries z_{ij}, u_{ij}, and v_{ij} are assumed to be errors originating from data collection, confidentiality, reporting, and sampling among other things (ten Raa, 2005; Rueda et al., 2013, Temurshoev, 2017).[2] According to West (1986), Jackson (1989), and Rueda et al. (2013), the objective of the stochastic I-O analysis is to study:

- Biases in Leontief's inverse matrix $\mathbf{L}$ and different multipliers.
- Moments, confidence intervals, and density functions in $\mathbf{L}$.
- Moments and confidence intervals for a_{ij}.
- Statistical estimation of a_{ij} and the multipliers.

For these purposes, various assumptions about the stochastic structure of the matrices are made primarily based on sampling theory (e.g. the normality assumption, the symmetric nature of variance, the (in)dependence of the entries, etc). Nevertheless, these studies do not explicitly attempt to find the statistical regularities of the I-O matrices. As of this writing, Jackson (1986) and Wibe (1982) are the best known studies that look into some statistical characteristics of the I-O coefficients; however, these authors focus on the input and output relations of the firms and establishments within a small set of

[2]Only a few studies, such as Jackson (1986, 1989) and Jackson and West (1989), consider that randomness in the I-O flows or coefficient entries could come from sources related to economic factors.

industries. More fundamentally, some of the assumptions conventionally made in the field of stochastic I-O analysis, especially the normality assumption of the coefficient a_{ij}, are not empirically grounded.

2.2 The Debate on Bródy's Conjecture

By choosing a probabilistic model to define **A** and estimating the dominant λ_1 and subdominant eigenvalue λ_2, Bródy (1997) puts forward the hypothesis that the spectral ratio $\frac{|\lambda_2|}{\lambda_1}$ will tend to zero as the size n of **A** increases. In this probabilistic model, each $a_{ij} \sim i.i.d.\left(\mu, \sigma^2\right)$.[3] This conjecture, which was later known as Bródy's Conjecture, implies that economic systems with large **A** matrices, like the ones representing everyday market economies, require fewer iterations to reach equilibrium output proportions —the proportions given by the right-hand PF eigenvector $\underset{n\times 1}{\mathbf{q}_1^{\mathrm{R}}} = \{q_{i,1}^{\mathrm{R}}\}$.

Bródy arrives at this result purely based on algebraic and probabilistic formalism —he neither uses statistical properties of actual **A** matrices to corroborate his probabilistic assumptions nor studies the empirical eigenvalues of these matrices. Bidard and Schatteman (2001) provides a proof of Bródy's conjecture while Sun (2008) and Schefold (2013) introduce theorems by Goldberg et al. (2000) and Goldberg and Neumann (2003) —from the mathematical random matrix literature— to further develop Bródy's Conjecture. None of these works, however, try to determine if the statistical characteristics of the theoretical a_{ij} and λ_t are consistent with the observed ones.

Bródy's Conjecture has also spurred a number of empirical works whose purpose is to empirically assess the conjecture. The path taken by these empirical studies computes the empirical spectral ratio at different levels of aggregation. Some statistical regularities been have successfully identified in the eigenvalues of observed **A** and $\mathbf{J} = \frac{1-\lambda_1}{\lambda_1}\mathbf{A}(\mathbf{I}-\mathbf{A})^{-1}$ matrices that are persistent for all the countries, years, and aggregation levels considered in their studies.[4] Their main findings show that:[5]

1. When looking at the complex plane, eigenvalues are clustered around zero.
2. Only a small proportion of subdominant eigenvalues have a magnitude different from zero, but this number is important —i.e. we cannot say that all subdominant eigenvalues are close to 0.
3. The absolute value of the eigenvalues exhibits a remarkable and persistent organization, which seem to follow a pattern of exponential decay —There is no evidence of a degenerate distribution of the subdominant eigenvalues around 0.

[3]Bródy proposes the following probabilistic model: Each column of **A** is a sample of $\{x\}$ of size n, where x is a positive and continuous random variable with unknown probability distribution $\{x\}$ but with a given mean μ and variance σ^2. Bródy implicitly assumes that elements within each sample are independent, hence each input coefficient $a_{ij} \sim i.i.d.\left(\mu, \sigma^2\right)$.

[4]See Mariolis and Tsoulfidis (2011, 2016b), Nassif and Shaikh (2015), and Schefold (2013).

[5]Some authors have been interested in the behavior of the spectral ratio and the corresponding eigenvalue distribution when changing the aggregation level. Mariolis and Tsoulfidis (2014), Gurgul and Wojtowicz (2015), and Nassif and Shaikh (2015) show that the spectral ratio $\frac{|\lambda_2|}{\lambda_1}$ tends to increase as the disaggregation level increases. In addition, the results from Nassif and Shaikh (2015) indicate that, as the level of disaggregation increases, 1) the number of eigenvalues with a considerable magnitude increases, 2) the magnitude of these relevant eigenvalues increases, and 3) the normalized rank-plot of eigenvalues' moduli tends to be more 'L-shaped.'

2.3 Recent Prices of Production Models

Empirical computations of prices of production models for different countries and years have consistently produced nearly linear price and wage curves, i.e. production prices and the wage rate (or wage share) as a function of the rate of profit.[6] This is considered a puzzle in the literature because the constraints on the exogenous variables in these models are not strong enough to persistently produce nearly linear price and wage curves. These constraints are the positivity ($a_{ij} \geq 0$) and productivity ($\lambda < 1$) of **A**, and the positivity ($l_j, x_j > 0$) of the labor coefficient vector $\underset{1\times n}{\mathbf{l}} = \{l_j\}$ and gross output vector **x**. In order to explain the observed near-linearity in the price and wage curves, recent works in the literature have advanced some hypotheses based on the statistical regularities in the structure of technology and demand. This structure consists of (Torres, 2018):[7]

(1) the properties of **A** and
(2) the relationship that **A** has with **l** and $\mathbf{x}^{\mathrm{T}}$

In the study of (1), Iliadi et al. (2014), Mariolis and Tsoulfidis (2011, 2016a, 2018), and Torres (2018, 2017) find the same characteristics in the eigenvalues as the ones in the Bródy's Conjecture literature. Torres (2018; 2017) finds in addition that its column sums $\sum_{i=1}^{n} a_{ij}$ display a smooth, unimodal, and highly symmetric ED. Regarding (2), Torres (2018) finds evidence of a statistical tendency toward the proportionality ($\propto$) between **l** and $\mathbf{x}^{\mathrm{T}}$ and the PF eigenvectors of **A**: $\mathbf{l} \propto \mathbf{q}_1^{\mathrm{L}}$ and $\mathbf{x}^{\mathrm{T}} \propto \mathbf{q}_1^{\mathrm{R}}$.[8] However, these regularities in the column sums and the proportionality of the vectors are only descriptive, without any attempt to characterize the observed patterns.

3 A Study of the Statistical Structure of the US I-O coefficients matrices

The regularities in the statistical behavior of the eigenvalues, column sums, and the relationship of the labor and output vectors with the PF eigenvectors summarized in the previous section indicate that **A** matrices might have persistent structural properties of a statistical nature. The objective of this section is to present empirical evidence for the *existence* of a persistent structure in **A** matrices and to *characterize* the identified statistical regularities.

To do so, we present statistical information for different variables involving some of or all the coefficients a_{ij}. These variables represent relevant economic and mathematical aspects of **A** matrices

[6] Prices of production are theoretical prices that equalize the rates of profits taken as given the technology and the wage rate (or the profit rate). See Pasinetti (1977, 1988) for an exposition of this model and Shaikh (2016) and Mariolis and Tsoulfidis (2016b) for a modern review of this literature.

[7] Iliadi et al. (2014) and Mariolis and Tsoulfidis (2011, 2016a, 2018) consider matrix **J** whereas Schefold (2013) and Torres (2017, 2018) considers matrix **A**. Both matrices share the same eigenvectors.

[8] The deviations between $l_j - l_1^* q_{1,j}^{\mathrm{L}}$ and $x_i - x_1^* q_{i,1}^{\mathrm{R}}$, have a smooth, unimodal, highly peaked, and symmetric ED, centered around 0. When **A** is diagonalizable, then it has n linearly independent (left- and right-hand) eigenvectors, so vectors **l** and $\mathbf{x}^{\mathrm{T}}$ can be represented as $\mathbf{l} = l_1^* \mathbf{q}_1^{\mathrm{L}} + \ldots + l_n^* \mathbf{q}_n^{\mathrm{L}}$ and $\mathbf{x}^{\mathrm{T}} = x_1^* \mathbf{q}_1^{\mathrm{R}} + \ldots + x_n^* \mathbf{q}_n^{\mathrm{R}}$, where the l_t^* and x_t^* are the coordinates of **l** and $\mathbf{x}^{\mathrm{T}}$ in the vector space given by the left- and right-hand eigenvectors, respectively. Hence, if $\mathbf{l} \propto \mathbf{q}_1^{\mathrm{L}}$ and $\mathbf{x}^{\mathrm{T}} \propto \mathbf{q}_1^{\mathrm{R}}$, then $\mathbf{0} = \mathbf{l} - l_1^* \mathbf{q}_1^{\mathrm{L}} = \sum_{t=2}^{n} l_t^* \mathbf{q}_t^{\mathrm{L}}$ and $\mathbf{0} = \mathbf{x}^{\mathrm{T}} - x_1^* \mathbf{q}_1^{\mathrm{R}} = \sum_{t=2}^{n} x_t^* \mathbf{q}_t^{\mathrm{R}}$ (Torres, 2018).

from different points of view. Based on the US benchmark I-O accounts (BIOA), compiled by the Bureau of Economic Analysis (BEA), 10 different **A** matrices are constructed, covering 44 years from 1963-2007, and with a level of detail ranging between 351 to 478 industries. All the matrices refer to totals (domestic plus imported figures). Non-negative, indecomposable **A** matrices are contructed with non-fictitious industries and with industries containing positive wage bill and value-added.[9]

3.1 The Set of Variables to Study

Seven variables that show important aspects of **A** matrices are investigated. The first four variables are related to the configurations of the coefficients a_{ij}, whereas the remaining three are related to the eigenvalues and eigenvectors. The following variables are considered:

1. *Pooled* a_{ij}. Leontief (1986, Ch.1,2,8) considers that there is a fundamental relationship between outputs x_j and inputs z_{ij}, measured by the a_{ij} coefficients, $0 \leq a_{ij} < 1$, , which is 'relatively invariable' and 'inflexible.' This input-output relationship is determined mainly by the technology —customs and institutional arrangements are complementary determinants of a_{ij}. Hence, the technological structure of an economy, regarding the intermediate commodity inputs, is represented by **A**.[10] Due to the size of the constructed **A** matrix, there are usually a great number of zero entries. For example, in the **A** matrix for the US in 1967, 88% of the observed entries are zero.[11] Since the existence of zero values is highly affected by the construction procedure, which has gone through substantial changes for a few times in the past decades, this study will use the positive entries only. As will be shown in the following sections, the distribution of pooled a_{ij} is extremely peaked close to zero with a heavy tail, even after removing zero entries.

2. *Diagonal entries* a_{ii}. Studies on multisectoral models take an interest in dominant diagonal matrices, i.e. matrices for which positive scalars $d_1, ..., d_n$ exist such that $d_j|a_{ii}| > \sum_{i \neq j} d_j|a_{ij}|$, for $j = 1, ..., n$ (Takayama 1974, Ch. 4.C). In addition, they have a mathematical connection with some aspects of the matrices, such as the trace and the sum of eigenvalues, $\mathrm{Tr}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii} = \sum_{t=1}^{n} \lambda_t$.

[9]For years 1963 to 1967, BEA publishes the industry-by-industry matrix **Z**. From 1972 onwards, the BIOA are constructed from a commodity-by-industry perspective and they are gathered in the Use **U** and the Make **V** tables. We use tables **U** and **V** after redefinitions (see BEA, 2009). Following Miller and Blair (2009), the industry-by-industry **A** matrices are constructed under the industry-based technology assumption. There were some minimal manipulations of matrices and tables **Z**, **U**, **V**, and **A**: Industries with zero intermediate commodity inputs and negative value-added or wage bill were eliminated. Fictitious industries were eliminated as well (like Commodity Credit Corporation; directly, transferred, and noncomparable imports; scrap, used and secondhand goods; rest of the world industry; inventory valuation adjustment; household industry). A small number of negative entries were replaced with zero in the use tables. Finally, non-basic industries were eliminated to obtain an indecomposable matrix, which ensures that $\mathbf{q}_1^{\mathrm{L}} > \mathbf{0}$ and $\mathbf{q}_1^{\mathrm{R}} > \mathbf{0}$. The years (number of industries) of the database are: 1963 (351), 1967 (438), 1972 (450), 1977 (478), 1982 (472), 1987 (456), 1992 (466), 1997 (461), 2002 (406), and 2007 (371).

[10]The labor coefficient vector **l** and the matrix of capital stock coefficients **B** complement the picture. We will not consider these aspects of technology in this paper.

[11]The following is the proportion of zero entries in each matrix.1963: 74.4%, 1967: 88.2%, 1972: 38.8%, 1977: 30.9%, 1982: 20.6%, 1987: 27.2%, 1992: 31.8%, 1997: 32.6%, 2002: 24.7%, 2007: 21.3%

3. *Column sums* $\kappa_j = \sum_{i=1}^{n} a_{ij}$. These sums are related to the intermediate commodity inputs per dollar of output and value-added per dollar of output for each industry va_j, $\kappa_j = 1 - va_j$. In linear price and growth models with circulating capital, $\kappa_j = \frac{1}{1+u_j}$ represents industries' capital output ratios, which are directly related to u_j, the capital-value added ratios, i.e. the inverse of capital productivities. Hence, regularities in industries' column sums indicate an organization on industries' capital productivities and the value-added per dollar of production.

4. *Row sums* $\rho_i = \sum_{j=1}^{n} a_{ij}$. Each a_{ij} represents the proportion of the i-th input outlay on the value of production of industry j. By taking the row sums of **A**, the comparison of the different ρ_i indicates the relative importance of each input in the productive structure of an economy: A high ρ_i implies that industry i's production is an indispensable input in the production network.[12]

5. *Eigenvalues' absolute value* $|\lambda_t|$. Studying the eigenvalues λ_t of **A** can help us to identify hidden economic and mathematical properties of the matrix that otherwise would be difficult to see. To show this, we assume **A** has n-distinct eigenvalues $\lambda_1, ..., \lambda_n$;[13] then, **A** can be represented as (Meyer, 2000, pp. 517-8):

$$\mathbf{A} = \frac{\lambda_1}{\mathbf{q}_1^{\mathrm{L}}\mathbf{q}_1^{\mathrm{R}}}\mathbf{q}_1^{\mathrm{R}}\mathbf{q}_1^{\mathrm{L}} + ... + \frac{\lambda_n}{\mathbf{q}_n^{\mathrm{L}}\mathbf{q}_n^{\mathrm{R}}}\mathbf{q}_n^{\mathrm{R}}\mathbf{q}_n^{\mathrm{L}} \quad (1)$$

$$= \lambda_1\mathbf{A}_1 + ... + \lambda_n\mathbf{A}_n, \quad (2)$$

where $\mathbf{A}_t = \frac{1}{\mathbf{q}_t^{\mathrm{L}}\mathbf{q}_t^{\mathrm{R}}}\mathbf{q}_t^{\mathrm{R}}\mathbf{q}_t^{\mathrm{L}}$ and $\mathbf{I} = \mathbf{A}_1 + ... + \mathbf{A}_n$. That is, the actual **A** matrix is a linear combination of rank-one $\mathbf{A}_t$ matrices; therefore, the observed strong concentration of λ_t around zero reported in Section 2, suggests that if λ_t are ordered by their absolute value, then $\mathbf{A} \approx \sum_{t=1}^{c} \lambda_t\mathbf{A}_t$, for a small integer $c < n$.[14] In multisectoral price and growth models, the PF eigenvalue λ_1 is directly associated with the maximum rates of return and growth an economy can achieve.[15] More generally, eigenvalue λ_s is the capital-gross output ratio of the economy when the vector of labor coefficients **l** is proportional to eigenvector $\mathbf{q}_s^{\mathrm{L}}$, $\mathbf{l} \propto \mathbf{q}_s^{\mathrm{L}}$, and/or when the output vector $\mathbf{x}^{\mathrm{T}}$ is proportional to eigenvector $\mathbf{q}_s^{\mathrm{R}}$, $\mathbf{x}^{\mathrm{T}} \propto \mathbf{q}_s^{\mathrm{R}}$.[16] Given that eigenvalues belong to the complex field, one way to represent their statistical information is to take their magnitude or absolute value: $|\lambda_t|$.

6. *Entries of the Perron-Frobenius eigenvectors* $q_{1,j}^{\mathrm{L}}$ and $q_{i,1}^{\mathrm{R}}$. In the same models where λ_1 is related

[12]From a mathematical point of view, both column and row sums are popular matrix norms and, according to Gerschgorin's circles, they could constitute relevant constraints for the eigenvalues. See chapter 5.2 and example 7.1.4 in Meyer (2000).

[13]'Any defective square matrix A [a matrix with repeated eigenvalues] can be made to have all distinct eigenvalues by choosing an appropriate perturbation direction F for a sufficiently small $\epsilon \neq 0$' (Aruka, 2011, p. 46)

[14]See Schefold (2013) and Mariolis and Tsoulfidis (2016b).

[15]See Pasinetti (1977, Ch. 5,7) and Takayama (1974, Ch. 6,7) for a detailed discussion.

[16]Goodwin (1974; 1976) introduces the notion of Normalized General Coordinates, that is, the change of basis from the commodity space from matrix **A** to the space given by the eigenvectors, which represent composite commodities of the original system in **A**.

to the maximum profit and growth rates, eigenvectors $\mathbf{q}_1^{\mathrm{L}}$ and $\mathbf{q}_1^{\mathrm{R}}$ correspond to the proportions of the corresponding price and output vectors. These vectors represent the direction of a self-expansion of the system. On the other hand, $\mathbf{q}_1^{\mathrm{L}}$ and $\mathbf{q}_1^{\mathrm{R}}$ define $\mathbf{A}_1$ —the first layer of the matrix $\mathbf{A}$ in Equation (2). The matrix $\lambda_1 \mathbf{A}_1$ could be seen as the core of the structure of technology, as is in the tiered approach to the fundamental economic structure in Jensen et al. (1991). Finally, the empirical evidence on the statistical tendency towards the proportionality between $\mathbf{l} \propto \mathbf{q}_1^{\mathrm{L}}$ and $\mathbf{x}^{\mathrm{T}} \propto \mathbf{q}_1^{\mathrm{R}}$ reviewed in Section 2 makes the study of the statistical properties of the PF eigenvectors relevant.

3.2 Descriptive results

Figure 1 shows the heat map of $\mathbf{A}$ matrices to visualize their general patterns. This plot presents the above 95% quantile components as a black cell and the remaining components as a white cell, effectively displaying the most salient pattern in the matrix. The result shows that almost all diagonal entries in each matrix contain large technical coefficients, which indicates that intra-industry trade is an important component. Also, large values are distributed disproportionally among rows. This means that the inputs from some industries are substantially more important than others in the production network. These patterns are persistent for the whole sample.

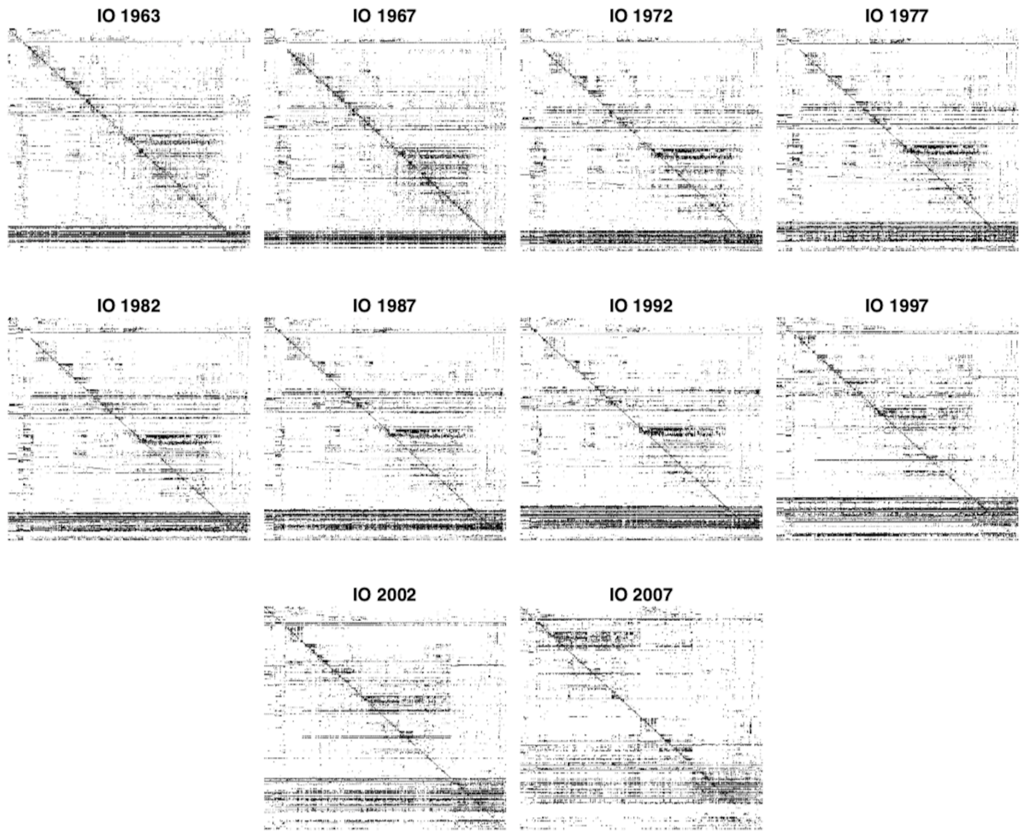


Figure 1: Heat map of the US input coefficient matrix, $\mathbf{A} = \{a_{ij}\}$*: 1963-2007, between 351-478 industries. The black cells represent the above 95% quantile components while the white cells represent the remaining ones. Visual inspection reveals two persistent regularities: the diagonal elements and certain rows have a prominent role in the signal of* **A**.

Figures 2 through 8 show the empirical distribution (ED), cumulative distribution function (CDF), and complementary cumulative distribution function (CCDF) for each variable and year. The plots for all the variables show smooth, unimodal, and highly peaked ED, whose shapes are remarkably constant for almost all the samples.[17]

The relevance of this time invariance in the ED is remarkable given the substantial technological and institutional changes of the US economy in the past decades.[18] Borrowing the terminology from statistical mechanics, each **A** matrix represents a particular *micro-state* of the interindustry structure, that is, a microscopic or detailed account of the value of each one of coefficients a_{ij}. Predictably, a numerous microstates are possible depending on different technological and institutional settings. Indeed, Foerster and Choi (2017), Gowdy (1991), and Verspagen (2004) report on the structural changes in the interindustry relationships of the US economy in the post-World War II period by tracking the

[17]The same general patterns are present for years 1947 and 1958. However, due to the smaller number of observations, the EDs are somewhat noisy.

[18]It is worthwhile to note the persistence of these regular patterns in **A** matrices despite of major changes in the methodology for the construction of the BIOA experienced since 1972.

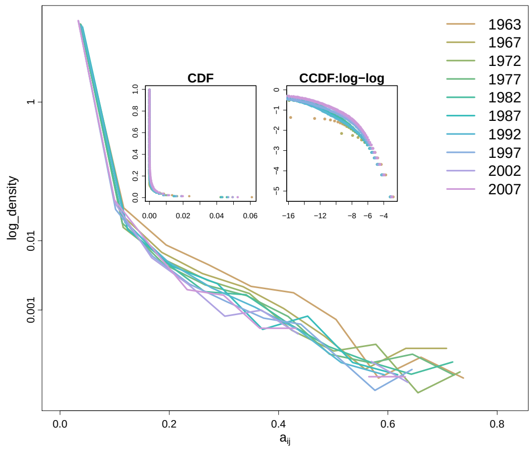


Figure 2: Pooled coefficients a_{ij}

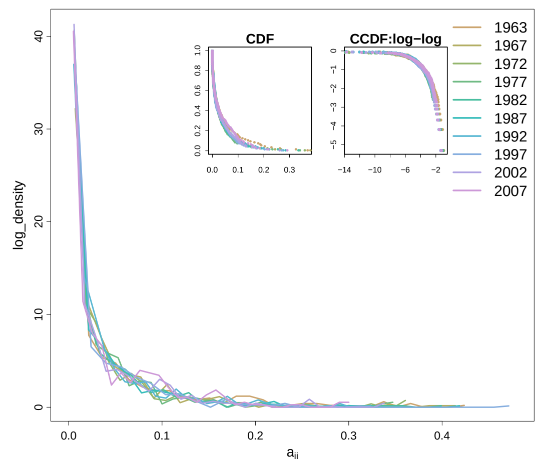


Figure 3: Diagonal coefficients a_{ii}

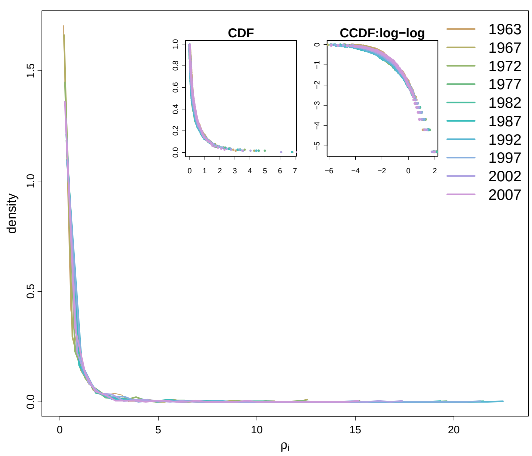


Figure 4: Row sums $\rho_i = \sum_{j=1}^{n} a_{ij}$

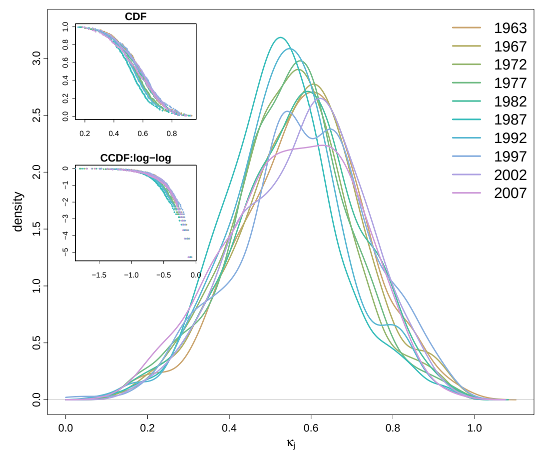


Figure 5: Column sums $\kappa_j = \sum_{i=1}^{n} a_{ij}$

Empirical distributions (ED), cumulative density functions (CDF), and complementary cumulative density functions (CCDF) of different elements of the US input coefficient matrices $\mathbf{A} = \{a_{ij}\}$: 1963-2007, between 351-478 industries. Each one of the figures shows smooth and highly peaked ED. Whereas there is a sharp rate of decay in the ED of the pooled elements a_{ij}, the diagonal elements a_{ii}, and the ρ_i, (Figures 2-4), the ED of the column sums κ_j is highly symmetric (Figure 5). The functional form of the ED, CDF, and CCDF shows a remarkable persistence in spite of the series of technological and organizational changes in the US economy.

changes in coefficients a_{ij}. However, despite of the potential infinite number of different microstates and the observed changes in a_{ij} in the $\mathbf{A}$ matrix for the US, our result shows that the *macro-state* of $\mathbf{A}$, which captures some macro properties of the system, such as the distribution of all the entries, has remained highly stable. That is, even though every coefficient a_{ij} changes over time, the statistical structure of $\mathbf{A}$, appears relatively invariant.

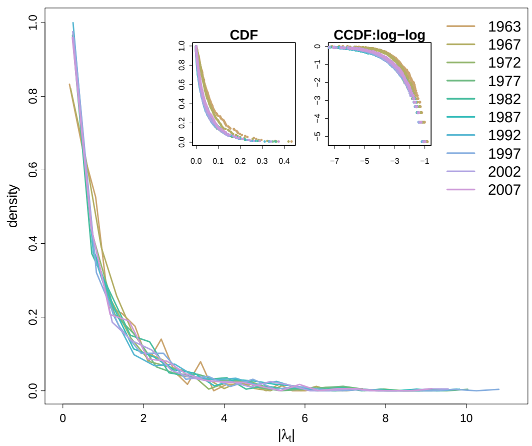


Figure 6: Eigenvalues' moduli $|\lambda_t|$

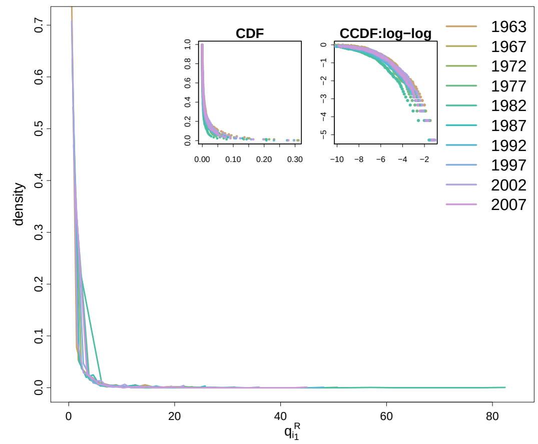


Figure 7: Right-hand eigenvector's coefficient $q^{\mathrm{R}}_{i,1}$

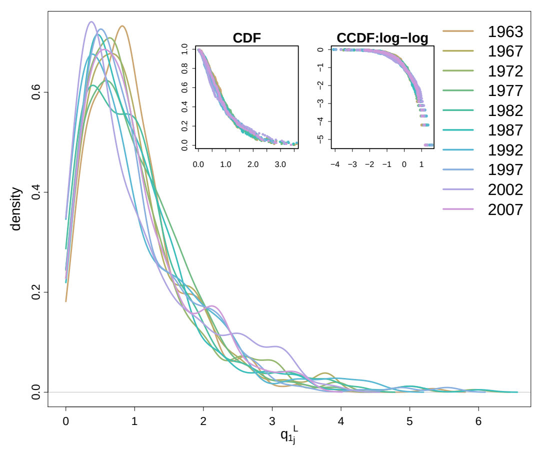


Figure 8: Normalized left-hand eigenvector's coefficient $q^{\mathrm{L}}_{1,j}/\overline{q^{\mathrm{L}}_{1,j}}$, where $\overline{q^{\mathrm{L}}_{1,j}}$ is the mean of $q^{\mathrm{L}}_{1,j}$.

Empirical distributions (ED), cumulative density functions (CDF), and complementary cumulative density functions (CCDF) of the absolute value of eigenvalues $|\lambda_t|$ and the coefficients of the right- $q^{\mathrm{R}}_{i,1}$ and left-hand $q^{\mathrm{L}}_{1,j}$ Perron-Frobenius eigenvectors of the US input coefficient matrices **A***: 1963-2007, between 351-478 industries. Each figure shows smooth and highly peaked ED. Whereas there is a sharp rate of decay in the ED of the $|\lambda_t|$ and $q^{\mathrm{R}}_{i,1}$ (Figure 6-7), the ED of $q^{\mathrm{L}}_{1,j}$ shows sharp increase and then a slower decrease (Figure 8).*

3.3 The Models

In this section we introduce several probabilistic models for our target variables. For each variable we estimate three to six probabilistic models, based on the support of the variable —some variables have support $[0, 1)$ and others $[0, \infty)$. The probabilistic models used for the estimation are Gamma (G), Beta (B), Weibull (W), Exponential (E), Pareto (P), Normal (N) and Log-Normal distributions (LN), whose functional forms are the following:

$$
\begin{aligned}
f_{\mathrm{G}}(x;\alpha,\beta) &= \frac{\beta^{\alpha}x^{\alpha-1}e^{-\beta x}}{\Gamma(\alpha)} \;\text{ for } x\in(0,\infty) && (3)\\
f_{\mathrm{B}}(x;\alpha,\beta) &= \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)} \;\text{ for } x\in[0,1] && (4)\\
f_{\mathrm{W}}(x;\alpha,\beta) &= \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1} e^{-(x/\beta)^{\alpha}} \;\text{ for } x\in[0,\infty) && (5)\\
f_{\mathrm{E}}(x;\alpha) &= \alpha e^{-\alpha x} \;\text{ for } x\in[0,\infty) && (6)\\
f_{\mathrm{P}}(x;x_m,\alpha) &= \frac{\alpha x_m^{\alpha}}{x^{\alpha=1}} \;\text{ for } x\in[x_m,\infty) && (7)\\
f_{\mathrm{LN}}(x;\mu,\sigma) &= \frac{1}{x\sigma\sqrt{2\pi}}e^{-\frac{(\log x-\mu)^2}{2\sigma^2}} \;\text{ for } x\in(0,\infty) && (8)\\
f_{\mathrm{N}}(x;\mu,\sigma) &= \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \;\text{ for } x\in(-\infty,+\infty). && (9)
\end{aligned}
$$

For those variables which have a bounded support such as $0 \leq a_{ij} \leq 1$, a truncated distribution is used. For notational convenience, we do not distinguish the truncated and non-truncated distributions. The following are the probabilistic models we use for each variable:

- For the distributions of the pooled $a_{ij} = [0,1)$, $f(a_{ij})$, and the diagonal entries $a_{ii} = [0,1)$, $f(a_{ii})$, which have a rapidly decaying pattern, we exclude zero values and fit six models: Exponential, Weibull, Beta, Gamma, Pareto, and Log-Normal distributions.

- For the distributions of the column sums $\kappa_j = (0,1)$, $f(\kappa_j)$, which have a symmetrical pattern, we fit three models: Beta, Gamma, and (truncated) Normal distributions.

- For the distributions of the row sums $\rho_i = (0,\infty)$, $f(\rho_i)$, which have a rapidly decaying pattern, we fit five models: Exponential, Weibull, Gamma, Pareto, and Log-Normal distributions.

- For the distributions of eigenvalues' absolute value $|\lambda_t| = [0,1)$, $f(|\lambda_t|)$, and the right-hand PF eigenvector $q^{\mathrm{R}}_{i,1}$, $f(q^{\mathrm{R}}_{i,1})$, which have a rapidly decaying pattern in our data, we fit 6 models after excluding zero values: Exponential, Weibull, Beta, Gamma, Pareto, and Log-Normal distributions.

- For the distribution of the left-hand PF eigenvector $q^{\mathrm{L}}_{1,j}$, $f(q^{\mathrm{L}}_{1,j})$, which have a right skewed pattern, we fit four models: Weibull, Beta, Gamma, and Log-Normal distributions.

3.4 Model Estimation Results

The different parameters of the distributions are estimated using Bayesian methods with a uniform prior. The Hamiltonian Monte Carlo (HMC) algorithm is used for the posterior simulation with 1,000 iterations and 3 chains. Then, we compute the accuracy of the model estimation using leave-one-out cross-validation (LOO-CV) and Widely Applicable Information Criteria (WAIC), which takes the log-likelihood of the model corrected for model complexity. WAIC is asymptotically equal to LOO.[19] See Appendices A and B for detailed discussions on the estimation results and model comparison methods.

Table 1 presents the results from the WAIC index, showing the rank of models in a decreasing order. The smaller the WAIC is, the more preferred the model is. The number within the parenthesis is the standard error of WAIC. For example, the distribution of all entries $f(a_{ij})$ for year 1963 prefers the Log-Normal distribution whose WAIC index is -2912 with the standard error 95. Detailed estimation results for all the models are available in Appendix B.

The most noticeable result we can see from Table 1 is that the probability density function that fits the ED of the variable best is relatively persistent for different years. This temporal persistence of the functional form of the EDs provides evidence for the existence of a statistical structure of the matrices, indicating that structural change in the US matrices **A** might be characterized by changes in the parameters determining the shape, scale, and/or location of the density function, not by changes in the underlying data generating process.

Figures 9 through 15 present the frequency distribution, the fit of different distributions, and the Q-Q plots for one representative example (year 1967), which typifies the patterns observed for other years.

All entries, $f(a_{ij})$: The distributions of all entries are well described by the Log-Normal distribution, $\log(a_{ij}) \sim N(\mu, \sigma^2)$. As is well known, the Log-Normal distribution is one of the fat tailed distributions, meaning that some of the technical coefficients in **A** have relatively large values while most of them are very small. As the heat map of **A** shows, this rapidly decaying pattern with a fat-tail results from the relative importance of the diagonal entries and a small number of rows, further suggesting that, without the diagonal part, the I-O network is highly concentrated on a few sectors.

Diagonal entries, $f(a_{ii})$: The distributions of diagonal entries are well explained by the Weibull, Gamma, and Beta distributions. The Weibull distribution, with its shape parameter below 1, is a subexponential distribution with a heavy tail. Our estimation results show that α_W is far below 1 in all cases. The Gamma distribution can also have significant kurtosis and have a fat tail as well. The kurtosis of the Gamma distribution is calculated by $6/\alpha_G$, the shape parameter. In our estimation result, α_G for the diagonal distribution is normally around 0.5, implying a high kurtosis. The two scale parameters of the Beta distribution α_B and β_B are estimated around 0.5 and 12, respectively, so that the expected

[19]For a detailed discussion on the HMC and its application see Gelman et al. (2013). See Vehtari et al. (2015) for leave-one-out cross-validation and Widely Applicable Information Criteria.

$f(.)$	1963	1967	1972	1977	1982	1987	1992	1997	2002	2007
$f(a_{ij})$	LN : -2912 (95)	LN : -2377 (72)	LN : -5575 (102)	LN : -5742 (104)	LN : -6117 (108)	LN : -5804 (109)	LN : -5785 (109)	LN : -4848 (126)	LN : -4970 (118)	LN : -4812 (125)
	W : -2873 (95)	W : -2321 (80)	W : -5502 (103)	W : -5671 (107)	W : -6023 (108)	W : -5723 (110)	W : -5695 (108)	W : -4814 (124)	W : -4915 (117)	W : -4770 (121)
	G : -2765 (107)	G : -2237 (91)	G : -5308 (114)	G : -5433 (140)	G : -5757 (156)	G : -5464 (153)	G : -5453 (136)	G : -4697 (125)	G : -4779 (118)	G : -4669 (118)
	B : -2748 (111)	B : -2224 (94)	B : -5301 (116)	B : -5418 (147)	B : -5727 (178)	B : -5434 (173)	B : -5436 (146)	B : -4695 (125)	B : -4777 (118)	B : -4668 (118)
	E : -2015 (226)	E : -1904 (137)	P : -4293 (110)	P : -4474 (112)	P : -4855 (114)	P : -4507 (115)	P : -4481 (113)	P : -3577 (130)	P : -3670 (122)	P : -3524 (127)
	P : -1373 (89)	P : -690 (70)	E : -3239 (275)	E : -3099 (471)	E : -3060 (642)	E : -3026 (614)	E : -3091 (483)	E : -3035 (214)	E : -3087 (202)	E : -3120 (166)
$f(a_{ii})$	W : -1537 (64)	W : -1635 (51)	W : -2044 (67)	G : -2224 (69)	W : -2141 (66)	W : -2224 (67)	B : -2186 (70)	G : -2290 (73)	B : -1933 (68)	B : -1605 (59)
	G : -1536 (63)	LN : -1629 (54)	G : -2042 (67)	W : -2223 (68)	G : -2135 (66)	G : -2219 (67)	G : -2185 (70)	B : -2288 (73)	G : -1933 (68)	G : -1603 (59)
	B : -1535 (62)	G : -1626 (51)	B : -2036 (68)	B : -2219 (69)	B : -2131 (66)	B : -2216 (67)	W : -2173 (68)	W : -2284 (71)	W : -1920 (67)	W : -1592 (59)
	LN : -1493 (70)	B : -1617 (52)	LN : -1943 (68)	LN : -2094 (69)	LN : -2053 (70)	LN : -2125 (70)	E : -2009 (58)	LN : -2104 (74)	E : -1755 (55)	LN : -1479 (63)
	E : -1340 (55)	E : -1602 (53)	E : -1870 (63)	E : -2054 (64)	E : -1980 (63)	E : -2057 (63)	LN : -1988 (70)	E : -2089 (62)	LN : -1740 (69)	E : -1471 (48)
	P : 1340 (79)	P : 1920 (56)	P : 1595 (86)	P : 1662 (90)	P : 1758 (84)	P : 1656 (86)	P : 1748 (98)	P : 1590 (99)	P : 1412 (97)	P : 1451 (82)
$f(\rho_i)$	LN : 231 (56)	LN : 213 (67)	LN : 77 (73)	LN : 93 (75)	W : 147 (77)	LN : 18 (76)	LN : 47 (77)	W : 194 (75)	W : 215 (68)	W : 193 (62)
	W : 262 (57)	W : 261 (68)	W : 116 (74)	W : 134 (76)	LN : 168 (84)	W : 34 (76)	W : 55 (78)	LN : 230 (74)	G : 256 (77)	G : 213 (69)
	G : 287 (61)	G : 309 (74)	G : 177 (78)	G : 200 (84)	G : 207 (85)	G : 89 (82)	G : 116 (86)	G : 244 (84)	LN : 262 (73)	LN : 277 (57)
	E : 325 (71)	E : 392 (91)	E : 365 (102)	E : 390 (112)	E : 418 (113)	E : 330 (108)	E : 366 (116)	E : 430 (111)	E : 375 (100)	E : 311 (82)
	P : 2527 (53)	P : 3036 (62)	P : 2848 (73)	P : 3047 (74)	P : 2996 (81)	P : 2746 (80)	P : 2826 (81)	P : 2988 (80)	P : 2721 (72)	P : 2481 (73)
$f(\kappa_j)$	N : -326 (27)	N : -412 (30)	N : -468 (31)	N : -483 (32)	B : -435 (29)	N : -515 (33)	N : -490 (33)	N : -369 (30)	B : -388 (25)	B : -323 (23)
	B : -317 (27)	B : -404 (29)	B : -457 (31)	B : -468 (33)	N : -434 (30)	B : -495 (34)	B : -473 (33)	B : -365 (32)	N : -379 (26)	N : -312 (24)
	G : -286 (32)	G : -373 (33)	G : -438 (34)	G : -439 (35)	G : -387 (34)	G : -495 (38)	G : -450 (40)	G : -290 (54)	G : -344 (29)	G : -277 (26)
$f(\lvert\lambda_i\rvert)$	E : -1066 (39)	B : -1492 (63)	B : -1999 (106)	B : -2368 (164)	B : -2196 (140)	B : -2461 (189)	B : -2724 (227)	B : -2502 (196)	B : -1779 (92)	B : -2049 (207)
	W : -1065 (39)	G : -1486 (60)	G : -1995 (104)	G : -2360 (163)	G : -2188 (138)	G : -2452 (188)	G : -2711 (226)	G : -2494 (194)	G : -1778 (90)	G : -2038 (206)
	G : -1065 (39)	W : -1469 (50)	W : -1942 (86)	W : -2246 (131)	W : -2100 (112)	W : -2317 (157)	W : -2535 (194)	W : -2351 (161)	W : -1744 (74)	W : -1885 (175)
	B : -1060 (39)	E : -1453 (45)	E : -1774 (55)	E : -1923 (57)	E : -1839 (55)	E : -1838 (57)	E : -1892 (58)	E : -1882 (59)	E : -1632 (52)	E : -1415 (48)
	LN : -1016 (44)	LN : -680 (94)	LN : -1283 (73)	LN : -1462 (143)	LN : -1353 (108)	LN : -1620 (175)	LN : -1829 (225)	LN : -1620 (184)	LN : -1094 (63)	LN : -1299 (209)
	P : 1337 (48)	P : 1284 (152)	P : 597 (182)	P : 232 (259)	P : 432 (229)	P : -92 (277)	P : -394 (321)	P : -106 (289)	P : 613 (169)	P : -181 (296)
$f\left(q^{\mathrm{R}}_{i,1}\right)$	LN : -2410 (77)	LN : -3318 (89)	LN : -3499 (90)	W : -3962 (97)	W : -4399 (103)	W : -3501 (94)	W : -3882 (100)	W : -3603 (96)	W : -2900 (87)	W : -2855 (86)
	W : -2364 (73)	W : -3263 (87)	W : -3477 (91)	LN : -3937 (93)	LN : -4345 (103)	LN : -3490 (93)	LN : -3859 (100)	G : -3533 (101)	G : -2849 (89)	G : -2801 (92)
	G : -2305 (72)	G : -3177 (86)	G : -3384 (92)	G : -3855 (105)	G : -4262 (123)	G : -3423 (94)	G : -3794 (104)	LN : -3525 (92)	B : -2836 (90)	LN : -2798 (83)
	B : -2294 (72)	B : -3162 (87)	B : -3366 (94)	B : -3823 (111)	B : -4203 (141)	B : -3408 (95)	B : -3768 (110)	B : -3513 (104)	LN : -2801 (94)	B : -2770 (101)
	E : -2004 (89)	E : -2698 (114)	E : -2852 (127)	E : -3235 (162)	E : -3485 (236)	E : -2844 (119)	E : -3078 (149)	E : -2943 (131)	E : -2413 (101)	E : -2330 (127)
	P : 634 (73)	P : 427 (86)	P : 295 (92)	P : 22 (102)	P : -518 (110)	P : 283 (101)	P : -69 (109)	P : 211 (110)	P : 496 (101)	P : 215 (98)
$f\left(q^{\mathrm{L}}_{1,j}\right)$	G : -1380 (31)	W : -1523 (34)	G : -1968 (40)	W : -1689 (31)	W : -2040 (36)	G : -1808 (40)	G : -1921 (40)	G : -1024 (39)	B : -782 (34)	G : -605 (31)
	B : -1380 (31)	B : -1522 (34)	B : -1965 (40)	B : -1686 (32)	B : -2039 (36)	B : -1803 (41)	B : -1919 (39)	W : -1020 (39)	G : -781 (36)	B : -604 (29)
	W : -1377 (32)	G : -1520 (34)	W : -1964 (40)	G : -1682 (32)	G : -2038 (36)	W : -1798 (41)	W : -1918 (39)	B : -1011 (41)	W : -780 (35)	W : -603 (30)
	LN : -368 (20)	LN : -387 (26)	LN : -723 (30)	LN : -416 (26)	LN : -740 (31)	LN : -558 (26)	LN : -715 (32)	LN : -35 (24)	LN : -10 (26)	LN : 128 (18)

Table 1: Widely Applicable Information Criteria index of the estimated probabilistic models for different components of the US input coefficient matrices $\mathbf{A} = \{a_{ij}\}$: 1963-2007, between 351-478 industries. The components of $\mathbf{A}$ are: the pooled entries a_{ij}, the diagonal elements a_{ii}, the row sums ρ_i, the column sums κ_j, the absolute value of the eigenvalues $|\lambda_i|$, and the entries of the left- and right-hand Perron-Frobenius eigenvectors $q^{\mathrm{L}}_{1,j}$ and $q^{\mathrm{R}}_{i,1}$. The initials for each probability model are G = Gamma, B = Beta, W = Weibull, E = Exponential, P = Pareto, LN = Log Normal, and N = Normal. The numbers in the parenthesis are the standard error of the estimated WAIC index. The lower the WAIC index is, the more the model is preferred.

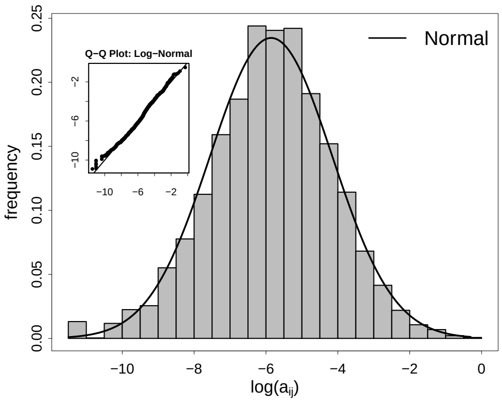


Figure 9: All coefficients, $\log(a_{ii})$*: 1967*

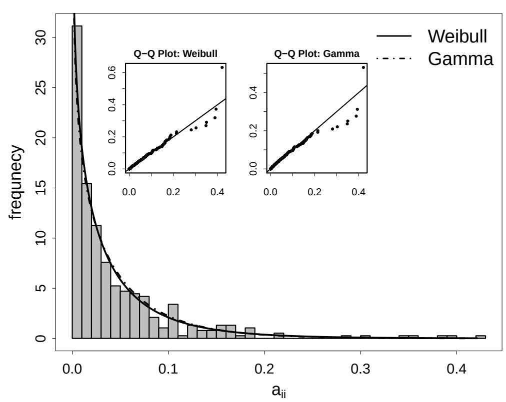


Figure 10: Diagonal coefficients a_{ii}*: 1967*

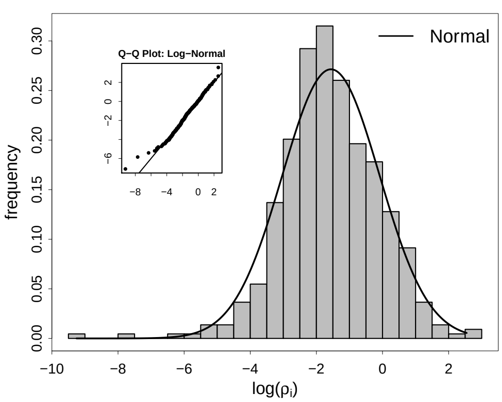


Figure 11: Row Sums ρ_i*: 1967*

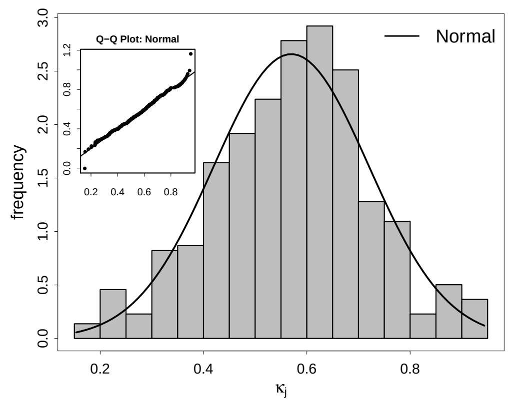


Figure 12: Column Sums κ_j*: 1967*

Histograms, fitted probability distributions models, and Q-Q plots for different aspects of the US input coefficient matrices $\mathbf{A} = \{a_{ij}\}$*, year 1967*

value is around 0.04 while having a high degree of right skewness and kurtosis, confirming its fat-tailed behavior. Considering the fact that the diagonal entries are among the most pronounced part of **A**, the fat tailed behavior in their distribution suggests that even among the relatively important entries, the relative difference between the a_{ii} entries is significant, with a few a_{ii} having very high values. It is worthwhile to note, however, that the tail of diagonal entry distribution is less heavy than that of all entries distribution whose kurtosis is normally three times higher than the diagonal counterpart.

Column sums, $f(\kappa_j)$: The distributions of column sums are well described by the (truncated) Normal distribution in almost all samples. Some years have the Beta distribution as the best performing distribution, whose two scale parameters α_B and β_B are very close to each other, implying the symmetry of the distribution. Since the column sum distribution represents the industry dispersion of the ratio of intermediate commodities to output (whose inverse is the productivity of intermediate commodities), the symmetric unimodal pattern of the distribution suggests that there is a cross-sectional convergence

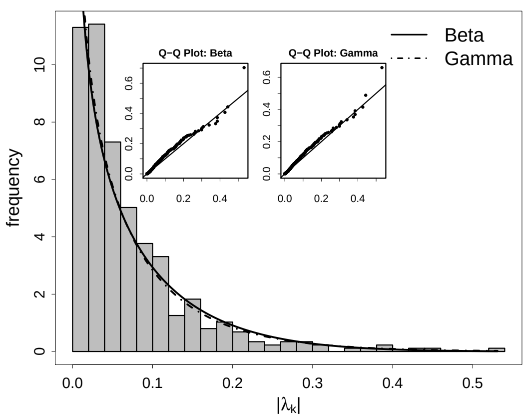


Figure 13: Eigenvalues' moduli $|\lambda_t|$*:1967*

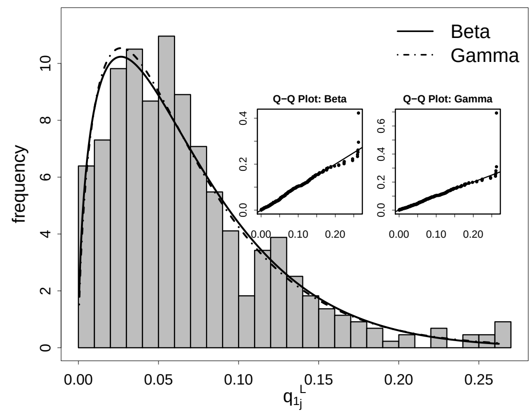


Figure 14: Left-hand Perron-Frobenius eigenvector's coefficient $q^{\mathrm{L}}_{1,j}$*:1967*

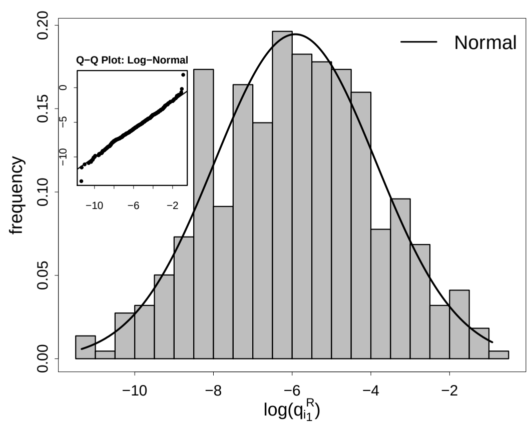


Figure 15: Right-hand Perron-Frobenius eigenvector's coefficient $\log q^{\mathrm{R}}_{i,1}$*:1967*

Histograms, fitted probability distributions models, and Q-Q plots for different aspects of the US input coefficient matrices $\mathbf{A} = \{a_{ij}\}$*, year 1967*

of the productivity of circulating capital. Also, in models of prices of production with uniform wage w and profit (or interest) rates r, this observed symmetry suggest a cross-sectional convergence in the capital-labor ratios k_j, because $\kappa_j = \frac{1}{1+wk_j+r}$.

Row sums, $f(\rho_i)$: The distributions of row sums are well described by the Log-Normal distribution and the Weibull distribution with the shape parameter α_W being far below 1. As discussed above, these two distributions capture the fat-tailed behavior of the row sums. Since the row sums of $\mathbf{A}$ represent the relative importance of industry i's product in the production of others, the result shows a production structure in which the supply chain network concentrates on a small number of industries.

Eigenvalues' moduli, $f(|\lambda_t|)$: The distributions of eigenvalues' moduli are well described by the Beta distribution. The Gamma and Weibull distributions are also a good models chasing the Beta with a small margin. The two scale parameters of the Beta distribution α_B and β_B are estimated around 0.5 and 9, respectively, so that the expected value is around 0.05 while having a high degree of right skewness and kurtosis. The persistent behavior of eigenvalues' moduli suggests a statistical tendency to a certain degree of proportionality among the columns of **A**, i.e. among industries' intermediate commodity inputs.[20] For each linearly dependent column in **A** there is a corresponding zero eigenvalue. The fact that the expected value of eigenvalues' moduli is close to zero suggests that **A** can be well approximated by a linear combination of a small number of matrices $\mathbf{A}_t = \frac{1}{\mathbf{q}_t^{\mathrm{L}}\mathbf{q}_t^{\mathrm{R}}}\mathbf{q}_t^{\mathrm{R}}\mathbf{q}_t^{\mathrm{L}}$, that is, as $\mathbf{A} \approx \sum_{t=1}^{c} \lambda_t \mathbf{A}_t$ for a small $c \ll n$. However, it is worthwhile to note that the relatively heavy tail behavior implies that there are several eigenvalues with a significant magnitude.

Perron-Frobenius eigenvectors, $f(q_{1,j}^{\mathrm{L}})$ and $f(q_{i,1}^{\mathrm{R}})$: The distributions of the right-hand eigenvectors prefer either the Weibull or the Log-Normal distributions as the best fitting model, while the distributions of the left-hand eigenvectors prefer either the Gamma or the Weibull distributions. All estimated Gamma and Weibull distributions consistently have a high degree of kurtosis and exhibit a fat-tailed behavior. In linear production models, the left-hand PF eigenvector is associated with the vectors of labor inputs **l** and prices whereas the right-hand PF eigenvector is associated with the output vectors $\mathbf{x}^{\mathrm{T}}$. Under the assumption that **A** is diagonalizable, vectors **l** and $\mathbf{x}^{\mathrm{T}}$ can be represented as a linear combination of the eigenvectors, $\mathbf{l} = l_1^* \mathbf{q}_1^{\mathrm{L}} + \ldots + l_n^* \mathbf{q}_n^{\mathrm{L}}$ and $\mathbf{x}^{\mathrm{T}} = x_1^* \mathbf{q}_1^{\mathrm{R}} + \ldots + x_n^* \mathbf{q}_n^{\mathrm{R}}$. Hence, the empirical evidence given in Torres (2017, 2018) on the statistical tendency towards the proportionality, $\mathbf{l} \propto \mathbf{q}_1^{\mathrm{L}}$ and $\mathbf{x}^{\mathrm{T}} \propto \mathbf{q}_1^{\mathrm{R}}$, suggests that there is a close relationship between the distribution of $\mathbf{q}_1^{\mathrm{L}}$ and **l** and the distribution of $\mathbf{q}_1^{\mathrm{R}}$ and $\mathbf{x}^{\mathrm{T}}$. The fact that eigenvector distributions have a fat-tailed behavior allows us to have a more systematic approach to studying the distributions of the labor and output vector since the latter could also have a fat tailed behavior.

4 Discussion

In the previous section we reported on the existence of smooth, unimodal, and highly peaked ED for all the aspects of **A** matrices for the US. It was shown that the functional form is time invariant for almost all samples. The use of the frequency and cumulative distributions gives an aggregate view of the matrices (an account of their macro-state) consistent with the holistic approach advocated by Jensen et al. (1988) and Jensen et al. (1991). Based on these results we conclude that our findings provide evidence for the existence of a statistical structure in **A** matrices for the US.

There are three theoretical aspects that we can now address based on the contrast between the empirical evidence obtained in Section 3 and some aspects of the literature reviewed in the paper.

[20]This tendency toward proportionality in the columns of **A** is used by Mariolis and Tsoulfidis (2011) and Schefold (2013) as a hypothesis to explain certain regularities in empirical computation of prices of production models.

The Origin of Randomness: Section 2.1 showed that, with few exceptions, for the stochastic I-O analysis literature, randomness of the elements in the tables (u_{ij}, v_{ij}) and matrices (z_{ij}, a_{ij}) comes from *errors* generated during the construction of the I-O accounts. However, it is hard to interpret the empirical regularities in the data as the outcome of systematic measurement errors since the construction methods of the I-O accounts in the US have changed drastically since 1972 while the statistical structure remained highly stable. The persistent patterns observed in many aspects of **A** matrices over 40 years suggest that a more fundamental force than just human error might be in place, which requires further research looking into the economic forces behind these patterns.

Probabilistic Assumptions vs. Empirical Regularities: The surveys of the stochastic I-O analysis literature by West (1986), Jackson (1989), and Temurshoev (2017) show that such studies rely frequently on probabilistic assumptions of the a_{ij} components such as an additive or multiplicative random structure, normality, symmetry, independence, identity in distribution, among others. For instance, Bródy (1997) and Bidard and Schatteman (2001) assume that $a_{ij} \sim i.i.d(\mu, \sigma^2)$, while Schefold (2013) uses the random matrix model of Goldberg and Neumann (2003), which is similar to that of Bródy's but allows for intracolumn dependence (but intercolumn independence). However, these assumptions are made for mathematical convenience without relying on empirical evidence. West mentions that '[n]ormality was assumed because this is the distribution commonly invoked with sampling theory, and is probably more amenable to mathematical manipulation ... many of their assumptions appear too restrictive to be of practical use' (West 1986, p. 364).

The statistical behavior of the pooled coefficients a_{ij}, the diagonal elements a_{ii}, and the column κ_j and row ρ_i sums from the US' **A** matrices contrasts with that derived by the random matrix model in Bródy (1997) and other authors within the Bródy's conjecture debate. In Bródy's model, each column of matrix **A** is a sample of size n from *the same* positive and continuous probability distribution $\{x\}$ with $E(x) = \mu$ and $E[(x - \mu)^2] = \sigma^2$. If **A** matrices are generated by this probabilistic model, then there should be no difference in the distribution between a_{ij} and a_{ii}, nor between κ_j and ρ_i. What we observe in our sample are persistent differences in the probability density functions between a_{ij} and a_{ii} and between κ_j and ρ_i. This contrast provides evidence against the idea that the columns' (or rows') entries from the US's **A** matrices have the same data-generating process.

Furthermore, unlike the assumption made in Bródy (1997), a_{ij} coefficients are far from being normally distributed. They are asymmetrically distributed with a particularly sharp decay with a fat-tail. Also, the assumption that a_{ij} are distributed as a Beta distribution, sometimes specified in the literature (see ten Raa and Steel, 1994), finds weak support against the alternative distributions such as the Log-Normal or Weibull distributions.

A Fresh Look at Structural Analysis of A matrices: The **A** matrix is a representation of the technological structure of the economy with respect to intermediate commodity inputs for a given year. Traditionally, changes in this technological structure are seen as changes in the a_{ij} coefficients (see the pioneer works by Leontief, 1951 and Carter, 1970), which we call different ”micro-states” of **A**. As we discussed, structural I-O analysis has focused on tracking the changes in the micro-state of the input-output relation but paid little attention to the macro-state stability represented by the persistent macro structure of **A**. However, the consistency of observed distributions of the elements in **A** calls for a more comprehensive I-O modeling that could account for micro-state fluctuations along with the macro-state stability in a unified manner.[21]

5 Conclusion

In this paper we showed that there is a statistical structure in the US' I-O coefficient matrices, **A**, and characterized the identified statistical regularities by carrying out a statistical exploratory data analysis and model estimation on 10 different **A** matrices from the period 1963-2007. We found a series of regularities. First, there are smooth, unimodal, and highly peaked empirical distributions for several aspects of **A** matrices —the pooled and diagonal elements, the column sums and row sums, eigenvalues' moduli, and the elements of the left- and right-hand Perron-Frobenius eigenvectors. Second, the distributions of all entries, diagonal entries, row sums, and the PF eigenvectors are well described by a fat-tailed distribution such as the Log-Normal, the Weibull with the below unity shape parameter, and the Gamma distribution with a high kurtosis. In contrast, the distributions of column sums are well explained by the Normal distribution due to its symmetry, while the distributions of the eigenvalues' moduli prefer the Beta distribution. Finally, the general characteristics of these frequency distributions are maintained for almost all samples. This result is informative considering the fact that the US economy experienced tremendous changes in its technology, institutions, and the organization of production. The general patterns are also invariant to the methodologies in the construction of the I-O accounts and matrices.

In Sections 3 and 4 we gave a primer interpretation of these results and pointed to some of their implications and potential uses for the I-O economics literature. However, further work needs to be done to evaluate if the identified statistical stylized facts of **A** matrices are universal across countries. In addition, it would be interesting to assess if matrices that capture other aspects of the interindustry relationships, like **Z**, **L**, **J**, or the capital stock coefficient **B** matrices (see Leontief and Others, 1953) have a persistent statistical structure, and if so, how these differ from those of **A**.

We want to thank Duncan Foley, Geoffrey Hewings, Michael Lahr, and Anwar Shaikh for stimulating conversations and detailed comments on early drafts of this paper. We would also like to thank Fidel Aroche, Tom Cender, Henryk Gurgul, Theodore Mariolis, and Lefteris Tsoulfidis for their comments. The first version was presented at the International Input-Output Conference at Seoul, South Korea, 4-8 July 2016. We thank all our commentators on that occasion. The usual disclaimer applies. Daniel Torres acknowledges financial support by CONACYT and DGRI-SEP.

[21]The statistical stability of different aspects of the matrix **A** will also serve as informative constraints for biproportional methods such as RAS or Maximum Entropy in estimating missing entries and updating/creating the matrix **A**.

References

Aruka, Y. (2011). *Complexities of Production and Interacting Human Behavior*. Physica-Verlag.

BEA (2009). *Concepts and Methods of the U.S. Input-Output Accounts*. Bureau of Economic Analysis, U.S. Department of Commerce, updated edition.

Bidard, C. and Schatteman, T. (2001). The spectrum of random matrices. *Economic Systems Research*, 13(3):289–298.

Bródy, A. (1997). The second eigenvalue of the leontief matrix. *Economic Systems Research*, 9(3):253–258.

Carter, A. P. (1970). *Structural Change in the American Economy*. Harvard University Press.

Foerster, A. and Choi, J. (2017). The changing input-output network structure of the u.s. economy. *Federal Reserve Bank of Kansas City*, Second Quarter:23–49.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013). *Bayesian Data Analysis. Third Edition*. CRC Press.

Gelmann, A., Hwang, J., and Vehtari, A. (2014). Understanding predictive information criteria for bayesian models. *Statistics and Computing*, 24(6):997–1016.

Goldberg, G. and Neumann, M. (2003). Distribution of subdominant eigenvalues of matrices with random rows. *SIAM Journal on Matrix Analysis and Applications*, 24(3):747–761.

Goldberg, G., Okunov, P., Neumann, M., and Schneider, H. (2000). Distribution of subdominant eigenvalues of random matrices. *Methodology and Computing in Applied Probability*, 2(2):137–51.

Goodwin, R. M. (1974). The use of normalized general coordinates in economic analysis. In Mitra, A., editor, *Economic Theory and Planning: Essays in Honour of A. K. Das Gupta*, chapter 3, pages 26–38. Calcutta: Oxford University Press.

Goodwin, R. M. (1976). Use of normalized general coordinates in linear value and distribution theory. In Polenske, K. and Skolka, J., editors, *Advances in Input-Output Economics*, chapter 29, pages 581–602. Cambridge, Mass.: Ballinger Publ. Co.

Gowdy, J. M. (1991). Structural change in the usa and japan. an extended input-output analysis. *Ecnomic Systems Reseach*, 3(4):413–423.

Gurgul, H. and Wojtowicz, T. (2015). On the economic interpretation of the brody's conjecture. *Economic Systems Research*, 27(1):122–131.

Iliadi, F., Mariolis, T., Soklis, G., and Tsoulfidis, L. (2014). Bienenfeld's approximation of production prices and eigenvalue distribution: further evidence from five european economies. *Contributions to Political Economy*, 33(1):35–54. Advance Access published May 25, 2014.

Jackson, R. W. (1986). The full-distribution approach to aggregate representation in the input-output modeling framework. *Journal of Regional Science*, 26(3):515–531.

Jackson, R. W. (1989). Probabilistic input-output analysis: Modeling directions. *Socio-Economic Planning Sciences*, 23(1-2):87–95.

Jackson, R. W. and West, G. R. (1989). Prospectives on probabilistic input-output analysis. In Miller, R. E., Polenske, K. R., and Rose, A. Z., editors, *Frontiers of Input-Output Analysis*. Oxford University Press.

Jensen, R., Dewhurst, J., West, G., and Hewings, G. (1991). On the concept of fundamental economic structure. In Dewhurst, J., Jensen, R., and Hewings, G., editors, *Regional Input-Output Modelling*, chapter 12. Avebury.

Jensen, R., West, G., and Hewings, G. (1988). The study of regional economic structure using input-output tables. *Regional Studies*, 22(3):209–220.

Leontief, W. (1951). *The Structure of American Economy, 1919-1939: An Empirical Application of Equilibrium Analysis*. 2nd Edition, Harvard University Press.

Leontief, W. (1986). *Input-Output Economics*. 2nd Edition, Oxford University Press.

Leontief, W. and Others (1953). *Studies in the Structure of the American Economy*. Oxford University Press.

Mariolis, T. and Tsoulfidis, L. (2011). Eigenvalue distribution and the production price-profit rate relationship in linear single-product systems: theory and empirical evidence. *Evolutionary and Institutional Economics Review*, 8(1):87–122.

Mariolis, T. and Tsoulfidis, L. (2014). On brody's conjecture: theory, facts and figures about instability of the us economy. *Economic Systems Research*, 26(2):209–223.

Mariolis, T. and Tsoulfidis, L. (2016a). Capital theory 'paradoxes' and paradoxical results. resolved or continued? *Evolutionary and Institutional Economics Review*, 13(2):297–322.

Mariolis, T. and Tsoulfidis, L. (2016b). *Modern Classical Economics and Reality. A Spectral Analysis of the Theory of Value and Distribution*. Springer.

Mariolis, T. and Tsoulfidis, L. (2018). Less is more. capital theory and almost irregular-uncontrollable actual economies. *Contributions to Political Economy*, 37(1):65–88.

Meyer, C. D. (2000). *Matrix analysis and applied linear algebra*. SIAM: Society for Industrial and Applied Mathematics.

Miller, R. E. and Blair, P. D. (2009). *Input-Output Analysis. Foundations and Extensions*. Cambridge University Press.

Nassif, L. and Shaikh, A. (2015). Eigenvalue distribution, matrix size and the linearity of wage-profit curves. 23rd IIOA conference, June 22-26 2015, Mexico City.

Pasinetti, L. L. (1977). *Lectures on the theory of production*. Palgrave Macmillan.

Pasinetti, L. L. (1988). Sraffa on income distribution. *Cambridge Journal of Economics*, 12(1):135–138.

Rueda, J. M., Dietzenbacher, E., Fernandez, E., and Amores, A. (2013). The bias of the multiplier matrix when supply and use tables are stochastic. *Economic Systems Research*, 25(4):435–448.

Schefold, B. (2013). Approximate surrogate production functions. *Cambridge Journal of Economics*, 37(5):1161–1184.

Shaikh, A. (2016). *Capitalism: Competition, Conflict, Crises*. Oxford: University Press.

Sun, G.-Z. (2008). The first two eigenvalues of large random matrices and brody's hypothesis on the stability of large input-output systems. *Economic Systems Research*, 20(4):429–432.

Takayama, A. (1974). *Mathematical Economics*. The Driden Press.

Tarancón, M. A. (2003). *Técnicas de Análisis Económico Input-Output*. Editorial Club Universitario.

Team, S. D. (2018). *RStan: the R interface to Stan, Version 2.17.3*.

Temurshoev, U. (2017). Uncertainty treatment in input-output analysis. In ten Raa, T., editor, *Handbook of Input-Output Analysis*, chapter 12. Edward Elgar Publishing.

ten Raa, T. (2005). *The Economics of Input-Output Analysis*. Cambridge University Press.

ten Raa, T. and Steel, M. F. (1994). Revised stochastic analysis of an input-output model. *Regional Science and Urban Economics*, 24:361–371.

Torres, L. D. (2017). Regularities in prices of production and the concentration of compositions of capitals. Working Paper 09/2017, Department of Economics, The New School for Social Research.

Torres, L. D. (2018). The near-linearity in the empirical relation between wages and the rate of profits: an explanation based on the statistical structure of technology and demand. Mimeo.

Vehtari, A., Gelmann, A., and Gabry, J. (2015). Efficient implementation of leave-one-out cross-validation and waic for evaluating fitted bayesian models. *Unpublished: http://www.stat.columbia.edu/ gelman/research/unpublished/*.

Verspagen, B. (2004). Structural change and technology. *Revue èconomique*, 55(6):1099–1126.

Watanabe, S. (2010). Asymptotic equivalence of bayes cross validation and widely application information criterion in singular learning theory. *Journal of Machine Learning Research*, 11:3571–3594.

Watanabe, S. (2013). A widely applicable bayesian information criterion. *Journal of Machine Learning Research*, 14(1):867–897.

West, G. R. (1986). A stochastic analysis of an input-output model. *Econometrica*, 54(2):363–374.

Wibe, S. (1982). The distribution of input coefficients. *Economics of Planning*, 18(2):65–70.

Appendices

A HMC and WAIC&LOO-CV

The Hamiltonian Monte Carlo (HMC) algorithm, which is also known as Hybrid Monte Carlo (Duane 1987), is a Markov Chain Monte Carlo (MCMC) method —a fundamental simulation technique in Bayesian Statistics. HMC is known to be more efficient than other random walk MCMC algorithms such as the Metropolis-Hastings due to the gradient information of a dynamical system constructed from the target density. This paper uses RStan software that implements an HMC algorithm along with a particular sampler, called No U-turn sampler (Team, 2018). For a detailed technical discussion on HMC with No U-turn sampler see Gelman et al. (2013).

For a model evaluation and comparison, we use a fully Bayesian method called Widely Applicable Information Criteria (WAIC) and Leave-one-out cross-validation (LOO-CV) (Watanabe, 2010, 2013). They are fully Bayesian because they utilize the entire posterior distribution rather than a point estimation. WAIC is asymptotically equal to LOO. For detailed survey on different model selection criteria, see Gelmann et al. (2014).

Here, we briefly summarize the WAIC calculation. First, the model fit or the posterior predictive accuracy is assessed using the log pointwise predictive density (LPPD). For each data point $i = 1, \dots, n$, the expected log-likelihood of the data is calculated given the posterior distribution and is summed over the entire data points:

$$\text{LPPD} = \sum_{i=1}^{n} \text{Log}\left[\int p[y_i|\theta]p[\theta|y]d\theta\right]. \tag{10}$$

where y is data and θ is a set of parameters. The LPPD can be evaluated by drawing from the posterior simulations. Denoting the simulated θ to be θ_s, $s = 1, \dots, S$, the computed LPPD, which I denote with an asterisk, LPPD*, is written as:

$$\text{LPPD}^* = \sum_{i=1}^{n} \text{Log}\left[\frac{1}{S}\sum_{s=1}^{S} p[y_i|\theta_s]\right] \tag{11}$$

The higher the LPPD is, the more accurately the model fits the data.

Since the LPPD of observed data y is solely based on the log-likelihood, it does not account for the number of parameters in the model and therefore is prone to an overestimation of the predictive accuracy. One way to deal with this problem is to augment the LPPD with a measure of bias correction. In WAIC, the correction is made by subtracting the 'effective number of parameters' of the model ep_{WAIC}, which is a measure of 'model complexity'. It is defined as the difference between the model deviance given the posterior distribution, $2\,\text{Log}[E[p[y_i|\theta]]]$, and the expectation of the model deviance, $E[2\,\text{Log}[p[y_i|\theta]]]$:

$$\text{EP}_{\text{WAIC}} = \sum_{i=1}^{n} (2\,\text{Log}[E[p[y_i|\theta]]] - E[2\,\text{Log}[p[y_i|\theta]]]), \tag{12}$$

which can be computed by:

$$\text{EP}^*_{\text{WAIC}} = 2\sum_{i=1}^{n}\left(\text{Log}\left[\frac{1}{S}\sum_{i=1}^{S} p[y_i|\theta_s]\right] - \frac{1}{S}\sum_{s=1}^{S}\text{Log}[p[y_i|\theta_s s]]\right). \tag{13}$$

Finally, WAIC is obtained by subtracting EP_{WAIC} from the model fit, LPPD. In deviation scale, the measure is multiplied by -2 so that WAIC is written as follows:

$$\text{WAIC} = -2(\text{LPPD} - \text{EP}_{\text{WAIC}}). \tag{14}$$

The lower WAIC, the more accurate the posterior prediction is for new data and the more parsimonious the model is.

B Parameter Estimation

The following table summarizes the results of parameter estimation for all models we use for each variable of **A**– all entries a_{ij}, the diagonal elements a_{ii}, the column sums κ_j and row sums ρ_i, eigenvalues' moduli $|\lambda_i|$, and the elements of the left- and right-hand Perron-Frobenius eigenvectors, the eigenvectors associated to the maximal eigenvalue, $q^{\text{L}}_{1,j}$ and $q^{\text{R}}_{i,1}$:

$f(.)$	Parameter	1963	1967	1972	1977	1982	1987	1992	1997	2002	2007
$f(a_{ij})$	α_E	154.43 (0.85)	90.62 (0.63)	502.28 (1.31)	602.32 (1.55)	659.67 (1.38)	633.95 (1.55)	587.88 (1.45)	534.42 (1.52)	528.21 (1.62)	526.56 (1.49)
	α_W	0.42 (0)	0.61 (0)	0.32 (0)	0.31 (0)	0.29 (0)	0.29 (0)	0.28 (0)	0.3 (0)	0.31 (0)	0.31 (0)
	β_W	0.0017 (0)	0.0066 (0)	0.0002 (0)	0.0001 (0)	0.0001 (0)	0.0001 (0)	0.0001 (0)	0.0002 (0)	0.0002 (0)	0.0002 (0)
	α_B	0.26 (0)	0.46 (0)	0.18 (0)	0.17 (0)	0.16 (0)	0.16 (0)	0.16 (0)	0.18 (0)	0.18 (0)	0.18 (0)
	β_B	38.58 (0.48)	39.11 (0.51)	85.5 (0.64)	99.21 (0.7)	98.62 (0.59)	96.09 (0.67)	88.9 (0.68)	91.58 (0.57)	93.92 (0.62)	92.45 (0.76)
	α_G	0.27 (0)	0.47 (0)	0.18 (0)	0.18 (0)	0.16 (0)	0.16 (0)	0.16 (0)	0.18 (0)	0.18 (0)	0.18 (0)
	β_G	41.66 (0.55)	43.05 (0.54)	90.84 (0.67)	105.8 (0.67)	104.48 (0.74)	101.51 (0.72)	93.69 (0.67)	95.76 (0.62)	97.79 (0.71)	96.02 (0.78)
	α_P	0.02 (0)	0.02 (0)	0.02 (0)	0.02 (0)	0.02 (0)	0.02 (0)	0.02 (0)	0.02 (0)	0.01 (0.01)	0.02 (0)
	μ_{LN}	-7.63 (0.01)	-5.86 (0.01)	-10.35 (0.01)	-10.67 (0.01)	-11.3 (0.01)	-11.2 (0.01)	-11.16 (0.01)	-10.46 (0.01)	-10.29 (0.01)	-10.35 (0.01)
	σ_{LN}	2.65 (0.01)	1.7 (0.01)	3.13 (0.01)	3.18 (0.01)	3.67 (0.01)	3.52 (0.01)	3.62 (0.01)	3.6 (0.01)	3.54 (0.01)	3.61 (0.01)
$f(a_{ii})$	α_E	20.25 (1.08)	22.29 (1.18)	25.9 (1.3)	27.77 (1.3)	25.61 (1.18)	28.15 (1.4)	25.64 (1.16)	28.38 (1.34)	26.66 (1.4)	22.6 (1.25)
	α_W	0.58 (0.03)	0.81 (0.03)	0.64 (0.02)	0.65 (0.02)	0.65 (0.02)	0.65 (0.02)	0.65 (0.03)	0.63 (0.02)	0.62 (0.02)	0.65 (0.03)
	β_W	0.03 (0)	0.04 (0)	0.03 (0)	0.03 (0)	0.03 (0)	0.03 (0)	0.03 (0)	0.03 (0)	0.03 (0)	0.03 (0)
	α_B	0.44 (0.03)	0.71 (0.04)	0.49 (0.03)	0.5 (0.03)	0.51 (0.03)	0.51 (0.03)	0.5 (0.03)	0.48 (0.03)	0.48 (0.03)	0.51 (0.03)
	β_B	8.49 (0.83)	14.78 (1.24)	12.07 (1.07)	13.35 (1.12)	12.61 (1.08)	13.66 (1.13)	12.29 (1.01)	13.13 (1.12)	12.31 (1.11)	11 (1.06)
	α_G	0.46 (0.03)	0.74 (0.05)	0.51 (0.03)	0.52 (0.03)	0.53 (0.03)	0.52 (0.03)	0.51 (0.03)	0.5 (0.03)	0.49 (0.03)	0.52 (0.03)
	β_G	9.3 (0.96)	16.57 (1.44)	13.25 (1.22)	14.49 (1.21)	13.72 (1.23)	14.78 (1.28)	13.2 (1.12)	14.13 (1.17)	13.05 (1.22)	11.79 (1.11)
	α_P	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
	μ_{LN}	-4.42 (0.12)	-3.91 (0.08)	-4.5 (0.1)	-4.56 (0.1)	-4.42 (0.09)	-4.54 (0.1)	-4.47 (0.11)	-4.64 (0.1)	-4.6 (0.12)	-4.33 (0.11)
	σ_{LN}	2.17 (0.08)	1.43 (0.05)	2.1 (0.07)	2.14 (0.08)	1.99 (0.06)	2.04 (0.07)	2.31 (0.07)	2.34 (0.08)	2.45 (0.08)	2.21 (0.09)
$f(\rho_i)$	α_E	1.73 (0.09)	1.75 (0.08)	1.83 (0.09)	1.83 (0.09)	1.77 (0.08)	1.91 (0.09)	1.86 (0.08)	1.72 (0.08)	1.74 (0.09)	1.81 (0.09)
	α_W	0.76 (0.03)	0.71 (0.02)	0.63 (0.02)	0.64 (0.02)	0.62 (0.02)	0.6 (0.02)	0.6 (0.02)	0.64 (0.02)	0.67 (0.02)	0.69 (0.03)
	β_W	0.48 (0.04)	0.43 (0.03)	0.36 (0.03)	0.36 (0.03)	0.37 (0.03)	0.33 (0.03)	0.33 (0.03)	0.4 (0.03)	0.42 (0.03)	0.42 (0.03)
	α_G	0.69 (0.04)	0.62 (0.03)	0.51 (0.03)	0.52 (0.03)	0.5 (0.03)	0.47 (0.03)	0.47 (0.03)	0.51 (0.03)	0.56 (0.03)	0.57 (0.04)
	β_G	1.19 (0.1)	1.08 (0.09)	0.93 (0.08)	0.94 (0.08)	0.88 (0.07)	0.9 (0.08)	0.88 (0.08)	0.89 (0.07)	0.97 (0.09)	1.04 (0.09)
	α_P	0.02 (0)	0.02 (0)	0.02 (0)	0.02 (0)	0.02 (0)	0.02 (0)	0.02 (0)	0.02 (0)	0.02 (0)	0.02 (0)
	μ_{LN}	-1.43 (0.07)	-1.56 (0.07)	-1.85 (0.08)	-1.84 (0.08)	-1.84 (0.08)	-2 (0.09)	-1.98 (0.08)	-1.78 (0.09)	-1.67 (0.09)	-1.68 (0.09)
	σ_{LN}	1.4 (0.05)	1.47 (0.05)	1.69 (0.06)	1.67 (0.05)	1.82 (0.06)	1.84 (0.06)	1.85 (0.06)	1.84 (0.06)	1.77 (0.06)	1.88 (0.07)
$f(\kappa_j)$	α_B	5.33 (0.4)	5.42 (0.38)	5.85 (0.39)	5.6 (0.36)	5.36 (0.34)	6.09 (0.39)	5.81 (0.38)	4.66 (0.3)	5.63 (0.4)	4.99 (0.36)
	β_B	3.84 (0.29)	4.05 (0.28)	4.81 (0.32)	4.6 (0.29)	4.08 (0.26)	5.51 (0.35)	4.92 (0.32)	3.37 (0.22)	4.1 (0.28)	4.02 (0.29)
	α_G	12.52 (0.99)	12.61 (0.85)	13.09 (0.88)	12.33 (0.77)	11.98 (0.73)	13.35 (0.9)	12.71 (0.82)	10.34 (0.62)	12.89 (0.9)	10.6 (0.78)
	β_G	21.54 (1.75)	22.05 (1.52)	23.89 (1.63)	22.48 (1.43)	21.09 (1.31)	25.51 (1.75)	23.51 (1.56)	17.79 (1.09)	22.28 (1.6)	19.13 (1.46)
	μ_N	0.58 (0.01)	0.57 (0.01)	0.55 (0.01)	0.55 (0.01)	0.57 (0.01)	0.52 (0.01)	0.54 (0.01)	0.58 (0.01)	0.58 (0.01)	0.55 (0.01)
	σ_N	0.15 (0.01)	0.15 (0.01)	0.14 (0)	0.15 (0)	0.15 (0)	0.14 (0)	0.14 (0)	0.16 (0.01)	0.15 (0.01)	0.16 (0.01)

Table 2: Summary statistics of the estimated parameters for all entries a_{ij}, the diagonal elements a_{ii}, the column sums and row sums ρ_i and κ_j. The subscripts E, B, G, W, N, LN represent Exponential, Beta, Gamma, Weibull, Normal, Log-Normal distributions, respectively.

$f(.)$	Parameter	1963	1967	1972	1977	1982	1987	1992	1997	2002	2007
$f(\|\lambda_i\|)$	α_E	12.45 (0.64)	14.37 (0.67)	19.71 (0.91)	20.45 (0.9)	19.16 (0.87)	20.34 (0.94)	20.77 (0.99)	21.1 (1.05)	20.49 (1)	18.38 (0.96)
	α_W	0.99 (0.04)	0.84 (0.03)	0.63 (0.03)	0.54 (0.02)	0.57 (0.02)	0.47 (0.02)	0.43 (0.02)	0.48 (0.02)	0.68 (0.03)	0.43 (0.02)
	β_W	0.08 (0)	0.07 (0)	0.04 (0)	0.04 (0)	0.04 (0)	0.03 (0)	0.03 (0)	0.03 (0)	0.04 (0)	0.04 (0)
	α_B	0.94 (0.07)	0.67 (0.04)	0.47 (0.03)	0.38 (0.02)	0.41 (0.02)	0.33 (0.02)	0.29 (0.01)	0.33 (0.02)	0.51 (0.03)	0.3 (0.02)
	β_B	10.71 (0.96)	8.92 (0.72)	8.73 (0.7)	7.46 (0.68)	7.55 (0.63)	6.55 (0.59)	5.99 (0.56)	6.76 (0.65)	9.89 (0.9)	5.4 (0.58)
	α_G	1.02 (0.07)	0.7 (0.04)	0.48 (0.03)	0.39 (0.02)	0.42 (0.02)	0.33 (0.02)	0.29 (0.02)	0.34 (0.02)	0.53 (0.03)	0.3 (0.02)
	β_G	12.81 (1.11)	10.11 (0.85)	9.41 (0.83)	8 (0.72)	8.11 (0.69)	6.85 (0.67)	6.16 (0.61)	7.11 (0.68)	10.77 (1)	5.58 (0.6)
	α_P	0.02 (0)	0.02 (0)	0.02 (0)	0.02 (0)	0.02 (0)	0.02 (0)	0.02 (0)	0.02 (0)	0.02 (0)	0.02 (0)
	μ_{LN}	-3.09 (0.07)	-3.53 (0.17)	-4.32 (0.2)	-4.73 (0.27)	-4.5 (0.23)	-5.06 (0.31)	-5.4 (0.34)	-5.06 (0.32)	-4.22 (0.2)	-5.2 (0.4)
	σ_{LN}	1.25 (0.05)	3.57 (0.12)	4.23 (0.14)	5.83 (0.19)	5.18 (0.17)	6.4 (0.21)	7.31 (0.25)	6.62 (0.23)	4.12 (0.15)	7.57 (0.27)
$f\left(q^{\mathrm{R}}_{i,1}\right)$	α_E	47.6 (2.4)	59.95 (2.79)	65.48 (3.2)	81.42 (3.79)	112.46 (5.33)	62.34 (2.75)	75.13 (3.46)	66.88 (3.26)	53.62 (2.58)	63.91 (3.25)
	α_W	0.54 (0.02)	0.51 (0.02)	0.5 (0.02)	0.49 (0.02)	0.47 (0.02)	0.49 (0.02)	0.46 (0.02)	0.49 (0.02)	0.51 (0.02)	0.49 (0.02)
	β_W	0.01 (0)	0.01 (0)	0.01 (0)	0.01 (0)	0 (0)	0.01 (0)	0.01 (0)	0.01 (0)	0.01 (0)	0.01 (0)
	α_B	0.39 (0.02)	0.35 (0.02)	0.34 (0.02)	0.33 (0.02)	0.29 (0.01)	0.33 (0.02)	0.3 (0.02)	0.33 (0.02)	0.36 (0.02)	0.33 (0.02)
	β_B	17.4 (1.77)	19.96 (1.96)	20.97 (1.94)	23.87 (2.3)	26.9 (2.65)	19.5 (1.78)	20.86 (2.09)	20.69 (2.05)	18.07 (1.78)	18.64 (2)
	α_G	0.4 (0.02)	0.37 (0.02)	0.35 (0.02)	0.34 (0.02)	0.31 (0.02)	0.34 (0.02)	0.32 (0.02)	0.34 (0.02)	0.37 (0.02)	0.34 (0.02)
	β_G	19.45 (2.04)	21.92 (2.13)	23.27 (2.25)	27.9 (2.57)	34.63 (3.4)	21.43 (2.04)	23.88 (2.24)	22.98 (2.05)	20.02 (1.94)	22.19 (2.32)
	α_P	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
	μ_{LN}	-5.52 (0.1)	-5.92 (0.1)	-6.09 (0.11)	-6.39 (0.11)	-6.96 (0.12)	-6.1 (0.11)	-6.49 (0.11)	-6.18 (0.12)	-5.81 (0.12)	-6.12 (0.13)
	σ_{LN}	1.94 (0.07)	2.05 (0.07)	2.17 (0.07)	2.34 (0.08)	2.53 (0.08)	2.35 (0.08)	2.52 (0.08)	2.56 (0.08)	2.51 (0.09)	2.54 (0.09)
$f\left(q^{\mathrm{L}}_{1,j}\right)$	α_W	1.51 (0.06)	1.35 (0.05)	1.24 (0.05)	1.43 (0.05)	1.27 (0.04)	1.31 (0.05)	1.15 (0.04)	1.29 (0.05)	1.13 (0.05)	1.39 (0.06)
	β_W	0.06 (0)	0.08 (0)	0.05 (0)	0.08 (0)	0.05 (0)	0.06 (0)	0.05 (0)	0.14 (0.01)	0.15 (0.01)	0.19 (0.01)
	α_B	1.94 (0.13)	1.55 (0.09)	1.42 (0.09)	1.66 (0.09)	1.42 (0.08)	1.57 (0.1)	1.25 (0.07)	1.36 (0.08)	1.09 (0.06)	1.46 (0.09)
	β_B	31.71 (2.4)	21.02 (1.47)	31.7 (2.34)	22.69 (1.48)	30.93 (2.16)	27.56 (1.96)	24.93 (1.8)	9.23 (0.62)	6.54 (0.46)	6.83 (0.52)
	α_G	2.04 (0.14)	1.63 (0.11)	1.48 (0.08)	1.75 (0.1)	1.47 (0.09)	1.67 (0.1)	1.3 (0.07)	1.57 (0.09)	1.23 (0.08)	1.77 (0.12)
	β_G	35.3 (2.78)	23.78 (1.82)	34.7 (2.26)	25.75 (1.81)	33.45 (2.32)	31.08 (2.08)	27.4 (1.87)	12.36 (0.88)	8.71 (0.68)	10.11 (0.79)
	μ_{LN}	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)
	σ_{LN}	3.24 (0.12)	3.18 (0.1)	3.69 (0.13)	3.15 (0.1)	3.66 (0.12)	3.39 (0.11)	3.65 (0.12)	2.61 (0.09)	2.67 (0.1)	2.24 (0.08)

Table 3: Summary statistics of the estimated parameters for eigenvalues' moduli $|\lambda_i|$, and the elements of the left- and right-hand Perron-Frobenius eigenvectors, the eigenvectors associated to the maximal eigenvalue, $f q^{\mathrm{L}}_{1,j}$ and $q^{\mathrm{R}}_{i,1}$. The subscripts E, B, G, W, N, LN represent Exponential, Beta, Gamma, Weibull, Normal, Log-Normal distributions, respectively.