

Reconstructing firm-level input-output networks from partial information

Andrea Bacilieri^{1,2,*} and Pablo Austudillo-Estevez^{3,4}

¹Institute for New Economic Thinking, University of Oxford, Oxford, OX1 3UQ, UK

²Smith School of Enterprise and Environment, University of Oxford, Oxford, OX1 3QY, UK

³School of Geography and the Environment, University of Oxford, Oxford, OX1 3QY, UK

⁴Universidad San Francisco de Quito & School of Economics, Quito, 170902, Ecuador

*Corresponding author. *E-mail address:* andrea.bacilieri@sant.ox.ac.uk (A. Bacilieri).

Abstract

There is a large consensus on the fundamental role of firm-level supply chain networks in macroeconomics. However, data on supply chains at the fine-grained, firm level are scarce and frequently incomplete. For listed firms, some commercial datasets exist but only contain information about the existence of a trade relationship between two companies, not the value of the monetary transaction. We use a recently developed maximum entropy method to reconstruct the values of the transactions based on information about their existence and aggregate information disclosed by firms in financial statements. We test the method on the administrative dataset of Ecuador and reconstruct a commercial dataset (FactSet). We test the method's performance on the weights, the technical and allocation coefficients (microscale quantities), two measures of firms' systemic importance and GDP volatility. The method reconstructs the distribution of microscale quantities reasonably well but shows diverging results for the measures of firms' systemic importance. Due to the network structure of supply chains and the sampling process of firms and links, quantities relying on the number of customers firms have (out-degrees) are harder to reconstruct. We also reconstruct the input-output table of globally listed firms and merge it with a global input-output table at the sector level (the WIOD). Differences in accounting standards between national accounts and firms' financial statements significantly reduce the quality of the reconstruction.

Keywords: Network reconstruction, supply chain, production network, input-output table, maximum entropy, missing information

JEL codes: C80, D57, E32, L14, F12

1 Introduction

Recent events, such as the war in Ukraine and the evermore frequent natural disasters, have highlighted the fragility of global supply chains. Shocks to individual firms or a cluster of firms, sometimes located in a particular geographical region, can quickly spread through the network with severe repercussions on the global economy. Most of the research has so far been conducted at the sector level (Acemoglu et al., 2012; Carvalho, 2014; Pichler and Farmer, 2021). However, analysing such shocks at the level of sectors can lead to misleading results (Diem et al., 2022). The value of firm-level data is thus increasingly recognised, but research efforts are constrained by data availability.

Due to confidentiality and the data collection process, supply chain data are scarce, hard to access, and frequently incomplete. Countries collect the best data sources through VAT filings, but only a handful of countries collect them (for a comprehensive review and discussion of the different datasets Bacilieri et al., 2023). Some datasets of broader global coverage derived from US disclosure requirements are available. Given their wider breadth of coverage and relatively easier access, such datasets are used in many studies (e.g., Wu, 2016; Wang et al., 2021; Pankratz and Schiller, 2019; Taschereau-Dumouchel, 2020; Boehm and Sonntag, 2022; Barrot and Sauvagnat, 2016; Atalay et al., 2011; Herskovic et al., 2020). In contrast to national datasets, which report the monetary values of the transactions (i.e., the weighted network), global datasets do not provide this valuable information.

Reconstructing firm-level production networks is thus an important topic, showing growing research interest. The reconstruction problem concerns mainly two features of the production network: supplier-customer relations and transaction values. Several studies develop methods to infer both links and transaction values (Reisch et al., 2021; Ialongo et al., 2022; Hooijmaaijers and Buitenhuis, 2019; Hillman et al., 2021) or links only (Brinrup et al., 2018; Mungo et al., 2022; Kosasih and Brinrup, 2022), and two focus on inferring weights given the binary topology (Inoue and Todo, 2019; Welburn et al., 2020). The assessment of the quality of the weights reconstruction is usually missing and sometimes carried out on aggregate quantities or compared to empirical facts about another country's network.

In this paper, we focus on reconstructing the transaction values using partial information on the supply chain relations and aggregate information about firms' revenues and expenditures. We provide a rigorous assessment of a recently developed maximum entropy reconstruction method (Parisi et al., 2020) using the administrative dataset of Ecuador. We evaluate the method on microscale, higher-order and macroscale quantities that are widely used in economic input-output (I-O) models.

We assess the method's performance at recovering the link weights, and the technical and allocation coefficients (i.e., normalised weights). We also use two indicators of firms' systemic importance, the output multipliers and the influence vector, that are prominent in macroeconomic I-O models of shock propagation. While the reconstruction method reproduces the weights (normalised or not) rather poorly, it reconstructs their distributions reasonably well. In contrast, the reconstruction shows diverging results for higher-orders quantities: the output multipliers are in remarkable agreement with the empirical values, while the influence vector is overestimated. We then use a general equilibrium I-O model (Acemoglu et al., 2012) to assess how shocks to firms' total factor productivity (TFP) propagate through the network, ultimately affecting aggregate GDP fluctuations. We show that aggregate volatility is overestimated by the reconstruction method we employ.

Our results suggest that quantities relying more prominently on the number of customers firms have are more adversely affected by missing firms and links due to the structure of supply chain networks (Bacilieri et al., 2023) and the sampling process underpinning the observed firms and links. We also find that including a proxy node, to represent the rest of the economy that is not captured by the network, is of help in predicting microscale and higher-order quantities but not for predicting aggregate volatility.

An additional contribution we make in this paper is to construct the I-O table of globally listed firms using the dataset collected by FactSet. We merge FactSet with the World Input-Output Database (WIOD). Key challenges related to differences in accounting standards between national accounts and firms' financial statements prevent us from (1) merging the two datasets at the desired country-sector or even sector level, and (2) carrying out an accurate quantification (at the firm level) of the key variables making up an I-O table. We then reconstruct and compute weights, coefficients and higher-order quantities for FactSet as well. The inability to accurately quantify the variables of the I-O table at the firm level dramatically reduces the quality of the final dataset and thus of the reconstruction.

The remainder of the paper is organised as follows. In Section 2, we explain the notation and

define the production network at the firm level. In Section 3, we discuss the two datasets we use. In Section 4, we review the literature on network reconstruction, briefly describe the maximum entropy method we employ and the metrics we use to assess the performance of the reconstruction method. We then show and discuss the results for microscale, higher-order and macroscale properties for Ecuador and FactSet. Section 4.5 shows results for different numbers of unknown links and Section 5 concludes.

2 Firm-level input-output tables

This section describes firm-level production networks and gives an example of the supply chain network of publicly listed firms we aim to reconstruct. We then explain I-O tables and outline key differences between I-O tables at the sector and firm level.

The production or supply chain network is composed of N firms (nodes) and links between firms indicate yearly trading relationships. Links may be weighted, where each weight w_{ij} represents firm j 's intermediate input expenditure on goods produced by i . We label the weighted adjacency \mathbf{W} and \mathbf{A} the binary adjacency matrix. Figure 1 shows the binary (left) and weighted (right) adjacency matrices depicting the empirical data collected by FactSet. Both matrices have on the i -th row the customers of firm i , while column j lists the suppliers of the j -th firm. The unknown weights are labelled as question marks. Given \mathbf{A} (and other aggregate information about firms that we outline below), we aim to reconstruct \mathbf{W} .

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix} \quad \mathbf{W} = \begin{bmatrix} 0 & ? & 0 & ? \\ 0 & 0 & ? & 0 \\ ? & 0 & 0 & ? \\ ? & ? & 0 & 0 \end{bmatrix}$$

Figure 1: Example of the data we aim to reconstruct. **Left:** Binary directed adjacency matrix. **Right:** Weighted directed adjacency matrix.

For each firm, we can define its (total) intermediate expenditure and sales as, respectively, the column and row sums of the weighted adjacency matrix. The column and row sums are also called the weighted in- and out-degrees or in- and out-strengths; they are given by

$$\mathbf{s}^{\text{in}} = \mathbf{W}^\top \mathbf{1}, \text{ and} \tag{1}$$

$$\mathbf{s}^{\text{out}} = \mathbf{W} \mathbf{1}, \tag{2}$$

where $\mathbf{1}$ is a vector of ones of appropriate size.

The supply chain network just described captures only a part of the economic activity, namely firm-to-firm trades. Transactions with other economic actors (e.g., households) are captured in an input-output table, usually at the sector level. One can also define an input-output table at the firm level, however with notable differences. For a more in-depth discussion, we refer to Appendix A.3 and Bacilieri et al. (2023).

In I-O studies of production networks, the weighted adjacency matrix is usually normalised using firms' total costs instead of their in-strengths. A firm's total costs are the costs of intermediate inputs plus value-added, which is itself composed of labour costs, depreciation, amortisation and profit (see Appendix A.1.3). The weights so normalised are called the *technical coefficients* and represent the percentage of inputs firm j buys from firm i . The technical coefficients are given by

$$T_{ij} = \frac{W_{ij}}{\sum_i W_{ij} + y_j}, \tag{3}$$

where y_j is value-added of firm j .

Similar to the technical coefficients, one can define the *allocation coefficients*, \mathbf{B} . B_{ij} tells us the percentage of output firm i sells to firm j . Letting f_i be the amount of final demand satisfied by firm i , the allocation coefficient is defined as

$$B_{ij} = \frac{W_{ij}}{\sum_j W_{ij} + f_i}. \tag{4}$$

3 Data

FactSet is the global production network that we aim to reconstruct and for which we do not know the values of the monetary transactions. Therefore, we test the reconstruction method on the administrative dataset of Ecuador, for which we know the monetary values of the transactions.

3.1 FactSet

We use three primary data sources provided by FactSet: Fundamentals, Supply Chain Relationships and Supply Chain Shipping Transactions.¹ FactSet covers mainly listed firms around the world. The supply chain relationships of these companies are collected through two primary sources: company filings required by US Federal Accounting Standards (Supply Chain Relationships) and import and export declarations at ports from the US Bill of Lading (Supply Chain Shipping Transactions).² FactSet also collects information on supply chain relationships from investor presentations, company websites and press releases. Due to the nature of the data collection process, coverage is biased toward companies listed on US stock exchanges, large firms and large transactions. For a more detailed description of FactSet see Appendix A.1 and [Bacilieri et al. \(2023\)](#).

We aggregate customer-supplier relations within a fiscal year to ensure time consistency between the formation of supplier-customer relations and financial statements.³ We further aggregate all three datasets at the parent company level. For each company, we also have information on the sector (NACE Rev.2 codes at the 4-digit level) and the country where the company's headquarters are located.

We use several variables from companies' income statements (FactSet Fundamentals): revenues, the cost of goods sold, labour expenses, earnings before interest and taxes (EBIT), depreciation and amortisation. We convert all the variables to USD using the currency conversion tables provided by FactSet. We define value-added as the sum of labour expenses, EBIT, amortisation and depreciation (Appendix A.1.3). Some firms do not disclose their labour costs and include them in the costs of goods sold; we estimate these firms' labour expenses (see Appendix A.1.2).

For simplicity, we limit ourselves to the 2014 network, which we call henceforth "FactSet". We keep firms with positive sales, intermediate expenses and value-added, and with non-negative labour costs (see Appendix A.1.4). We exclude firms in financial and insurance, extraterritorial organisations and bodies and activities of households as employers. The number of firms in the 2014 cleaned dataset is 5,442; these are involved in 15,916 trading relations. The average degree is 2.9.

Evaluation of coverage and proxy node. In 2014, FactSet captures around 16.4% of world gross output as reported in the WIOD (see Appendix A.2.4). To capture the rest of the economic activity that we do not capture in FactSet, we introduce a "proxy" node in the network to which all firms are connected. We construct the proxy node's variables (gross output, intermediate sales and expenditure, etc.) using the WIOD aggregated at the world level.

A good approach would be to integrate FactSet with the WIOD at the country and sector level. However, due to differences between national accounting standards and firms' financial statements, we had to aggregate the WIOD at the world level. We refer to Appendix A.3 for a detailed discussion of how we integrated the two datasets.

3.2 Ecuador

Ecuador collects customer-supplier relations through VAT filings, which are mandatory for firms and natural persons. We do not have access to firms' financial statements, so we do not know their revenues, labour costs or profit, but we know the sectors' firms are in (ISIC Rev. 4 codes). The Ecuador dataset was provided by Ecuador's government to one of the authors. We refer to [Astudillo-Estevez \(2021\)](#) and [Bacilieri et al. \(2023\)](#) for more information about the dataset.

¹The datasets were downloaded in April 2020.

²The Statement of Financial Accounting Standards No. 131 requires publicly traded firms on US stock exchanges to report customers that account for 10% or more of their annual revenues, formally called *major customers*.

³The fiscal year goes from June to May, meaning that if a company's fiscal year end-month falls between January and May, the fiscal year is the current calendar year minus one; otherwise, it is the current calendar year.

Constructing the test network. While Ecuador’s dataset has comprehensive coverage, FactSet does not. Therefore, we construct a test network that mimics the missing firms and links in FactSet. To mimic the missing firms in FactSet, we keep the same number of firms in Ecuador that we have in FactSet. We choose to keep the largest firms (in terms of out-strength) since we observed predominantly large firms in FactSet. We also require firms to have positive in-strengths, meaning that firms need to buy some inputs from the other firms in the subgraph.⁴ The resulting network, which we call “test network” (third column in Table 1), has a much higher average degree compared to FactSet, meaning that the average Ecuadorian firm is connected to many more firms than the average firm in FactSet.

To mimic the missing links in FactSet, we eliminate links at random in the “test network” until we match FactSet’s average degree. Since FactSet’s supply chain relations mostly cover customers that account for 10% or more of a firm’s annual revenues, we delete links with a smaller weight with a higher probability: we set the link deletion probability to be inversely proportional to the link weight $p_{ij} \propto 1/W_{ij}$. We do this procedure 50 times and reconstruct each of the 50 randomised networks. The summary statistics for these networks, which we call “trimmed test network”, are shown in the last column of Table 1. In deleting links to match FactSet’s average degree, 96% of the links among firms in the test network are deleted. Consequently, our results can be interpreted as an approximate lower bound on the quality of the reconstruction.

We aggregate the firms and transactions left out of the test network in one proxy node representing the rest of the economy. As done for FactSet, we establish an incoming and outgoing link between each firm and the proxy node.

Summary statistics	Full network	Test network	Trimmed test network
N. nodes	84,978	5,440	5,440
N. edges	3,439,975	432,910	15,776
Average degree	40.5	79.6	2.9

Table 1: Summary statistics for the 2014 Ecuador network. The first column “Full network” is the network composed of all the firms except those in sectors: finance, insurance, activities of households as employers, activities of extraterritorial organisations and bodies, and those that have no sectoral code. The third column, “Test network”, refers to our test network, which is composed of the top 5,440 firms with the largest total intermediate sales and positive intermediate expenses. The last column “Trimmed test network” refers to the network built from the “Test network” by further deleting links at random to match FactSet’s average degree. The summary statistics of the test and trimmed test network do not include the proxy node.

Inferring missing data. For Ecuador, we do not have information on final demand, revenues and the variables that compose value-added (i.e., labour costs, depreciation, amortisation and profits). To carry out the analyses described in Section 4.2, we need final demand and value-added of each firm. Therefore, we simulate final demand and value-added using the 2014 I-O table of Ecuador at the sector level.⁵ Consider value-added (a similar procedure is done for final demand), for each sector s , we calculate the ratio of value-added to intermediate expenditure $\nu_s = y_s/s_s^{\text{in}}$. Assuming that a firm’s ratio is the same as that of the sector the firm is in, a firm’s value-added is given by $y_i = \nu_s \cdot s_i^{\text{in}}$.

4 Network reconstruction

Methods for reconstructing networks with missing information have mostly been developed for financial or trade networks (e.g., Moussa, 2011; Mastrandrea et al., 2014; Cimini et al., 2015b; Gandy and Veraart, 2017; Anand et al., 2015) and I-O tables at the sector level (e.g., Golan et al., 1994; Robinson et al., 2001; Lenzen et al., 2009). Only a few studies develop methods for reconstructing firm-level production networks (Inoue and Todo, 2019; Welburn et al., 2020; Reisch et al., 2021; Hooijmaaijers and Buiten, 2019; Ialongo et al., 2022; Hillman et al., 2021). We start with a brief overview of the different reconstruction methods developed in the literature, which we divide into deterministic and

⁴4 firms are further dropped because they do not have any links with the other firms in the sampled subgraph.

⁵The sector-level I-O table is available at <https://contenido.bce.fin.ec/documentos/PublicacionesNotas/Catalogo/CuentasNacionales/Anuales/Dolares/MenuMatrizInsumoProducto.htm>.

ensemble methods, and subsequently give a more detailed account of the methods developed to reconstruct firm-level networks. We refer to [Squartini et al. \(2018\)](#) and [Cimini et al. \(2021\)](#) for reviews on reconstruction methods developed mostly for financial and trade networks, and to [Miller and Blair \(2009\)](#), [McDougall \(1999\)](#) and [Lahr and De Mesnard \(2004\)](#) for reviews on sector-level reconstruction methods and matrix balancing problems.

The network reconstruction problem, being it for financial, trade or production networks, boils down to inferring a matrix of bilateral flows among entities (e.g., banks, countries or sectors) given constraints on the total in- and out-flows of each entity and other information when available (e.g., degrees or a prior bilateral flows matrix). Most of the reconstruction methods are based on the maximum entropy principle, of which there are two strands: deterministic and ensemble methods. Deterministic methods yield a single reconstruction of the weighted network while meeting the constraints exactly. Instead, ensemble methods sample many networks from a distribution that is constructed to respect the constraints on average. Therefore, while ensemble methods generate a probability distribution over the likely networks, deterministic methods assign a probability of one to the reconstructed network and a zero probability to all the other networks – which likely include the true network ([Parisi et al., 2020](#)). The shortcomings that most deterministic and ensemble methods share are that they tend to create a fully connected network with weights distributed as uniformly as possible given the imposed constraints on the in- and out-flows. To create sparser networks, algorithms with tunable ([Moussa, 2011](#); [Upper, 2011](#); [Mastromatteo et al., 2012](#)) or exact network density ([Mastrandrea et al., 2014](#); [Cimini et al., 2015b](#)) have been developed.

The most well-known deterministic method is known as *MaxEnt*. It maximises an entropy-like functional subject to constraints on the in- and out-strength of each node. The solution to this maximisation yields the well-known gravity model (without distance) in the international trade literature (first proposed by [Tinbergen, 1962](#) and [Pöyhönen, 1963](#); see also [Squartini and Garlaschelli, 2014](#) for a discussion). MaxEnt displays all the shortcomings mentioned above: it generates a fully connected network and the weights are distributed as equally as possible given the constraints. To enhance the MaxEnt reconstruction, if some prior information about the binary topology or the weights is available, it can be integrated using a cross-entropy method ([Golan et al., 1994](#); [Di Gangi et al., 2018](#); [Upper, 2011](#); [Wells, 2004](#)). The cross-entropy method reconstructs a network that has minimum distance to the prior while accounting for the imposed constraints. The cross-entropy method is equivalent to the iterative proportional fitting (IPF) algorithm, which iteratively distributes the weights (coming from the MaxEnt solution or any other prior) among the non-zero edges until the row and column sums are satisfied. The IPF algorithm is also known as the RAS technique in the I-O literature ([Miller and Blair, 2009](#)). If the network is fully connected, the IPF algorithm is equivalent to MaxEnt.

There are different ensemble methods depending on the information used for the reconstruction (e.g., in- and out-degrees or strengths sequences). We discuss the method developed by [Cimini et al. \(2015b\)](#) since it is one of the best performing ([Anand et al., 2018](#); [Lebacher et al., 2021](#)). To enhance the reconstruction of the weighted network, [Cimini et al. \(2015b\)](#) impose constraints on both the in- and out-strength sequences and on the degrees. Their strategy is motivated by recent results showing that the strengths do not encode information about the binary topology (although they are correlated with degrees) and that the degrees are “fundamental” local structural properties of weighted networks ([Mastrandrea et al., 2014](#)). Combining constraints on the degrees and strengths thus greatly enhances the reconstruction of weighted networks since the degrees provide information about the binary topology that strengths do not, enabling to identify better the matrix of link probabilities as well as higher-order properties ([Mastrandrea et al., 2014](#); [Gandy and Veraart, 2017](#)).

The method proposed by [Mastrandrea et al. \(2014\)](#) requires knowledge of the degrees and strengths of all the nodes, which are not always available. [Cimini et al. \(2015b\)](#) note that in financial networks, the in- and out-strengths are usually known while the degrees might be known for a subset of nodes only. To account for these two pieces of information, [Cimini et al. \(2015b\)](#) restore to the *fitness* ansatz. The fitnesses are nodes’ non-topological futures that relate to the ability of nodes to establish connections: nodes with higher fitness attract more connections and are thus likely to become hubs ([Squartini et al., 2018](#); [Mazzarisi and Lillo, 2017](#)). Given the empirical correlation frequently observed among strengths and degrees, strengths are often used as a proxy for nodes’ fitnesses. To estimate the binary topology, [Cimini et al. \(2015b\)](#) thus develop a fitness-induced configuration model that estimates link probabilities using the nodes’ fitnesses and the degrees of only a few nodes. The weights are estimated in a second step using a degree-corrected gravity model that accounts for the sparsity

of the adjacency matrix inferred in the first step.

How does one choose which reconstruction method to use? [Anand et al. \(2018\)](#), [Lebacher et al. \(2021\)](#) and [Ramadiah et al. \(2020\)](#) find that the choice of the reconstruction method ultimately depends on the feature of the network one aims to reconstruct, which usually boils down to connectivity structure versus weights. If one cares about inferring links, methods that focus on reconstructing a sparse connectivity structure are better suited. [Anand et al. \(2018\)](#), [Lebacher et al. \(2021\)](#) and [Mazarisi and Lillo \(2017\)](#) conclude that the best-performing ones are the degree-corrected gravity model ([Cimini et al., 2015b](#)), the minimum density ([Anand et al., 2015](#)) and the Bayesian hierarchical fitness ([Gandy and Veraart, 2017](#)). If one cares about inferring the link weights, methods based on MaxEnt or the IPF algorithm perform best. Since financial networks have many link weights of relatively equal size (something that is less likely to be the case in firm-level networks; [Bacilieri et al., 2023](#)), they score well on weight-based similarity measures ([Anand et al., 2015](#)). In their horse races, [Anand et al. \(2018\)](#) and [Lebacher et al. \(2021\)](#) identify the methods developed by [Cimini et al. \(2015b\)](#) and [Baral and Fique \(2012\)](#) to be the best performing. Altogether, the reconstruction method proposed by [Cimini et al. \(2015b\)](#) seems to be the best in reconstructing both binary and weighted topological features.

Firm-level network reconstruction. [Welburn et al. \(2020\)](#) reconstruct the network of listed firms in the US available through a dataset similar to FactSet but covering only their major customers.⁶ Therefore, they know the binary topology only partially. To reconstruct the weighted network, they use a two-step procedure. In the first step, they infer missing links using a logistic regression. In the second step, they develop a linear programming method to reconstruct the unknown weights given the links. Since they do not observe the whole economy, they introduce a proxy node that captures the rest of the economy and to which each firm is linked. The cumulative in- and out-flows of the proxy node are then minimised subject to constraints on firms' revenues and the cost of goods sold. While there are almost 6,000 firms in their network, they only reconstruct the network composed of 1,000 firms due to the high computational complexity of the second step of their procedure. [Inoue and Todo \(2019\)](#) reconstruct the weighted network of Japanese firms given the binary topology.⁷ Firstly, they assume that the link weight is proportional to the supplier and customer's sales. Subsequently, they re-adjust the estimated weights using the sector-level I-O tables to ensure that if the firm-level network is aggregated at the sector level, it is consistent with national accounts. In a similar fashion and using the same Japanese dataset, [Carvalho et al. \(2021\)](#) assume that the technical coefficient between two firms is proportional to the technical coefficient between the sectors those two firms are in.

[Hillman et al. \(2021\)](#) reconstruct the global network of private and public firms in the ORBIS database. As done in other reconstruction methods, they build the firm-level network so that it is consistent with sector-level I-O data (OECD). In each step, a firm i is chosen at random and its sales are split into n units that are then sold to n different customers. i 's customers are chosen according to the industry they are in based on sector-level I-O tables, meaning that if firm i is in industry s , its customers need to be in one of the industries to which s sells (in the sector-level I-O table). They further calibrate their model on the observation that larger firms tend to have more customers ([Bernard et al., 2019](#)). This feature can be controlled by changing how, at the beginning of each step, firm i 's sales are split into n units. For computational reasons, they only use 5,000 firms among the more than 200 million firms in the Orbis database. [Hooijmaaijers and Buitenhuis \(2019\)](#) reconstruct the supply chain network of the Netherlands using several microdata sources available to the Office of National Statistics. They describe their method as being akin to maximum entropy methods with exact link density (as classified in [Squartini et al., 2018](#)), but they pose additional constraints thanks to the richness of their microdata and to findings in the literature about empirical facts of firm-level supply chain networks. A novel feature of their reconstruction is the disaggregation of firms' output into different goods.

None of the studies just described can assess how well their method recovers the empirical weighted network because none of them has access to it. Two studies assess their reconstruction method, at least to some extent. First, [Reisch et al. \(2021\)](#) reconstruct a firm-level production network using mobile communication data. To guarantee anonymity, the company providing the data and the country are

⁶They retrieved the customer-supplier relations from firms' filings available through the EDGAR database.

⁷The network is collected by a private company. Although the coverage is extensive, comprising almost 890,000 firms, it is not exhaustive.

not disclosed. Roughly speaking, links are inferred by assuming that if two firms communicate with each other, they are involved in a supply-chain relationship. To determine the link direction, they use the national I-O table at the sector level and information about the sectors the customer and supplier are in. A gravity model is then used to estimate the link weights, where a firm's size is given by its total assets. To assess the reconstruction of the binary topology, they compare to the Hungarian network, whereas to assess the performance of the method regarding the weights reconstruction, they use the Economic Systemic Risk Index and compare with results obtained for Hungary by [Diem et al. \(2022\)](#). They find similarities between results obtained for Hungary and their reconstructed network. Second, [Ialongo et al. \(2022\)](#) develop the stripe-corrected gravity model, which builds on the degree-corrected gravity model ([Cimini et al., 2015a](#)) by adding constraints on the in-strength of each industry. They test their method on two transaction data made available by two Dutch banks. The assessment is carried out on the degree and strength distributions, degree-strength correlations and average nearest neighbour strength.

4.1 Method

To reconstruct the weighted network given the binary topology, we use the conditional maximum entropy ensemble reconstruction method developed by [Parisi et al. \(2020\)](#). We do not use any of the previously developed methods for reconstructing firm-level networks because either they are too computationally expensive ([Welburn et al., 2020; Hillman et al., 2021](#)), demand too many data inputs that we do not have ([Hooijmaaijers and Buitenhuis, 2019](#)) or would imply constraining the firm-level network with sector-level I-O tables ([Inoue and Todo, 2019; Ialongo et al., 2022](#)). We disregard the latter methodologies because we think using sector-level data in the way proposed by [Inoue and Todo \(2019\)](#) or in the spirit of [Ialongo et al. \(2022\)](#) could bias the reconstruction in unwanted ways given the underlying differences in accounting standards between national I-O tables and firms' financial statements (see Appendix A.3, but also [Bacilieri et al., 2023](#)).

[Parisi et al. \(2020\)](#) develop a maximum entropy method that reconstructs an ensemble of likely weighted networks given some prior information about the binary ensemble and aggregate information about each node. The procedure is flexible in that it allows to use of an observed binary topology or to infer it in a previous step and account for the additional uncertainty. They proposed two methods that use different constraints. One method constrains the in- and out-strength sequences, while the other one constrains the expected link weights (and the in- and out-strengths indirectly). We choose the second method for two reasons. First, it is computationally more efficient since it involves solving m (the number of links) decoupled equations, while the method constraining the in- and out-strengths require solving $2N$ coupled equations, where N is the number of firms. Second, the authors show that it predicts the weights better compared to the model constraining the in- and out-strengths.

The method consists of two steps. The first step constrains the in- and out-strengths (total intermediate expenditure and sales) of each firm and determines the values of the expected link weights used as constraints in the second step. As discussed in Section 4, MaxEnt reconstructs the weights best among the other methods ([Anand et al., 2018; Lebacher et al., 2021](#)); thus, MaxEnt is used in the first step. The second step allows us to account for any prior information about the binary topology and generates an ensemble of weighted networks compatible with this information and with the constraints on the expected link weights.

As discussed, MaxEnt assumes a fully connected network; however, the conditional maximum entropy method assumes no self-loops. Additionally, in our case, we know the binary topology. To redistribute the weights corresponding to $A_{ij} = 0, \forall (i, j) \notin \mathcal{E}$, where \mathcal{E} is the edge set, we employ the IPF algorithm. The IPF algorithm redistributes the weights in an iterative procedure until the constraints on the in- and out-strengths are met (see Appendix B.1). [Parisi et al.'s \(2020\)](#) method turns the IPF algorithm into a probabilistic method and allows calculating confidence intervals around each reconstructed weight. We give a brief outline of the method below and describe it in detail in Appendix B.1.

First step. The method derives values for the expected link weights to enforce as constraints in the second step. The values of the weights are derived by solving the MaxEnt problem, which maximises an entropy-like functional subject to constraints on intermediate sales and costs of each firm. The

value of each weight is given by

$$W_{ij}^{\text{ME}} = \frac{s_i^{\text{out}*} s_j^{\text{in}*}}{W^{\text{tot}*}},$$

where $W^{\text{tot}*} = \sum_i s_i^{\text{out}*} = \sum_j s_j^{\text{in}*}$ is the total weight of the empirical network.

Second step. The method maximises the conditional entropy defined over the probability density function of the weighted networks compatible with the prior on the binary ensemble and subject to constraints on the expected weights. Solving the conditional maximum entropy problem yields that the probability of observing a weight $W_{ij} > 0$ given that there is a link between i and j is of exponential form with parameter λ_{ij} (which also corresponds to the Lagrange multiplier):

$$Q_{ij}(W_{ij} | A_{ij} = 1) = \lambda_{ij} e^{-\lambda_{ij} W_{ij}}, \quad W_{ij} > 0. \quad (5)$$

To find the values of the λ_{ij} 's, one maximises the log-likelihood function, which leads to the first order conditions

$$\langle W_{ij} \rangle = \frac{p_{ij}}{\lambda_{ij}}, \quad \forall i \neq j. \quad (6)$$

Since for each link the expected weight $\langle W_{ij} \rangle = W_{ij}^{\text{ME}}$, the Lagrange multipliers are given by

$$\lambda_{ij}^* = p_{ij} \frac{W^{\text{tot}*}}{s_i^{\text{out}*} s_j^{\text{in}*}}, \quad \forall i \neq j. \quad (7)$$

We set $p_{ij} = 1, \forall (i, j) \in \mathcal{E}$.

Confidence interval on the expected edge weight. For each expected weight, the confidence interval is $[w^-, w^+]$. The lower bound is given by

$$w^- = -\frac{\ln[e^{-1} + q^-]}{\lambda_{ij}^*}, \quad (8)$$

where q^- is a desired confidence level and the upper bound is given by

$$w^+ = -\frac{\ln[e^{-1} - q^+]}{\lambda_{ij}^*}. \quad (9)$$

We set $q^+ = q^- = 0.25$. We refer to Appendix E in [Parisi et al. \(2020\)](#) for the derivation.

4.2 Assessing the reconstruction

We assess how well the reconstruction method can recover the empirical network at three scales. First, we look at microscale quantities: weights, and technical and allocation coefficients. Second, we evaluate the reconstruction of higher-order properties (i.e., multipliers): the output multipliers and the influence vector. Third, we turn to macroscale properties and use a general equilibrium I-O model to study how the propagation of shocks through the network affects GDP volatility. We conclude this section by defining the statistical indicators used to compare the empirical and reconstructed quantities.

4.2.1 Weights

As discussed in Section 2, technical and allocation coefficients (Equation 3 and 4, respectively) are normalised weights frequently used in economic models. Therefore, we assess the reconstruction of both coefficients and the weights: their exact values and their distributions.

4.2.2 Higher-order properties

We look at two higher-order properties widespread in the economic literature: the output multipliers and the influence vector. These are two centrality measures that quantify firms' contributions to economy-wide fluctuations.

The output multipliers. The output multipliers are derived from the Leontief model (Miller and Blair, 2009) and are defined as

$$\mathcal{O} \equiv (\mathbf{I} - \mathbf{T}^\top)^{-1} \mathbf{1}, \quad (10)$$

where \mathbf{I} is the identity matrix and $\mathbf{1}$ a vector of ones, both of appropriate size. The output multiplier captures the upstream propagation channel of an exogenous shock to a firm's final demand and its economy-wide impacts (where final demand increases by one monetary unit; Miller and Blair, 2009). It can also be seen as the average length of a firm's production chain (McNerney et al., 2022; Fally, 2012; Miller and Temurshoev, 2017). The higher the output multiplier is, the longer the production chain is on average and the greater the impact of a change in a firm's final demand is on the whole economy.

The influence vector. The influence vector is derived from the Cobb-Douglas model proposed by Acemoglu et al. (2012). The influence vector is defined as

$$\mathbf{v} \equiv \frac{\alpha}{N} [\mathbf{I} - (1 - \alpha)\mathbf{\Omega}]^{-1} \mathbf{1}, \quad (11)$$

where $\mathbf{\Omega}$ is the matrix of input shares with $\omega_{ij} \in [0, 1]$ being the share of input i used in j 's production process ($\sum_i \omega_{ij} = 1$), $\alpha \in (0, 1]$ is the share of labour and N is the number of firms.

Contrary to the output multipliers, the influence vector captures the downstream propagation channel of TFP shocks and it gauges the contribution of firms to fluctuations in aggregate GDP. Positive TFP shocks can be thought of as firms' innovating their production processes and becoming more efficient in using their inputs. The influence vector is equivalent to a Reverse Weighted PageRank with a damping factor equal to $(1 - \alpha)$.

4.2.3 Macroscale properties

As highlighted by recent crises and natural disasters, assessing how different shocks affect the economy is of paramount importance to be better prepared in preventing or alleviating crises. Therefore, after assessing the weights and the multipliers, we study how supply-side shocks at the firm level propagate through the network to downstream firms, ultimately leading to fluctuations in aggregate GDP.

Supply-side shocks and aggregate volatility. We model supply shocks as shocks to firms' TFP. We use the model developed by Acemoglu et al. (2012) and refer to the paper for a derivation. In the competitive equilibrium, aggregate volatility is given by

$$\sigma_{\Delta y} = \sqrt{\sum_i \text{Var}(\Delta \epsilon_i) v_i^2}, \quad (12)$$

where v_i is the influence of firms i (Equation 11), $\Delta \epsilon_i$ is the TFP shock of firm i and ϵ_i is an i.i.d. random variable with mean zero and bounded variance. Equation 12 shows that productivity shocks at the firm level affect aggregate value-added through the production network. Initially, the shock propagates to the customers of the affected firm and, subsequently, propagates downstream to the customers' customers and so on, potentially spreading through the whole supply chain network.

Note that the model gives a static representation of the economy in equilibrium and firms' input shares are exogenous; therefore, the network structure and hence the influence vector are constant over time. We use Equation 12 to assess the impact of firms' TFP shocks on GDP volatility using either the empirical or the reconstructed weighted network given the TFP shocks.

Simulating TFP shocks. To estimate TFP shocks, we cannot use an econometric technique (e.g., Magerman et al., 2016) because it requires knowledge of several variables that we do not observe for Ecuador. Besides requiring all the variables describing a firm's production function, the estimation of TFP requires a time series of these variables. Since our goal is not to empirically validate the economic model but to assess the discrepancy in the predicted GDP volatility when the reconstructed influence vector is used instead of the empirical one, we simulate firm-level TFP shocks. For FactSet, we do not perform this test since it would entail estimating TFP, which is outside of the scope of this paper.

We simulate TFP shocks from a normal distribution with mean zero and standard deviation of 6. We chose a zero mean in line with the empirical mean reported for the Belgian production network by Magerman et al. (2016). Since they do not report the variance, we set the variance to 6 so that GDP volatility, calculated using the model with the true network, matches the observed country-level volatility.⁸ We set the TFP shock of the proxy sector equal to the median of the TFP shocks of firms that were excluded from our test network. We simulate 10 time periods. Figure 2 shows the distribution of the simulated TFP growth rates in Panel (a) and the distribution of their standard deviation in Panel (b). Once we simulate the TFP volatilities, we use Equation 12 to predict the fluctuation in aggregate GDP using either the empirical or the reconstructed influence vector. We then compare the empirical volatility with that predicted by the reconstruction.

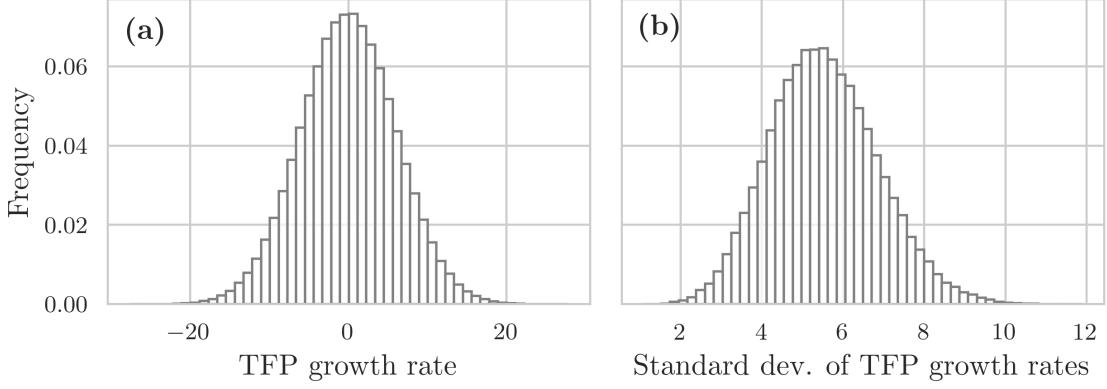


Figure 2: Distribution of TFP (a) growth rates and (b) their standard deviation. The growth rates are pooled over the 10 simulated years. We binned the data into 50 equally-spaced bins.

4.2.4 Statistical indicators

To compare how well the reconstruction method can recover the quantities defined in Section 4.2.1, 4.2.2 and 4.2.3, we employ metrics that are standard in the literature: the L_1 -error, the root-mean-squared error (RMSE), the mean and median absolute error (MAE and MedAE, respectively) and the cosine similarity.

The L_1 -error assesses the degree to which constraints on intermediate sales and costs are violated. It is defined as

$$L_1 = \sum_i |s_i^{\text{in}} - s_i^{\text{in}*}| + \sum_i |s_i^{\text{out}} - s_i^{\text{out}*}|,$$

s_i^{in} is firm i 's in-strength in the reconstructed network and s_i^{out} its out-strength in the reconstructed network; quantities with a * refer to observed, empirical values.

In what follows, we define the error measures using the technical coefficients, but they similarly apply to any other quantity of interest. We do not use the RMSE, the MAE and the MedAE to assess the raw weights because their distribution has heavy tails. For the technical and allocation coefficients, and the multipliers, we further normalise the metrics to allow their comparison across variables that have different scales. We rescale by ϕ , the difference between the maximum and minimum value (excluding the zeros) of the empirical quantity of interest. Our rescaled measures compare the variation in the residuals to the range of the empirical data. For instance, a normalised RMSE of 0.1 means that the variation in the residuals is 10% of the range of variation of the empirical data. The lower the normalised error metric is, the better the reconstruction is. The normalised root-mean-square error is given by

$$\text{RMSE} = \frac{1}{\phi} \sqrt{\frac{1}{m} \sum_{ij} (T_{ij} - T_{ij}^*)^2},$$

⁸To calculate the GDP volatility in Ecuador, we use data from the IMF; available at <https://data.imf.org/?sk=388DFA60-1D26-4ADE-B505-A05A558D9A42&sId=1479331931186>. We use nominal GDP since our data are not adjusted for inflation. Since we simulate 10 years, we calculate GDP volatility for the period 2005-2015, which is 6.35%.

We tried different parametrisations of the normal distribution, which do not match the volatility in the empirical data, and results do not change, at least qualitatively.

where m is the number of links. The normalised mean absolute error is given by

$$\text{MAE} = \frac{1}{\phi m} \sum_{i,j} |T_{ij} - T_{ij}^*| .$$

The normalised median absolute error is defined as

$$\text{MedAE} = \frac{\text{Median}(|\mathbf{T} - \mathbf{T}^*|)}{\phi} .$$

The cosine similarity is defined as

$$\vartheta = \frac{\sum_{ij} T_{ij}^* T_{ij}}{\sqrt{\sum_{ij} T_{ij}^{*2}} \sqrt{\sum_{ij} T_{ij}^2}} .$$

4.3 Results: link weights

In this section, we show the results of the reconstruction method for the weights, and the technical and allocation coefficients. We start by discussing the results for our test network, Ecuador, and then show the results for FactSet, for which we do not know the ground truth.

We compare weights (normalised or not) for firms only and not those of the proxy node since the proxy node can be thought of as a sink node and does not meaningfully represent either a firm or a sector. For parsimony, we show plots for one of the 50 randomised reconstructions (always the same one throughout the paper) since they all yield virtually identical results; the same holds for the summary statistics. Regarding the statistical indicators described in Section 4.2, we compute them for each of the 50 randomised test networks and report the average value of each metric across the 50 randomised networks.

4.3.1 Ecuador

The constraints on the intermediate sales and costs are always satisfied (L_1 -error = 10^{-4}). The reconstruction of individual weights is rather poor, as shown in Figure 3a; perfect prediction is achieved when points lie on the 45-degree line (grey dashed line). The reconstruction method tends to under-predict weights of high values and overpredict weights with intermediate or low values, although there is significant dispersion. (We show the histogram of the relative prediction errors and the empirical weights against their prediction errors in Figure C.2.) On average, 47% of the weights fall in the 50% confidence interval (CI).

Although the reconstruction cannot recover individual weights particularly well, the weight distribution is recovered quite well. The weight distribution has heavy tails in both the empirical and the reconstructed networks (Figure 3b; see Figure C.1a for a comparison of the weight distribution across the 50 randomised networks, the empirical test network and the full network). As expected from maximum entropy methods, the expected weight distribution is less heterogeneous than the empirical distribution. [Bacilieri et al. \(2023\)](#) find that the weight distribution is likely to follow a power-law with an exponent that varies between 1.0 and 1.4 depending on the country, year and estimation method. Therefore, we check whether we can recover a similar power-law exponent. We can recover it pretty well (see Figure C.1b). The power-law exponent is 1.1 for the empirical weight distribution and 1.3 for the reconstructed one.⁹ Although the exponent of the reconstructed distribution is slightly higher, it still implies a divergent second moment.

⁹To fit a power-law distribution to our data, we use the method of [Clauset et al. \(2009\)](#) since (1) it yields a single exponent estimate and (2) it is the most widely used estimator in the literature. See [Bacilieri et al. \(2023\)](#) for an in-depth discussion. [Bacilieri et al. \(2023\)](#) use also the estimators developed by [Voitalov et al. \(2019\)](#), which are based on extreme value theory. We abstain from such an analysis in this paper.

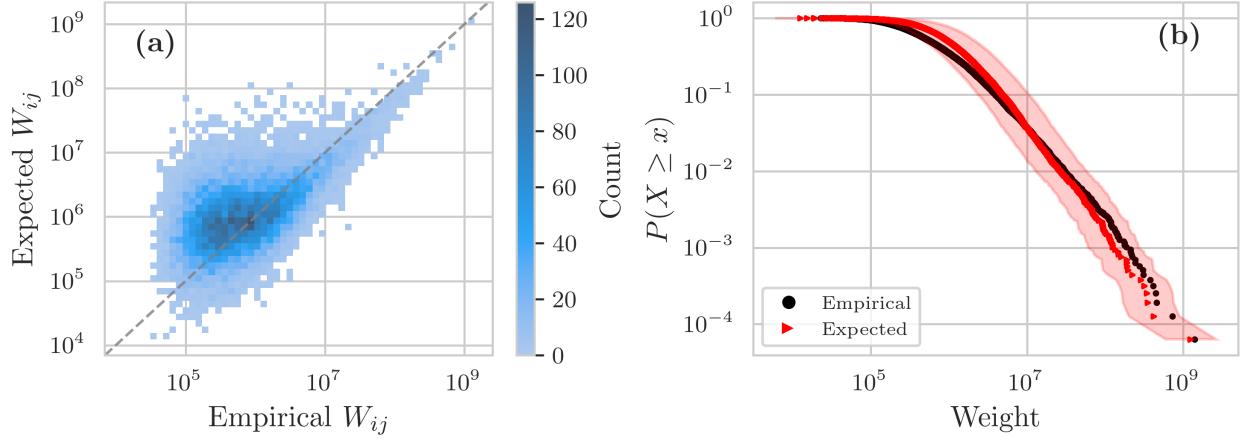


Figure 3: (a) 2D histogram of the empirical (x -axis) and expected (y -axis) weights for Ecuador. We divide both axes into 50 log-spaced bins and then count the number of data points falling in each square. Perfect prediction is achieved when points lie on the identity line (dashed grey line). (b) CCDF of the empirical (black dots) and expected (red triangles) weights for Ecuador. The shaded area represents the 50% confidence bounds. We compute the CCDF as $\bar{F}_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}(X_i \geq x)$, where $\mathbf{1}$ is the indicator function. Values are in USD.

In Table 2, we do not report the RMSE, MAE or MedAE for the weights since there is too much variation to obtain a meaningful metric; additionally, the weights likely have a diverging second moment. Therefore, we report only the cosine similarity, which is 0.93.

Type	RMSE	MAE	MedAE	Cosine similarity
Weight	—	—	—	0.928 (0.006)
Technical	0.081 (0.001)	0.041 (0.000)	0.013 (0.000)	0.723 (0.004)
Allocation	0.105 (0.001)	0.054 (0.000)	0.019 (0.000)	0.758 (0.003)

Table 2: Statistical indicators for the weights, and the technical and allocation coefficients for Ecuador. As defined in the main text, RMSE denotes the root mean squared error, MAE the mean absolute error and MedAE the median absolute error. For each metric, we show its mean value across the 50 randomised reconstructions. Below the mean value, the standard deviation in parenthesis. We excluded the proxy node from the calculations.

Figure 4a and 4c show, respectively, the empirical technical and allocation coefficients on the x -axis and their expected values on the y -axis. As for the weights, the reconstruction method does not perform particularly well in recovering either of the coefficients. Although less pronounced for the coefficients than for the weights, the method tends to overpredict coefficients of small values and underpredict coefficients with high values, again with considerable dispersion. This is further highlighted in Figure 4b and 4d, showing the empirical and reconstructed CCDF of the technical and allocation coefficients, respectively. Figure 4b and 4d further show that we can reconstruct the technical coefficients better, which also have smaller error metrics (Table 2) and for which we can recover the moments more accurately (Table 3). However, the cosine similarity is slightly higher for the allocation coefficients: 0.76 compared to 0.72.

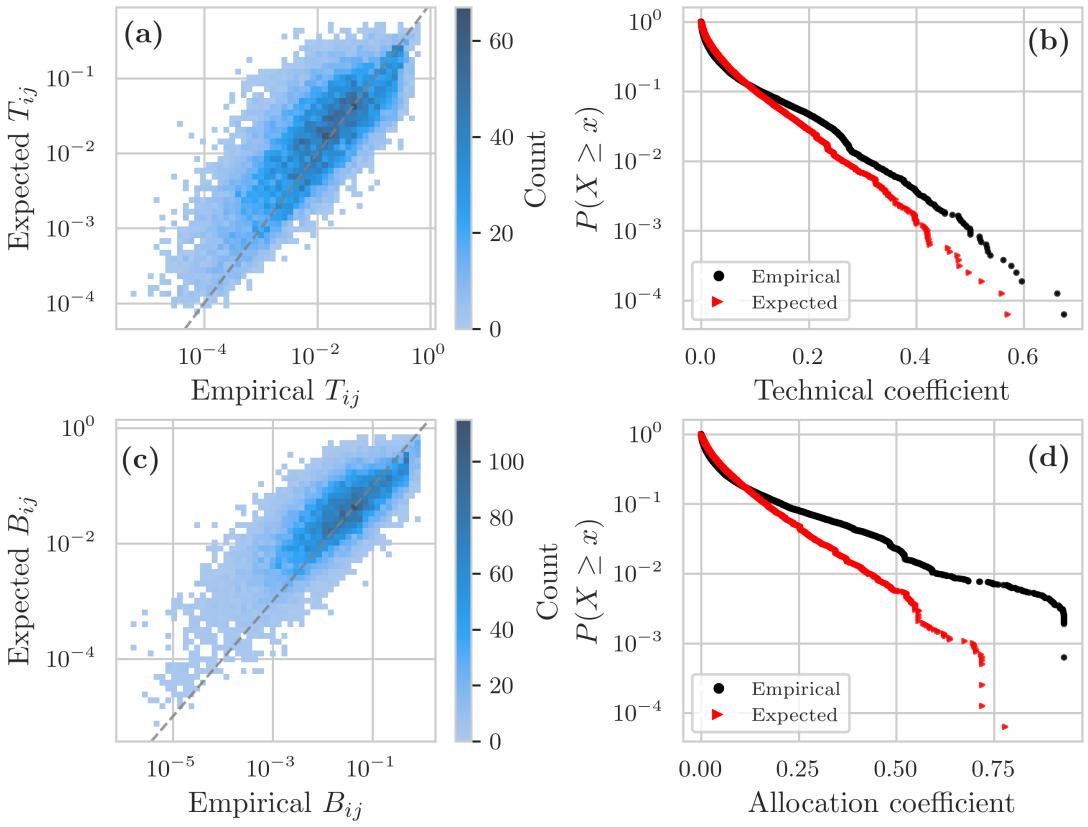


Figure 4: (a), (c) 2D histogram of the empirical (x -axis) and the expected (y -axis) technical and allocation coefficients, respectively, for Ecuador. We use 50 log-spaced bins and then count the number of data points falling in each square. (b), (d) CCDF of the empirical (black dots) and expected (red triangles) technical and allocation coefficients, respectively, in semi-log scale for Ecuador.

	Technical coefficients		Allocation coefficients	
	Empirical	Expected	Empirical	Expected
Mean	0.037	0.038	0.071	0.065
Median	0.010	0.016	0.019	0.031
Standard dev.	0.067	0.057	0.132	0.088

Table 3: Summary statistics of the technical and allocation coefficients for Ecuador. For each coefficient, the first column reports summary statistics for the empirical coefficients and the second column for the reconstructed ones. We show results for one of the reconstructions; all three quantities have virtually the same summary statistics across the 50 randomised empirical and reconstructed networks. We excluded the proxy node from the calculations.

4.3.2 FactSet

The constraints on the total intermediate expenditure and sales are satisfied (L_1 -error = 6×10^{-4}). Figure 5a shows the reconstructed weight distribution, which visually appears to have heavy tails. As done for Ecuador, we fit a power-law distribution. The estimated power-law exponent is lower than that recovered for Ecuador (1.0 for FactSet and 1.3 for Ecuador, see Figure C.3) and on the lower end of what is observed for other supply chain networks (Bacilieri et al., 2023). The reconstructed technical coefficients (Figure 5b) and allocation coefficients (Figure 5c) have a narrow range of variation and tend to be much smaller than those reconstructed for Ecuador. For FactSet, the technical coefficients reach a maximum value of approximately 0.008, while for the expected Ecuadorian network, they can be as high as 0.570. Similarly, the allocation coefficients are not higher than ~ 0.007 in FactSet, while in Ecuador, they reach a maximum value of 0.778.

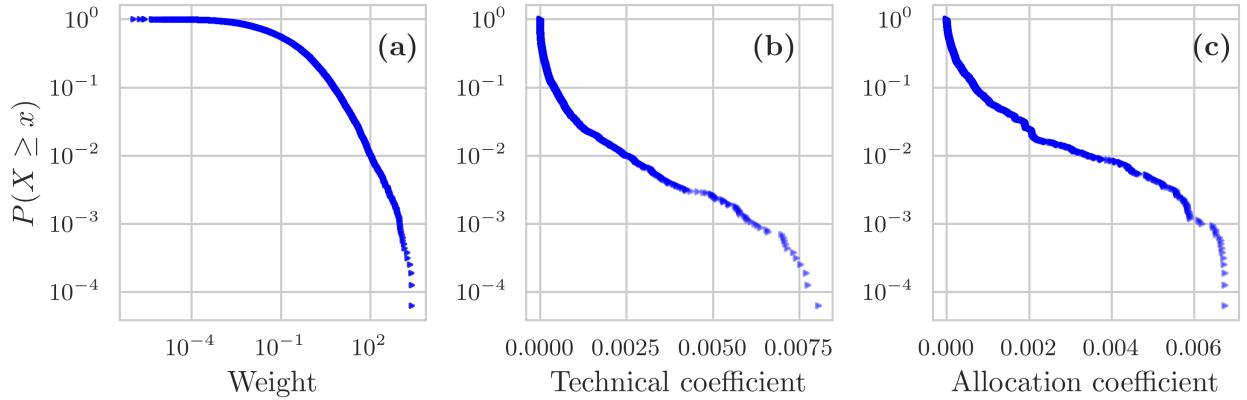


Figure 5: (a) CCDF of the expected weights, values in thousand USD. (b), (c) CCDF of, respectively, the expected technical and allocation coefficients in semi-log scale for FactSet.

Given the data collection method of customer-supplier relations in FactSet, there is a bias towards observing links with customers that account for 10% or more of a firm’s annual revenues. Therefore, we would expect the CCDF of the allocation coefficients to have most of the mass around, or at the very least include this 10% threshold. Instead, the maximum value is around 0.7%, well below this threshold. The average expected allocation coefficient is 3×10^{-4} while the median is 8×10^{-5} (Table 4), both of which are three to four orders of magnitude smaller than what we would have expected given the data collection method.

Expected		
	Technical coefficient	Allocation coefficient
Mean	2×10^{-4}	3×10^{-4}
Median	3×10^{-5}	8×10^{-5}
Standard dev.	5×10^{-4}	6×10^{-4}

Table 4: Summary statistics of the technical and allocation coefficients for FactSet.

All three quantities have a narrow range of variation and are smaller than expected because total intermediate sales and expenditure of the proxy node (which represent the constraints that need to be satisfied in the weights allocation) are much bigger than those of the other firms. The proxy node accounts for 80% of intermediate expenditure and 75% of intermediate sales. To compare, in Ecuador, the proxy node accounts for 29% and 14%, respectively. Therefore, in FactSet, firms make most of their trades with the proxy node (Figure 6). The maximum transaction value for transactions between firms and the proxy node is \$407 million. But, the overall maximum is between the proxy node and itself, where it is \$48 billion, 3 orders of magnitude bigger than the maximum value among firms as well as among firms and the proxy node. Consequently, the coefficients representing trades among firms are much smaller than those representing the trades between firms and the proxy node, as shown in Figure 6.

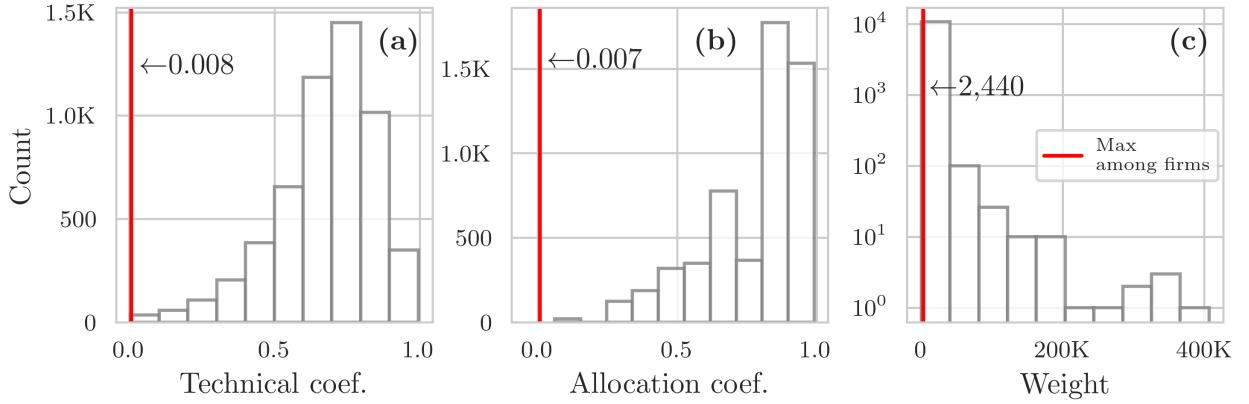


Figure 6: Trades between firms and the proxy node. (a) Percentage of inputs that firms buy from the proxy node (i.e., their technical coefficients with the proxy node). (b) Percentage of sales that firms make to the proxy node (i.e., their allocation coefficients with the proxy node). (c) Weights among firms and the proxy node (in both directions); values in thousand USD. We do not show quantities regarding trades of the proxy node with itself. The red line marks the maximum value of the quantity among firms only.

4.3.3 Discussion

Since the other studies that reconstruct firm-level production networks do not assess how their network reconstruction method performs on reconstructing weights, technical and allocation coefficients (Inoue and Todo, 2019; Welburn et al., 2020; Hooijmaaijers and Buitendijk, 2019; Reisch et al., 2021; Ialongo et al., 2022; Hillman et al., 2021), we cannot compare our results with these studies. Moreover, although several studies assess network reconstruction methods on the international trade network (ITN) or financial networks, few assess the reconstruction of the weights. Most of the studies look at higher-order network properties or dynamic indicators of systemic risk, which we discuss in Section 4.4.

Table 5 shows a comparison of our results with those of the literature. Parisi et al. (2020) test the reconstruction method on the ITN and the Electronic Market for Interbank Deposits (e-MID). They find a similar percentage of empirical weights that fall in the 50% CI. For the e-MID network, depending on the year, between 35% and 55% of the empirical weights fall in the 50% CI, while around 30% of the weights fall in the 50% CI for the ITN. They find that the empirical and expected link weights have a Pearson correlation of 0.50 for the e-MID and 0.75 for the ITN. The only other study we could find providing a comparison metric for link weights is Ramadiah et al. (2020). The authors test several reconstruction methods on Japan's bipartite bank-firm credit network. For the most disaggregated network, their reconstruction yields a cosine similarity of around 0.68 for the MaxEnt method (for bi-partite networks) and of 0.63 for the configuration fitness model with weights allocated using the IPF algorithm; these are lower than what we find for the weights but similar to our results for the technical and allocation coefficients.

While the power-law exponent of FactSet's weight distribution is in the ranges of what had been found for other supply chain networks (Bacilieri et al., 2023), the reconstructed weights and coefficients among firms are much smaller than expected. There are two main factors that deteriorate the quality of the reconstruction and act mainly through the constraints on intermediate sales and expenditure: the data cleaning procedure and the data imputations. On the one hand, the data cleaning procedure implies that we had to exclude many firms from the network (see Appendix A.1.4); this leads to (1) a higher share of the proxy node in the economy and (2) excluding many existing supply-chain relations. On the other hand, the data imputations concerning final demand and labour costs affect the values of firms' intermediate sales and expenditures. Because firms report the cost of goods sold, which often includes labour costs, we do not always know intermediate costs exactly. Similarly, firms disclose their revenues (intermediate sales plus sales to final demand), so for all firms, we do not know their intermediate sales exactly. As shown in Appendix A.1.2, labour costs tend to be overestimated. Although we cannot test whether final demand is over or underestimated, we think we are very likely overestimating it for the majority of firms. The main reason for overestimating final demand is the use of national I-O tables at the sector level, which have a very different treatment of the wholesale and retail sectors compared to firm-level data. National accounts treat wholesale and retail

Data set	Method	Quantity	Pct in 50% CI	Measure	Score	Source
Ecuador SC	CReM	Weight	47%	Cosine	0.93	This paper
Ecuador SC	CReM	Technical coefficient		Cosine	0.72	This paper
Ecuador SC	CReM	Allocation coefficient		Cosine	0.76	This paper
ITN	CReM	Weight	30%	Pearson	0.75	Parisi et al. (2020)
e-MID	CReM	Weight	35-55%	Pearson	0.50	Parisi et al. (2020)
Japan's bank-firm credit	MaxEnt	Weight		Cosine	0.68	Ramadiah et al. (2020)
Japan's bank-firm credit	BFiCM + IPF	Weight		Cosine	0.63	Ramadiah et al. (2020)

Table 5: Microscale quantities: comparison of our results with the literature. In column “Data set”, “Ecuador SC” stands for Ecuador supply chain network, “ITN” for international trade network and e-MID for electronic market for interbank deposits. In column “Method”, CReM stands for the conditional reconstruction method that we employ in this paper (Parisi et al., 2020). Besides the MaxEnt method, Ramadiah et al. (2020) use a two-step procedure. In the first step, the binary topology is reconstructed using a bipartite fitness-induced configuration model (BFiCM) (Squartini et al., 2017), which extends the fitness-induced configuration model of Cimini et al. (2015b) to the bipartite case. In the second step, weights are allocated using the iterative proportional fitting algorithm. For Ramadiah et al. (2020), we show results for the most disaggregated network and only for the two methods with the highest cosine similarity.

as “pass-through” sectors, accounting only for the trade margin they make and re-distribute the rest of their output among the other industries.¹⁰ The final demand of these other industries thus becomes (fictitiously) higher. In firm-level data quite the opposite happens, with many firms selling to retail and wholesale firms that then sell to final demand. We refer to Appendix A.3 and Bacilieri et al. (2023) for a longer discussion on differences between national accounts and firm-level data.

The data cleaning and data imputation problems just discussed imply that the constraints on intermediate sales and expenditure of the proxy sector are much bigger than those of other firms. Therefore, to satisfy the constraints posed in the maximum entropy procedure, the weights allocated to the proxy node need to be much bigger than those among firms, meaning that firms buy most of their inputs and sell most of their output to the proxy node. This deteriorates the reconstruction of the weights, and of the technical and allocation coefficients among firms.

Lastly, the findings for both Ecuador and FactSet suggest that additional mechanisms, not captured by the constraints we pose, may be at work that are essential to the formation of network weights. Further unravelling what these mechanisms are could improve the reconstruction.

4.4 Results: higher-order and macroscale properties

First, this section discusses the results for the output multipliers and the influence vector; we start with Ecuador and then FactSet. Second, we show the results for aggregate volatility. We do not assess aggregate volatility for FactSet as that would entail estimating TFP using an econometric technique such as that outlined in Magerman et al. (2016), which is outside of the scope of this paper. We conclude this section with a discussion.

4.4.1 Multipliers

Ecuador. The reconstruction method performs well at reproducing the empirical output multipliers but not that well at reconstructing the influence vector. Figure 7a and 7b show the empirical (x -axis) and the expected (y -axis) output multipliers and influence vector, respectively. For the output multipliers, points cluster fairly tightly around the identity line (dashed grey line), while for the influence vector most points are located at the bottom left corner, suggesting that the influence is consistently overestimated. However, there are a few exceptions, which tend to be firms with higher influence (Figure C.7). These findings are further confirmed in Figure 7c and 7d, showing the CCDF

¹⁰If wholesale is trading service goods, then all of its output is distributed to other products.

of the (empirical and reconstructed) output multipliers and influence vector, respectively. It also highlights that the minimum of the empirical influence vector is around one order of magnitude smaller than the reconstructed one. The cosine similarity of the output multipliers is higher than that of the influence vector (0.99 and 0.56, respectively); however, the other error metrics are less clear cut (Table 7). The reconstruction method can recover the first two moments and the median of the output multipliers (Table 6). For the influence vector, we can recover only the standard deviation.

In Figure 7a clusters tend to form. Clusters arise because, for the output multipliers, we simulated firms' value-added using sector-level data. The interaction of firms' intermediate expenditures, sectoral ν_s 's and the binary topology produces those clusters; see Appendix C.3.3 for more details.

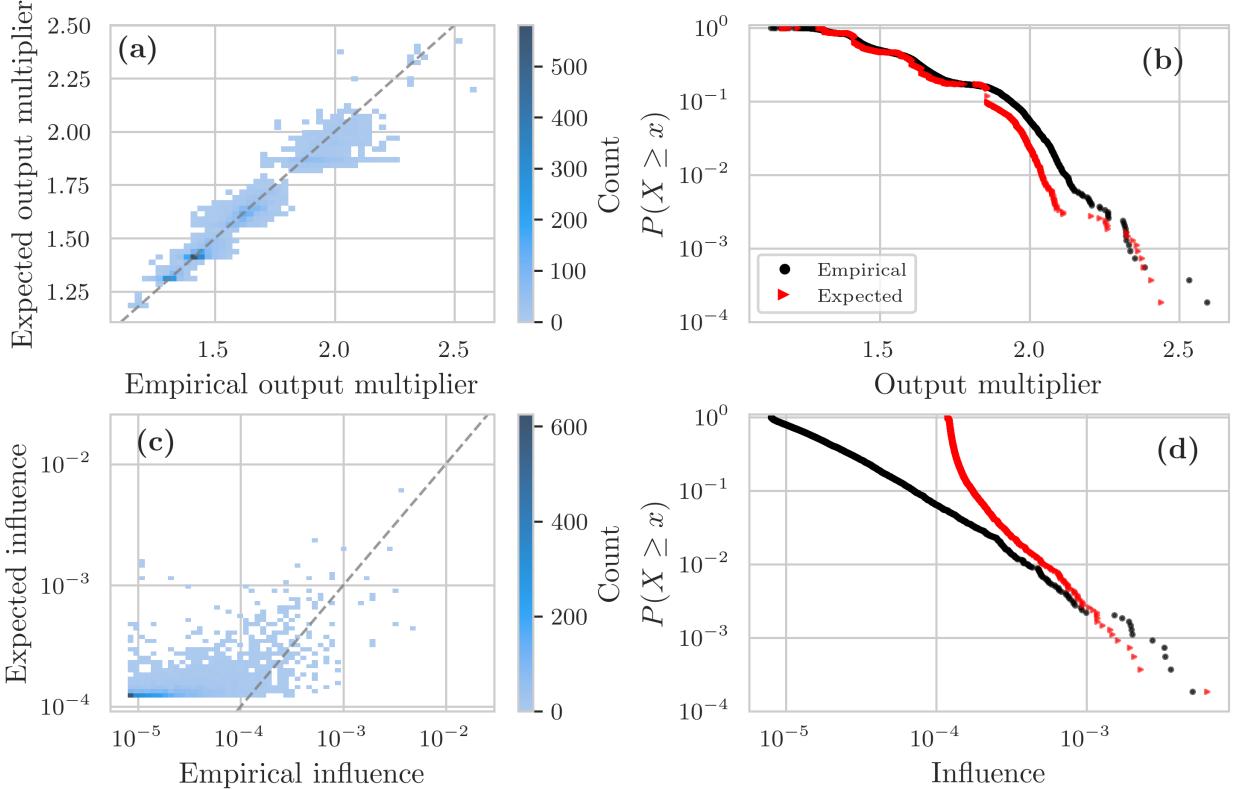


Figure 7: (a), (c) 2D histogram of the empirical (x -axis) and reconstructed (y -axis) output multipliers and influence vector, respectively, for Ecuador. Perfect prediction is achieved when points lie on the 45-degree line (dashed grey line). We use 50 log-spaced bins for both axes and count the number of points falling in each square. (b), (d) CCDF of the output multipliers and influence vector, respectively, for Ecuador. Black dots refer to the empirical CCDF and red triangles to the CCDF of the expected network. For calculating the empirical multipliers, we use the full network and consider the multipliers of firms in our test network only.

	Output multiplier		Influence vector	
	Empirical	Expected	Empirical	Expected
Mean	1.423	1.553	4×10^{-5}	1×10^{-4}
Median	1.420	1.475	2×10^{-5}	1×10^{-4}
Standard dev.	0.247	0.202	1×10^{-4}	1×10^{-4}

Table 6: Summary statistics of the output multipliers and influence vector for Ecuador. For each multiplier, the first column reports summary statistics for the multipliers calculated on the empirical network while the second column on the reconstructed network. We show results for one of the reconstructions; all three quantities have virtually the same summary statistics across the 50 randomised empirical and reconstructed networks. We excluded the proxy node from the calculations of the expected multipliers. For calculating the empirical multipliers, we use the full network and consider the multipliers of firms in our test network only.

Type	RMSE	MAE	MedAE	Cosine similarity
Output multipliers	0.036 (3×10^{-4})	0.024 (1×10^{-4})	0.015 (2×10^{-4})	0.999 (8×10^{-6})
	0.033 (2×10^{-4})	0.024 (2×10^{-5})	0.022 (1×10^{-5})	0.560 (3×10^{-3})
Influence vector				

Table 7: Comparison metrics for Ecuador’s multipliers. As defined in the main text, RMSE denotes the root mean squared error, MAE the mean absolute error and MedAE the median absolute error. For each multiplier, we show its mean value across the 50 randomised reconstructions. Below the mean value, the standard deviation in parenthesis. We excluded the proxy node from the calculations of the expected multipliers. For calculating the empirical multipliers, we use the full network and consider the multipliers of firms in our test network only.

FactSet. Figure 8a and 8b show the CCDF of, respectively, the expected output multipliers and the expected influence vector for FactSet. While the CCDF of the influence vector displays heavy tails, that of the output multipliers does not. As done for the weight distribution, we compare with empirical findings of other networks where the influence vector is found to have heavy tails and likely follows a power-law with a divergent second moment (Bacilieri et al., 2023). The estimated power-law exponent is equal to 1.9, higher than what found for Belgium (1.12, Magerman et al., 2016), Hungary and Ecuador (around 1.3-1.5 and 1.2-1.4, respectively, Bacilieri et al., 2023). We show the CCDF and its power-law fit in Appendix C.3.1.¹¹ The output multipliers have a much higher median, first and second moment in FactSet (Table 8) compared to Ecuador (Table 6), while the influence vector has lower moments.

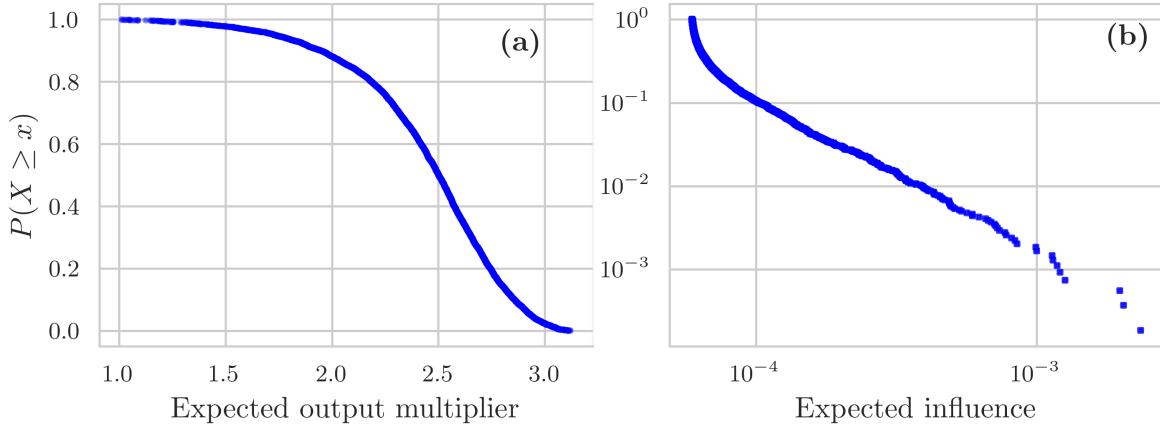


Figure 8: CCDF of the expected (a) output multipliers and (b) influence vector for FactSet. The multipliers of the proxy node are excluded.

Expected		
	Output multiplier	Influence vector
Mean	2.443	8×10^{-5}
Median	2.501	6×10^{-5}
Standard dev.	0.364	8×10^{-5}

Table 8: Summary statistics of the output multipliers and the influence vector for FactSet. The multipliers of the proxy node are excluded.

4.4.2 Supply-side shocks and aggregate volatility

We measure how firm-level TFP shocks affect macroeconomic output by propagating through the supply chain network using Equation 12. We simulate TFP shocks as explained in Section 4.2.3. We

¹¹We also fit a power-law distribution to Ecuador’s influence vector and recover an exponent of 2.0. However, the fit was rather poor. We show results for Ecuador in Appendix C.3.1.

then use the empirical influence vector to calculate the empirical volatility and the reconstructed influence vector to calculate the predicted volatility. The volatility predicted by our reconstructed network is much higher than the empirical one. Using the empirical network, GDP volatility is 6.4%, while the mean across the 50 reconstructions is 102.3% with a standard deviation of 0.002 (Table 9).

To understand why we predict such a high GDP volatility, we do a variance decomposition analysis and look at the role of the proxy node in explaining aggregate volatility:

$$\sigma_{\Delta y}^2 = \sum_{i \neq \kappa} \text{Var}(\Delta \epsilon_i) v_i^2 + \text{Var}(\Delta \epsilon_\kappa) v_\kappa^2,$$

where the first term on the RHS is the contribution to total GDP volatility of the firms in the test network and the second term is the contribution of the proxy node, indexed by κ . We further look at shares

$$\Lambda_S = \frac{\sum_{i \in S} \text{Var}(\Delta \epsilon_i) v_i^2}{\text{Var}(\Delta \epsilon_\kappa) v_\kappa^2 + \sum_{j \neq \kappa} \text{Var}(\Delta \epsilon_j) v_j^2}.$$

where S is the set of observations for which we are computing the share of the variance. We use the variance to ensure that shares sum up to 1. The proxy node contributes to 99.3% of the variance, while all the other firms to 0.7%. If we calculate GDP volatility without including the proxy node, the predicted GDP volatility drops to 8.8%.

	Reconstructed			Benchmark	
	Empirical	Proxy node	No proxy node	Proxy node	No proxy node
GDP volatility	6.4%	102.3% (0.0020)	8.8% (0.0010)	94% (0.0030)	7.0% (0.0002)

Table 9: Predicted aggregate GDP volatility using the empirical, the reconstructed and the benchmark influence vector (Equation 12). For the reconstructed and benchmark, “Proxy node” reports the predicted aggregate volatility calculated including the proxy node, while “No proxy node” excludes the proxy node from the calculations. Values for the reconstructed and the benchmark show the average taken over each of the 50 randomised networks. Below the mean value, we show the standard deviation in parenthesis.

Benchmark. To benchmark our results, we first calculate the influence vector assuming that each firm buys the same proportion of inputs from its suppliers (i.e., we modify Ω in Equation 11) and then compute aggregate volatility using Equation 12. Assigning homogeneous input shares for each firm still satisfies the constraints on the intermediate costs, but the constraints on intermediate sales are not guaranteed to be satisfied. Our benchmark yields a volatility of 94% when the proxy node is included and 7.0% when excluded (Table 9). The volatility predicted by the benchmark is 0.6 percentage points higher than the empirical volatility, while the volatility of the reconstruction is 2.4 percentage points higher than the empirical volatility.

4.4.3 Discussion

As noticed in Section 4.3.3, we cannot compare with previous results of other studies reconstructing firm-level production networks. For financial networks and the ITN, different reconstruction methods perform reasonably well in reconstructing higher-order network properties such as the weighted clustering coefficient (e.g., Mastrandrea et al., 2014; Cimini et al., 2015b,a; Parisi et al., 2020). Table 10 reports some of the findings in the literature. For financial networks, Ramadiah et al. (2020) find that all the reconstruction methods they employ underestimate the level of systemic risk, except for a small region of the parameter space. Anand et al. (2015) report similar findings for MaxEnt, but find that the minimum density method overestimates systemic risk. Di Gangi et al. (2018) find that the cross-entropy capital asset pricing model can reproduce very well systemic risk while the other ensemble methods overestimate or underestimate it depending on the shock scenario. Differently, individual banks’ systemic risk and indirect vulnerability, which could be seen as akin to the multipliers we test, are consistently underestimated across all the methods they assess. The degree-corrected gravity model can reproduce DebtRank (also akin to the multipliers), with a Person correlation equal to 1 (Cimini et al., 2015b).

Dataset	Method	Finding	Source
Ecuador SC	CReM	Overestimates influence vector, cosine sim. = 0.56	This paper
Ecuador SC	CReM	Under/Overestimates output multipliers, cosine sim. = 1.	This paper
e-Mid ITN	DcGM	DebtRank: correlation = 1	Cimini et al. (2015b)
	DcGM	DebtRank: correlation = 1	Cimini et al. (2015b)
US bank-asset	CE CAPM, Max. entropy CAPM, BWCM, BECM	Underestimate banks' measures of systemic risk	Di Gangi et al. (2018)
Ecuador SC	CReM	Overestimates aggregate volatility	This paper
German banks	MaxEnt	Underestimates systemic risk	Anand et al. (2015)
German banks	Minimum density	Overestimates systemic risk	Anand et al. (2015)
US bank-asset	Cross-entropy CAP model	Reproduces well aggregate vulnerability	Di Gangi et al. (2018)
US bank-asset	Max. entropy CAP model, BWCM, BECM	Over/underestimate aggregate vulnerability depending on the shock scenario	Di Gangi et al. (2018)
Japan bank-firm credit	MaxEnt, Min. density, CF + IPF, BFICM + IPF	Underestimate the average probability of default	Ramadiah et al. (2020)

Table 10: Higher-order and macroscale quantities: comparison of our results with the literature. In column “Dataset”, “Ecuador SC” stands for Ecuador supply chain network, “ITN” for international trade network and “e-MID” for electronic market for interbank deposits. In column “Method”, CReM stands for the conditional reconstruction method that we use in this paper ([Parisi et al., 2020](#)). “DcGM” stands for degree-corrected gravity model ([Cimini et al., 2015b](#)), “CE CAPM” for cross-entropy capital asset pricing model (it is a deterministic method), “Max. entropy CAPM” for maximum entropy capital asset pricing model, “BWCM” for bipartite weighted configuration model and “BECM” for bipartite enhanced configuration model (all of the last three methods are ensemble methods). “Min. density” is the minimum density method developed by [Anand et al. \(2015\)](#). “CF” is the bipartite maximum entropy configuration model and “BFICM” is the bipartite fitness-induced configuration model ([Squartini et al., 2017](#)); in both cases, weights are allocated in a second step using the iterative proportional fitting algorithm.

While the vast majority of the studies find that systemic risk is underestimated, for Ecuador, we find that the conditional maximum entropy reconstruction method we employ overestimates aggregate volatility, similar to what Anand et al. (2018) find for the minimum density method. First, while including the proxy node in the calculations of aggregate volatility is supposed to capture the volatility of the firms that were excluded from the test network, there is a fundamental difference between the role the proxy node has in the network and that of the firms that we excluded. The proxy node is connected to all the other firms in the test network. In contrast, the firms excluded from the test network are less well connected in the empirical network. Additionally, the proxy node has the biggest influence, which equals 0.19. Therefore, when the proxy node is shocked, it immediately passes 19% of the shock to all the other firms in the network. To compare, the firm with the biggest influence in the empirical network has an influence of 0.005, with the biggest influence among the firms excluded from the test network being much lower and equal to 0.0002. Noting also that the firms excluded from the test network contribute to a mere 7% of aggregate volatility while the firms included in the test network to 93%, it makes sense to exclude the proxy node from the calculation of aggregate volatility as it is not a good aggregate representation of the rest of the economy, at least in this context. In fact, while the inclusion of the proxy node degrades the prediction of aggregate volatility, it enhances the reconstruction of the multipliers; see Appendix C.3.4.

Second, aggregate volatility is still overestimated even when the proxy node is excluded because the influence vector tends to be overestimated. Notice that the benchmark can predict aggregate volatility more accurately since it tends to overestimate influences lightly less than the reconstruction (see Appendix C.3.5). The reconstructed influence vector is overestimated for a combination of two factors, one of which is related to the binary topology and the other to the weighted topology. We can see how the binary topology affects the influence vector by looking at the out-degrees, which are highly affected by the selection of firms and the link deletion. In creating the test network, firms lose more customers than suppliers (see Appendix C.4.1). Already at the first step (selecting the firms to keep in our test network), the maximum number of customers (out-degree) decreases almost 7 folds, while the maximum number of suppliers only 3 folds. Since the influence vector calculates the weighted sum of the number of walks from firm i to firm j (so following the outgoing edges starting at i) for walks of various lengths and the out-degrees have been considerably truncated, it cannot capture walks of longer lengths. Instead, the output multipliers, relying on the incoming edges and thus the in-degrees, are not affected. The weighted topology comes into play because the reconstruction method tends to overestimate weighted quantities for intermediate and small values, which are numerous. Although higher (and other) weights can be underestimated, these are not abundant enough and the magnitude of the underestimation is not big enough compared to the quantity and the size of the overestimation of low and intermediate weights.

For FactSet, a significant result is that the estimated power-law exponent of the distribution of the influence vector implies a divergent second moment. Acemoglu et al. (2012) show, at a theoretical level, that the influence vector affects aggregate volatility, so it is important that we can recover an exponent more or less in line with empirical observations.¹²

4.5 Results: different numbers of unknown links

We now discuss how the reconstruction method performs when the number of unknown links in the test network changes; we keep the number of firms constant. We delete 0%, 10%, 20% and so on up to 90% of the links; we also show results for the reconstruction matching the mean degree of FactSet, which has 96% of unknown links. For each percentage of unknown links, we simulate 50 randomised networks and investigate the number of weights that fall into the 50% CI, and the median absolute error and the cosine similarity for microscale and higher-order quantities. We also assess aggregate volatility.

¹² Acemoglu et al.'s (2012) theoretical result is exemplified in Equation 12, which shows that aggregated volatility scales with the Euclidean norm of the influence vector. Luca's diversification argument implies that aggregate volatility decays as $N^{1/2}$. However, if the CCDF of the influence vector is Pareto distributed with parameter $\gamma \in (1, 2)$, the diversification argument no longer holds and aggregate volatility decays much more slowly, as $N^{1/(\gamma-1)}$ (Carvalho and Tahbaz-Salehi, 2019).

Microscale and higher-order quantities. Figure 9 shows the different metrics used for assessing the quality of the reconstruction of the weights, and the technical and allocation coefficients, and the multipliers. For the weights, we show the percentage of weights that fall in the 50% CI (Figure 9a) and the cosine similarity (Figure 9b). For all the other quantities, we show the median absolute error (left column) and the cosine similarity (right column). The overall trend is that as the number of unknown links increases, so do the metrics.

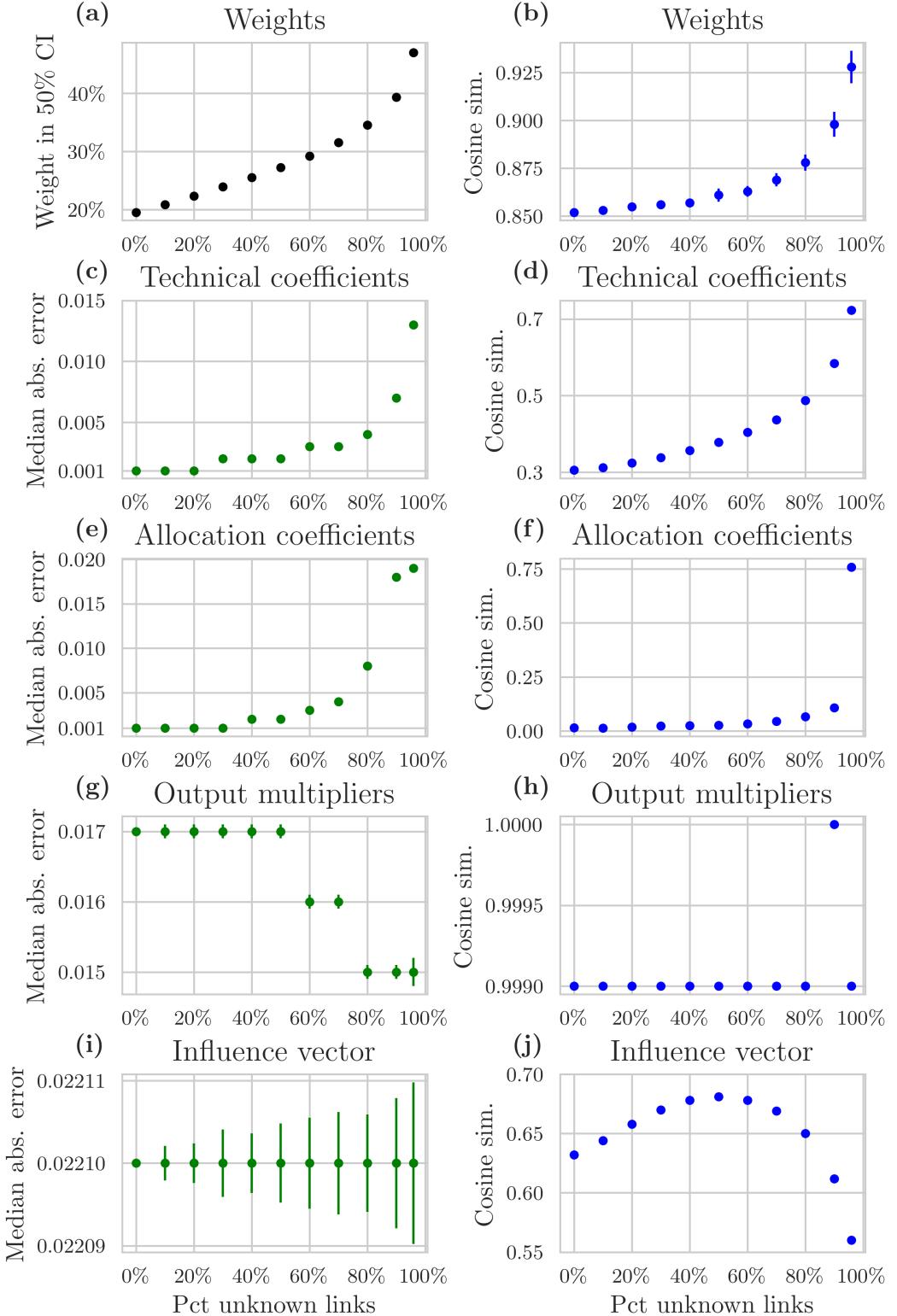


Figure 9: Error metrics and similarity measures as the number of unknown links increases. (a) Percentage of empirical weights in the 50% confidence interval. (b), (d), (f), (h), (j) cosine similarity for, respectively, weights, technical and allocation coefficients, output multipliers and influence vector. (c), (e), (g), (i), Mean absolute error for, respectively, technical and allocation coefficients, output multipliers and influence vector. Bars show the standard deviation across the 50 randomised networks.

As the number of unknown links increases, the cosine similarity increases, with the two multipliers being the only exceptions. Such a counter-intuitive finding arises from the link deletion mechanism. We are deleting links with a smaller weight with a higher probability, hence as the number of unknown links increases, the empirical weights associated with the known links are less heterogeneous. Since maximum entropy methods allocate weights as uniformly as possible given the constraints, the higher the number of unknowns, the better it can reconstruct the weights associated with the known links. Consequently, also more weights fall into the 50% confidence interval.

For the allocation coefficients, the cosine similarity is considerably lower than that of all the other quantities and it jumps from 0.05 to 0.76 when the number of unknown links increases from 90% to 96%. This jump is associated with a considerable decrease in the number of customers firms have when the number of unknown links increases from 90% to 96% (see Appendix C.4.1). Therefore, the allocation coefficients, which gauge the percentage of total output firm i sells to firm j , are highly affected by a drastic decrease in the number of customers (see Appendix C.4.2). However, we would have expected this to have a negative effect. Instead, it enhances the predictions. On the contrary, the technical coefficients, which measure the percentage of inputs that j buys from i , do not experience such a jump because the same drastic change does not happen for the number of suppliers. Moreover, firms tend to have more customers than suppliers (see Appendix C.4.1 and Bacilieri et al., 2023), making it easier to guess the weights correctly from the supply side and thus reconstruct the technical coefficients better than the allocation coefficients throughout.

The output multipliers always have the same cosine similarity. Since the output multipliers are derived from the technical coefficient matrix and depend on the incoming edges, hence the number of suppliers firms have, they are not that affected by the number of unknown links. Instead, the influence vector relies on the number of customers firms have (out-degree), which we saw being more affected by the deletion of firms and links, so the influence vector has a much lower cosine similarity than the output multipliers. It is not so clear why the cosine similarity has an inverted U-shaped curve, with the maximum value of 0.68 reached when 50% of the links are unknown. It is however the case that the cosine similarity increases from 0.63 when all links are known to 0.68 (50% unknown links) and then decreases again until it drops to 0.56 when 96% of the links are unknown. It might be that these are small fluctuations of no particular value and that the reconstruction of the influence vector starts deteriorating when more than 50% of the links are unknown because the out-degrees start being too affected by the deletion of the links.

The median absolute error slightly increases for the technical and allocation coefficients because higher errors (in absolute value) tend to occur for higher weights. The median absolute error is constant for the multipliers.

Aggregate volatility. Figure 10 shows the predicted aggregated volatility (calculated excluding the proxy node) as the number of unknown links increases for the reconstruction (red dots), the benchmark (green squares) and the empirical volatility (black dashed line). In line with the results for the case of 96% of unknown links, the reconstruction always predicts a much higher volatility, as does the benchmark. Although the benchmark's predictions are closer to the empirical volatility. One would think that the higher the number of known links, the better we get at predicting aggregate volatility. However, this is not the case. In fact, the fewer links we know, the better we can predict GDP volatility. This is because the more links we know, the more we overestimate the influences (especially bigger influences, see Appendix C.4.2).

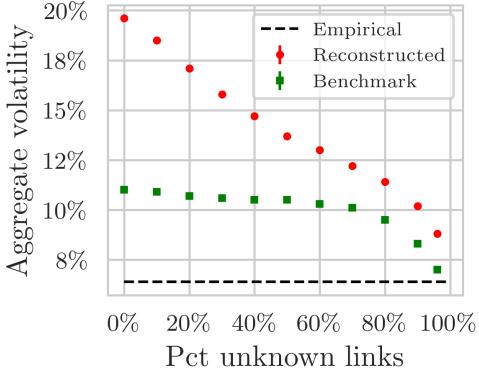


Figure 10: Aggregate volatility as the number of unknown links increases using the reconstructed network (red dots) and the benchmark (green squares). The black dashed line marks the empirical volatility. Bars show the standard deviation across the 50 randomised networks. We show the volatility calculated excluding the proxy node.

5 Conclusions

There is widespread interest in modelling the global economy from the bottom up. There is also widespread agreement that supplier-customer relations are an essential feature of such modelling efforts. However, data are scarce, have missing information and are not easily accessible. In this paper, we have made a couple of first steps in bringing this agenda forward.

First, we provided the first rigorous assessment of a network reconstruction method (Parisi et al., 2020) on the administrative dataset of Ecuador. We focused on reconstructing the weighted network given the binary topology when many links and firms are missing. An interesting finding is that the quality of the reconstruction of different quantities seems to depend on network features that are particularly sensitive to the sampling strategy of firms and links, something that future research should explore further. Second, we assessed whether a global dataset of listed firms, where many links and firms are missing, can be enhanced by merging it with sector-level data. We then used this “augmented” dataset for inferring the link weights using a conditional maximum entropy method (Parisi et al., 2020). Our results show that further work needs to be done, especially in reconciling firms’ financial accounts with national accounts, which is essential for better reconstructing the weighted production network, in particular when many firms are missing.

In our study, we assumed to know the binary topology (although partially) to cover a use case that could help reconstruct commercial datasets. A natural next step would be to predict links and then reconstruct weights, possibly with different degrees of knowledge of the production network. It is of particular value to understand the performance of reconstruction methods when one does not know the binary topology, as it could unlock the study of economies for which no such data is available. The reconstruction method we employed can easily accommodate the prediction of the binary topology, either partially or in its entirety.

Our assessment of macroscale quantities was rather limited and restricted to a standard general equilibrium I-O model (Acemoglu et al., 2012). Therefore, there is much research to be done on different models (and scenarios), especially agent-based models, where we believe that the accuracy of the reconstruction of microscale quantities matters more for the model’s outcomes than for general equilibrium models.

While many reconstruction methods are available, the high number of nodes and links in firm-level networks renders almost all of them infeasible. Future research is thus necessary to develop different reconstruction methods suited for large-scale firm-level networks. It is also important to conduct similar analyses on other firm-level datasets for which the ground truth network is available, as ours was only one of the first initial steps.

Acknowledgments

We thank François Lafond, Jose Moran, R. Maria del Rio-Chanona, Anton Pichler, Luca Mugo and Fabian Dablander for useful comments. We thank Doyne J. Farmer, Jangho Yang, Giancarlo Antonucci

and the Complexity group at the Institute for New Economic Thinking at the Oxford Martin School for inspiring discussions. We have benefited from many comments from audiences at Cambridge IfM and CSH Vienna. We also thank the Servicio de Rentas Internas (SRI) and its Centro de Estudios Fiscales that provided the raw data of Ecuador for research purposes. This work was supported by the Oxford Martin Programme on the Post-Carbon Transition, the Institute for New Economic Thinking at the Oxford Martin School and Baillie Gifford. This research is based upon work supported in part by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), via contract no. 2019-1902010003. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of ODNI, IARPA, or the U.S. Government. The U.S. Government is authorised to reproduce and distribute reprints for governmental purposes notwithstanding any copyright annotation therein. FactSet had the opportunity to review the paper.

Author contributions

A. Bacilieri suggested the research topic, designed the work, implemented the reconstruction method and associated tests, analysed the data and wrote the paper. Pablo Astudillo-Estevez provided the cleaned Ecuador dataset and relevant information about it. They also declare that there were no conflicts of interest.

Research data

We do not have permission to share the two firm-level datasets. For FactSet because it is a commercial dataset, while for Ecuador because it is a confidential, administrative dataset with restricted access.

References

- Acemoglu, D., Carvalho, V. M., Ozdaglar, A., and Tahbaz-Salehi, A. (2012). The network origins of aggregate fluctuations. *Econometrica*, 80(5):1977–2016.
- Anand, K., Craig, B., and von Peter, G. (2015). Filling in the blanks: Interbank linkages and systemic risk. *Quantitative Finance*, 15:625–636.
- Anand, K., van Lelyveld, I., Banai, Á., Friedrich, S., Garratt, R., Halaj, G., Fique, J., Hansen, I., Jaramillo, S. M., Lee, H., et al. (2018). The missing links: A global study on uncovering financial network structures from partial data. *Journal of Financial Stability*, 35:107–119.
- Astudillo-Estevez, P. A. (2021). *Towards a Post-Oil Economy: A Complexity Approach to Understanding Natural Resource Dependency and Economic Diversification in Ecuador*. PhD thesis, University of Oxford.
- Atalay, E., Hortacsu, A., Roberts, J., and Syverson, C. (2011). Network structure of production. *Proceedings of the National Academy of Sciences*, 108(13):5199–5202.
- Bacilieri, A., Borsos, A., Astudillo-Estevez, P., and Lafond, F. (2023). Firm-level production networks: what do we (really) know? Unpublished results, University of Oxford.
- Baral, P. and Fique, J. (2012). Estimation of bilateral connections in a network: Copula vs. maximum entropy. In CIRANO https://cirano.qc.ca/conferences/public/pdf/networks2012/02-BARAL-FIQUE-Estimation_of_Bilateral_Exposures-A_Copula_Approach.pdf.
- Barrot, J.-N. and Sauvagnat, J. (2016). Input specificity and the propagation of idiosyncratic shocks in production networks. *The Quarterly Journal of Economics*, 131(3):1543–1592.
- Bernard, A. B., Dhyne, E., Magerman, G., Manova, K., and Moxnes, A. (2019). The origins of firm heterogeneity: A production network approach. Technical report, National Bureau of Economic Research.
- Boehm, J. and Sonntag, J. (2022). Vertical integration and foreclosure: Evidence from production network data. *Management Science*.

- Brintrup, A., Wichmann, P., Woodall, P., McFarlane, D., Nicks, E., and Krechel, W. (2018). Predicting hidden links in supply networks. *Complexity*, 2018:1–12.
- Carvalho, V. M. (2014). From micro to macro via production networks. *Journal of Economic Perspectives*, 28(4):23–48.
- Carvalho, V. M., Nirei, M., Saito, Y. U., and Tahbaz-Salehi, A. (2021). Supply chain disruptions: Evidence from the great east japan earthquake. *The Quarterly Journal of Economics*, 136(2):1255–1321.
- Carvalho, V. M. and Tahbaz-Salehi, A. (2019). Production networks: A primer. *Annual Review of Economics*, 11:635–663.
- Cimini, G., Mastrandrea, R., and Squartini, T. (2021). *Reconstructing networks*. Cambridge University Press.
- Cimini, G., Squartini, T., Gabrielli, A., and Garlaschelli, D. (2015a). Estimating topological properties of weighted networks from limited information. *Physical Review E*, 92(4):040802.
- Cimini, G., Squartini, T., Garlaschelli, D., and Gabrielli, A. (2015b). Systemic risk analysis on reconstructed economic and financial networks. *Scientific reports*, 5:15758.
- Clauzel, A., Shalizi, C. R., and Newman, M. E. (2009). Power-law distributions in empirical data. *SIAM review*, 51(4):661–703.
- Di Gangi, D., Lillo, F., and Pirino, D. (2018). Assessing systemic risk due to fire sales spillover through maximum entropy network reconstruction. *Journal of Economic dynamics and control*, 94:117–141.
- Diem, C., Borsos, A., Reisch, T., Kertész, J., and Thurner, S. (2022). Quantifying firm-level economic systemic risk from nation-wide supply networks. *Scientific reports*, 12(1):1–13.
- Fally, T. (2012). Production staging: measurement and facts. University of Colorado Boulder, Boulder, Colorado.
- Gandy, A. and Veraart, L. A. (2017). A Bayesian methodology for systemic risk assessment in financial networks. *Management Science*, 63(12):4428–4446.
- Golan, A., Judge, G., and Robinson, S. (1994). Recovering information from incomplete or partial multisectoral economic data. *The Review of Economics and Statistics*, pages 541–549.
- Herskovic, B., Kelly, B., Lustig, H., and Van Nieuwerburgh, S. (2020). Firm volatility in granular networks. *Journal of Political Economy*, 128(11):4097–4162.
- Hillman, R., Barnes, S., Wharf, G., and MacDonald, D. (2021). A new firm-level model of corporate sector interactions and fragility: the Corporate Agent-Based (CAB) Model. Economics Department Working Papers No. 1675, OECD. Available at: https://www.oecd-ilibrary.org/economics/a-new-firm-level-model-of-corporate-sector-interactions-and-fragility-the-corporate-agent-based-cab-model_e9de0097-en.
- Hooijmaaijers, S. and Buitenhuis, G. (2019). A methodology for estimating the dutch interfirm trade network, including a breakdown by commodity. Technical report, Statistics Netherlands.
- Ialongo, L. N., de Valk, C., Marchese, E., Jansen, F., Zmarrou, H., Squartini, T., and Garlaschelli, D. (2022). Reconstructing firm-level interactions in the dutch input–output network from production constraints. *Scientific reports*, 12(1):1–12.
- Inoue, H. and Todo, Y. (2019). Firm-level propagation of shocks through supply-chain networks. *Nature Sustainability*, 2(9):841–847.
- Kosasih, E. E. and Brintrup, A. (2022). A machine learning approach for predicting hidden links in supply chain with graph neural networks. *International Journal of Production Research*, 60(17):5380–5393.

- Lahr, M. and De Mesnard, L. (2004). Biproportional techniques in input-output analysis: table updating and structural analysis. *Economic Systems Research*, 16(2):115–134.
- Lebacher, M., Cook, S., and Klein, N. (2021). In search of lost edges: a case study on reconstructing financial networks. *Journal of Network Theory in Finance*.
- Lenzen, M., Gallego, B., and Wood, R. (2009). Matrix balancing under conflicting information. *Economic Systems Research*, 21(1):23–44.
- Magerman, G., De Bruyne, K., Dhyne, E., and Van Hove, J. (2016). Heterogeneous firms and the micro origins of aggregate fluctuations. NBB Working Paper No. 312, National Bank of Belgium, Brussels.
- Mastrandrea, R., Squartini, T., Fagiolo, G., and Garlaschelli, D. (2014). Enhanced reconstruction of weighted networks from strengths and degrees. *New Journal of Physics*, 16(4):043022.
- Mastromatteo, I., Zarinelli, E., and Marsili, M. (2012). Reconstruction of financial networks for robust estimation of systemic risk. *Journal of Statistical Mechanics: Theory and Experiment*, 2012(03):P03011.
- Mazzarisi, P. and Lillo, F. (2017). Methods for reconstructing interbank networks from limited information: A comparison. In *Econophysics and Sociophysics: Recent Progress and Future Directions*, pages 201–215. Springer.
- McDougall, R. A. (1999). Entropy theory and RAS are friends. GTAP Working Paper No. 300, Center for Global Trade Analysis, Department of Agricultural Economics, Purdue University.
- McNerney, J., Savoie, C., Caravelli, F., Carvalho, V. M., and Farmer, J. D. (2022). How production networks amplify economic growth. *Proceedings of the National Academy of Sciences*, 119(1).
- Miller, R. E. and Blair, P. D. (2009). *Input-output analysis: foundations and extensions*. Cambridge University Press.
- Miller, R. E. and Temurshoev, U. (2017). Output upstreamness and input downstreamness of industries/countries in world production. *International Regional Science Review*, 40(5):443–475.
- Moussa, A. (2011). *Contagion and systemic risk in financial networks*. PhD thesis, Columbia University.
- Mungo, L., Lafond, F., Astudillo-Estevez, P., and Farmer, J. D. (2022). Reconstructing production networks using machine learning. INET Oxford Working Paper No. 2022-02, Institute for New Economic Thinking at the Oxford Martin School. Available at: <https://www.inet.ox.ac.uk/publications/no-2022-02-reconstructing-production-networks-using-machine-learning/>.
- Pankratz, N. M. and Schiller, C. M. (2019). Climate change and adaptation in global supply-chain networks. Proceedings of Paris December 2019 Finance Meeting EUROFIDAI - ESSEC, European Corporate Governance Institute – Finance Working Paper No. 775/2021. Available at SSRN: <https://ssrn.com/abstract=3475416> or <http://dx.doi.org/10.2139/ssrn.3475416>.
- Parisi, F., Squartini, T., and Garlaschelli, D. (2020). A faster horse on a safer trail: generalized inference for the efficient reconstruction of weighted networks. *New Journal of Physics*, 22(5):053053.
- Pichler, A. and Farmer, J. D. (2021). Simultaneous supply and demand constraints in input–output networks: The case of covid-19 in germany, italy, and spain. *Economic Systems Research*, pages 1–21.
- Pöyhönen, P. (1963). A tentative model for the volume of trade between countries. *Weltwirtschaftliches Archiv*, pages 93–100.
- Ramadiah, A., Caccioli, F., and Fricke, D. (2020). Reconstructing and stress testing credit networks. *Journal of Economic Dynamics and Control*, 111:103817.

- Reisch, T., Heiler, G., Diem, C., and Thurner, S. (2021). Inferring supply networks from mobile phone data to estimate the resilience of a national economy. *arXiv preprint arXiv:2110.05625*.
- Robinson, S., Cattaneo, A., and El-Said, M. (2001). Updating and estimating a social accounting matrix using cross entropy methods. *Economic Systems Research*, 13(1):47–64.
- Squartini, T., Almog, A., Caldarelli, G., Van Lelyveld, I., Garlaschelli, D., and Cimini, G. (2017). Enhanced capital-asset pricing model for the reconstruction of bipartite financial networks. *Physical Review E*, 96(3):032315.
- Squartini, T., Caldarelli, G., Cimini, G., Gabrielli, A., and Garlaschelli, D. (2018). Reconstruction methods for networks: the case of economic and financial systems. *Physics Reports*, 757:1–47.
- Squartini, T. and Garlaschelli, D. (2014). Jan tinbergen’s legacy for economic networks: from the gravity model to quantum statistics. In *Econophysics of Agent-Based Models*, pages 161–186. Springer.
- Stolowy, H. and Lebas, M. (2006). *Financial accounting and reporting: a global perspective*. Cengage Learning EMEA.
- Taschereau-Dumouchel, M. (2020). Cascades and fluctuations in an economy with an endogenous production network. Available at SSRN: <https://ssrn.com/abstract=3115854> or <http://dx.doi.org/10.2139/ssrn.3115854>.
- Timmer, M. P., Dietzenbacher, E., Los, B., Stehrer, R., and De Vries, G. J. (2015). An illustrated user guide to the world input–output database: the case of global automotive production. *Review of International Economics*, 23(3):575–605.
- Tinbergen, J. (1962). *Shaping the World Economy: Suggestions for an International Economic Policy*. New York: Twentieth Century Fund.
- United Nations (2010). *System of National Accounts 2008*. European Communities, International Monetary Fund, Organisation for Economic Co-operation and Development, United Nations, World Bank.
- Upper, C. (2011). Simulation methods to assess the danger of contagion in interbank markets. *Journal of Financial Stability*, 7(3):111–125.
- Voitalov, I., van der Hoorn, P., van der Hofstad, R., and Krioukov, D. (2019). Scale-free networks well done. *Physical Review Research*, 1(3):033034.
- Wang, Y., Li, J., Wu, D., and Anupindi, R. (2021). When ignorance is not bliss: An empirical analysis of subtler supply network structure on firm risk. *Management Science*, 67(4):2029–2048.
- Welburn, J. W., Strong, A., Nekoul, F. E., Grana, J., Marcinek, K., Osoba, O., Koirala, N., and Setodji, C. M. (2020). Systemic Risk in the Broad Economy: Interfirm Networks and Shocks in the US Economy. Santa Monica, CA: RAND Corporation, 2020. Available at: https://www.rand.org/pubs/research_reports/RR4185.html.
- Wells, S. J. (2004). Financial interlinkages in the United Kingdom’s interbank market and the risk of contagion. Bank of England Working Paper No. 230, Bank of England. Available at SSRN: <https://ssrn.com/abstract=641288> or <http://dx.doi.org/10.2139/ssrn.641288>.
- Wu, J. (2016). *Firm performance and risk in supply chain networks*. PhD thesis, The University of Chicago.
- Yang, J., Heinrich, T., Winkler, J., Lafond, F., Koutroumpis, P., and Farmer, J. (2019). Measuring productivity dispersion: a parametric approach using the Lévy alpha-stable distribution. INET Working Paper No. 2019-14, Institut of New Economic Thinking at the Oxford Martin School, University of Oxford.

Appendices

A Data

A.1 FactSet: description of the dataset

We use three primary data sources provided by FactSet: Fundamentals, Supply Chain Relationships and Supply Chain Shipping Transactions. FactSet covers mostly listed firms around the world. We downloaded the datasets in April 2020. We refer to [Bacilieri et al. \(2023\)](#) for a more detailed description of the dataset. We remark only that due to the nature of the data collection process, coverage is biased towards companies listed on US stock exchanges, big firms and big transactions.

The monetary values of customer-supplier transactions are rarely recorded. When recorded, the money flows are reported as a revenue percentage earned by the seller from a specific customer. However, it is unknown to what disclosed revenue figure that percentage refers. For example, it could be the quarterly or the annual income statement, but also some interim revenue disclosure not tied to any specific financial statement. Similarly, in the shipment dataset, the cumulative value of the goods shipped, when disclosed, is in most cases above the seller's revenues and/or the customers' costs (for more information on why this is the case, see [Bacilieri et al., 2023](#)). Therefore, we dismiss this information when reconstructing the money flows. There are other, less comprehensive datasets, such as Compustat, which report the revenue percentage earned by the seller. These could be potentially merged to enhance the reconstruction. We leave this for future work.

We consider industrial firms only, hence we exclude firms in the financial and insurance sector, and firms classified as extraterritorial organisations. We aggregate customer-supplier relations within a year. We use the fiscal year instead of the calendar year to ensure time consistency between the formation of supplier-customer relations and financial statements. The fiscal year goes from June to May, meaning that if a company's fiscal year end-month falls between January and May, the fiscal year is the current calendar year minus one, otherwise it is the current calendar year.

We further aggregate all three FactSet datasets at the parent company level, meaning that we use consolidated income statements and attribute subsidiaries' supplier-customer relations to the parent company. We delete self-loops (i.e., supply chain relations among the parent and its subsidiaries) as these stem from intra-group sales that cancel out in consolidated income statements. We rely on the latest available information on a company's ownership structure since it is impossible to know the evolution of companies' ownership structures (mergers, acquisitions, buy-backs, etc.). For each company, we also have information on the sector (NACE Rev.2 codes at the 4-digit level) and the country where the company's headquarters are located.

A.1.1 Methods for writing the income statement

There are two main methods companies can follow to write their income statements: the *function of expense* method or the *nature of expense* method – some companies adopt a hybrid method. As the names suggest, the nature of expense method lists expenses based on their nature, while the function of expense method lists expenses based on their function. Thus, the nature of expense method breaks down expenses based on the inputs used to perform the business activity, e.g., materials, delivery charges, changes in inventory, rent, labour expenses and employee benefits. On the other hand, the function of expense method allocates expenses based on the activity for which the expense arises. Therefore, “if the expense did not contribute to the creation of the [good] or service that is the underlying source of sales revenue, they are not part of the cost of goods sold.” ([Stolowy and Lebas, 2006](#), p. 208).

A.1.2 Labour expenses

Depending on the method followed in writing the income statement (explained in Appendix A.1.1), companies may disclose or not their labour expenses as a separate line item. If a company follows the *nature of expense* method, labour costs and the cost of intermediate inputs are broken down into two separate items on the income statement. However, if a company follows the *function of expense* methods, labour and intermediate inputs expenses are lumped together in the cost of goods sold.

Moreover, even if they follow a function of expenses method, some companies may disclose their labour costs in a footnote.

FactSet collects the labour expenses disclosed in the footnote but without flagging the method followed by the company in writing its income statement. Therefore, there is no systematic way to know whether a company follows the nature of expense or the function of expense method. The only reasonable assumption is that if labour expenses are not observed in the dataset, it is because the company followed the nature of expenses method and thus the cost of goods sold correctly represents intermediate inputs costs. Consequently, it is impossible to extract labour costs from the cost of goods sold only when necessary.

Estimating missing labour expenses. For constructing the I-O table, we must discern labour and intermediate inputs expenses as they correspond to different parts of the I-O table. Labour costs are part of value-added, while intermediate input expenses correspond to in-strengths, i.e., the column sums of the weighted adjacency matrix (also referred to as inter-industry transaction matrix in sector-level data). Having an exact quantification of intermediate expenses is necessary to reconstruct the firm-to-firm money flows (i.e., network weights).

To estimate labour expenses for those firms that do not report them, we use the cost of goods sold share of labour expenses disclosed by other firms in the same sector. For each sector, we calculate the sector-level cost of goods sold share of labour expenses using four different estimation strategies, all of which use a rolling window approach. In this paper, we use the method that minimises the root mean squared error, the average absolute error and the median absolute error. On the one hand, method 2a and 3 achieve the minimum RMSE and MAE while method 2b achieves the lowest MedAE regarding the cost of goods sold share of labour expenses (Table A.2). On the other hand, if we look at the estimated labour costs and intermediate costs, it is method 2b that has the minimum error (Table A.3). Therefore, we use method 2b in this paper.

We estimate labour expenses only from 2013 onward and start the rolling window in 2011 because before 2011 the number of firms in each sector was too low. For firms that disclose their labour expenses only for a few years, we extrapolate the missing values of the cost of goods sold share of labour expenses from those that we observe: (1) if we observe labour expenses up to time t and the missing values are only from $t + 1$ onward, we use the last value of the cost of goods sold share of labour expenses and propagate it forward; (2) if we observe labour expenses from t onward but do not observe labour expenses from 2013 till t , we backpropagate the first available observation; and (3) if we do not observe labour costs in the middle, meaning that we do observe labour costs from 2013 to t and from $t + \tau$ until the end, we estimate the cost of goods sold share of labour expenses with linear interpolation. Table A.1 reports the percentage of firms that disclose labour costs for each sector average over time (second column) and its standard deviation (last column), while Figure A.1 shows the density of firms' cost of goods sold share of labour expenses for each sector (NACE Rev. 2 at the 2-digit level) over time.

NACE division description	Av. pct firms	Standard dev.
Accommodation	0.618	0.038
Activities of head offices; management consultancy activities	0.528	0.022
Activities of membership organizations	0.750	0.274
Advertising and market research	0.575	0.026
Air transport	0.882	0.042
Architectural and engineering activities; technical testing and analysis	0.732	0.024
Civil engineering	0.663	0.015
Computer programming, consultancy and related activities	0.530	0.022
Construction of buildings	0.608	0.021
Creative, arts and entertainment activities	0.944	0.167
Crop and animal production, hunting and related service activities	0.653	0.052
Education	0.554	0.052
Electricity, gas, steam and air conditioning supply	0.639	0.018
Employment activities	0.541	0.029
Extraction of crude petroleum and natural gas	0.489	0.040

Fishing and aquaculture	0.707	0.068
Food and beverage service activities	0.430	0.020
Forestry and logging	0.722	0.105
Gambling and betting activities	0.758	0.035
Human health activities	0.692	0.016
Information service activities	0.450	0.028
Land transport and transport via pipelines	0.562	0.060
Legal and accounting activities	0.669	0.093
Libraries, archives, museums and other cultural activities	0.750	0.289
Manufacture of basic metals	0.677	0.023
Manufacture of basic pharmaceutical products and pharmaceutical preparations	0.610	0.026
Manufacture of beverages	0.685	0.035
Manufacture of chemicals and chemical products	0.603	0.031
Manufacture of coke and refined petroleum products	0.739	0.023
Manufacture of computer, electronic and optical products	0.633	0.027
Manufacture of electrical equipment	0.680	0.050
Manufacture of fabricated metal products, except machinery and equipment	0.634	0.022
Manufacture of food products	0.680	0.014
Manufacture of furniture	0.604	0.028
Manufacture of leather and related products	0.709	0.022
Manufacture of machinery and equipment n.e.c.	0.575	0.047
Manufacture of motor vehicles, trailers and semi-trailers	0.629	0.043
Manufacture of other non-metallic mineral products	0.657	0.021
Manufacture of other transport equipment	0.639	0.036
Manufacture of paper and paper products	0.661	0.015
Manufacture of rubber and plastics products	0.665	0.024
Manufacture of textiles	0.776	0.027
Manufacture of tobacco products	0.698	0.062
Manufacture of wearing apparel	0.643	0.037
Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials	0.656	0.047
Mining of coal and lignite	0.649	0.053
Mining of metal ores	0.719	0.037
Mining support service activities	0.594	0.061
Motion picture, video and television programme production, sound recording and music publishing activities	0.679	0.025
Office administrative, office support and other business support activities	0.647	0.029
Other manufacturing	0.548	0.025
Other mining and quarrying	0.711	0.029
Other personal service activities	0.548	0.048
Other professional, scientific and technical activities	0.612	0.038
Postal and courier activities	0.813	0.089
Printing and reproduction of recorded media	0.720	0.007
Programming and broadcasting activities	0.586	0.047
Public administration and defence; compulsory social security	0.593	0.206
Publishing activities	0.549	0.026
Real estate activities	0.493	0.044
Remediation activities and other waste management services	0.717	0.076
Rental and leasing activities	0.515	0.033
Repair and installation of machinery and equipment	0.681	0.022
Repair of computers and personal and household goods	1	0
Residential care activities	0.588	0.041
Retail trade, except of motor vehicles and motorcycles	0.500	0.021
Scientific research and development	0.569	0.038
Security and investigation activities	0.688	0.033
Services to buildings and landscape activities	0.508	0.045

Sewerage	0.483	0.059
Social work activities without accommodation	0.680	0.096
Specialized construction activities	0.592	0.037
Sports activities and amusement and recreation activities	0.689	0.044
Telecommunications	0.690	0.019
Travel agency, tour operator, reservation service and related activities	0.621	0.045
Veterinary activities	0.972	0.083
Warehousing and support activities for transportation	0.690	0.031
Waste collection, treatment and disposal activities; materials recovery	0.651	0.034
Water collection, treatment and supply	0.603	0.057
Water transport	0.577	0.030
Wholesale and retail trade and repair of motor vehicles and motorcycles	0.658	0.019
Wholesale trade, except of motor vehicles and motorcycles	0.575	0.013

Table A.1: Time average of the percentage of firms that disclose labour costs per sector.

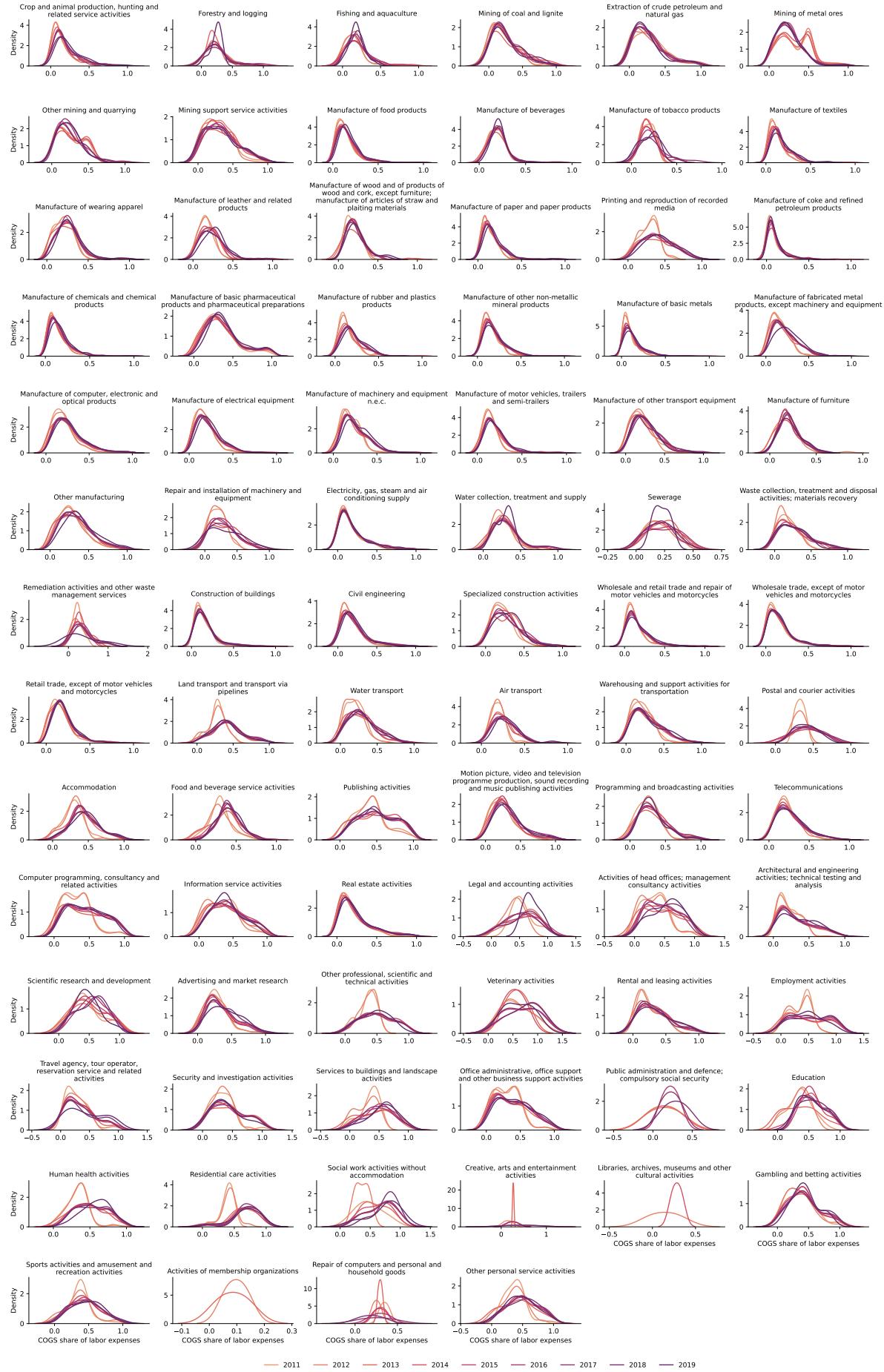


Figure A.1: Density of the cost of goods sold share of labour expenses for firms in different sectors and over time (2011-2019). The titles show the NACE Rev. 2 sectoral classification at the 2-digit level.

Let $w_{i,t}$ be labour expenses of firm i at time t and $g_{i,t}$ its cost of goods sold excluding depreciation and amortisation (for firms employing a *by nature* method in writing their income statements), then the cost of goods sold share of labour expenses of firm i at time t is defined as

$$\alpha_{i,t} = \frac{w_{i,t}}{g_{i,t} + w_{i,t}}. \quad (\text{A.1})$$

For some firms, we observed labour expenses for some years but not for other years. Thus, we assumed that their $\alpha_{i,t}$ does not change and kept it constant over time.

Estimation method 1. Firstly, we calculate a firm's average cost of goods sold share of labour expenses over a three-year period starting from 2013. For example, $\alpha_{i,2013}$ is computed using a rolling window going from 2011 to 2013, while for $\alpha_{i,2014}$ the rolling window goes from 2012 to 2014.¹³ A firm's average cost of goods sold share of labour expenses is given by

$$\bar{\alpha}_{i,t} = \frac{1}{3} \sum_{\tau=t-2}^t \alpha_{i,\tau}.$$

Afterwards, for each year t , we compute sector s ' cost of goods sold share of labour expenses by averaging over the α_i 's of the firms in sector s at time t :

$$\tilde{\alpha}_{s,t} = \frac{1}{N_{s,t}} \sum_{i \in \mathcal{S}} \bar{\alpha}_{i,t},$$

where $N_{s,t}$ is the number of firms in sector s at time t and \mathcal{S} is the set of firms in sector s .

Estimation method 2. We calculate a sector-level time-varying cost of goods sold share of labour expenses in the following two ways.

- 2.a** We compute a sector's cost of good sold share of labour expenses for each time step t by averaging over the α_i 's of firms in sector s :

$$\alpha_{s,t}^a = \frac{1}{N_{s,t}} \sum_{i \in \mathcal{S}} \alpha_{i,t}.$$

- 2.b** We calculate a sector's cost of goods sold share of labour expenses by summing the labour expenses of firms in sector s at time t and dividing these by the sum of firms' labour and intermediate costs:

$$\alpha_{s,t}^b = \frac{\sum_{i \in \mathcal{S}} w_{i,t}}{\sum_{i \in \mathcal{S}} w_{i,t} + \sum_{i \in \mathcal{S}} g_{i,t}}.$$

After using either of the two methods above (2.a or 2.b), to compute the sector-level cost of goods sold share of labour expenses at time t , we employ a three-year rolling window to calculate the time average:

$$\bar{\alpha}_{s,t}^k = \frac{1}{3} \sum_{\tau=t-2}^t \alpha_{s,\tau}^k,$$

for $k = \{a, b\}$, indicating whether we used method 2.a or 2.b in the first step.

Estimation method 3. We start from $\alpha_{s,t}^a$ calculated using method 2.a and then compute a weighted average using a three-year rolling window:

$$\bar{\alpha}_{s,t}^* = \sum_{\tau=t-2}^t \theta_{s,\tau} \alpha_{s,\tau}^a,$$

the $\theta_{s,t}$'s are weights that sum up to one. $\theta_{s,t}$ is the share of sector s at time t in the total expenditure of the sector on labour and intermediate inputs during the three years. $\theta_{s,t}$ is defined as:

¹³For some firms, picked at random, we look at how their α 's change over time; changes are negligible.

$$\theta_{s,t} = \frac{\sum_{i \in \mathcal{S}} g_{i,\tau} + w_{i,\tau}}{\sum_{\tau=t-2}^t \sum_{i \in \mathcal{S}} g_{i,\tau} + w_{i,\tau}}.$$

Assessing the methods. Figure A.2 shows the empirical against the predicted cost of goods sold share of labour expenses for our four estimation methods. They all seem to perform very similarly except for method 2b, which has the lowest median absolute error but the highest root mean squared error and median absolute error. Among the others, method 2a and 3 have the same and lowest error metrics (Table A.2).

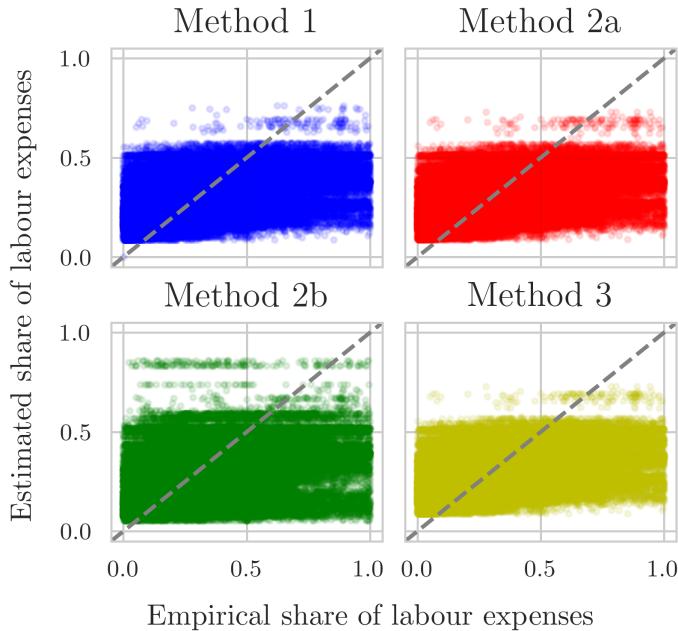


Figure A.2: Empirical (x -axis) and predicted (y -axis) cost of goods sold share of labour expenses for our four estimation methods.

	RMSE	MAE	MedAE
Method 1	0.184	0.138	0.106
Method 2a	0.184	0.137	0.105
Method 2b	0.201	0.144	0.098
Method 3	0.184	0.137	0.105

Table A.2: Error metrics for the cost of goods sold share of labour expenses for the four estimation methods we devised. RMSE denotes the root mean squared error, MAE the mean absolute error and MedAE the median absolute error.

Figure A.3 shows the empirical against the predicted labour expenses (top) and intermediate expenses (bottom) for our four estimation methods. They all seem to perform very similarly. Method 2b has the lowest error metrics (Table A.3). The error metrics are the same for the labour and intermediate expenses because they mirror each other. Had we not taken the absolute values, for instance, the error metrics would have opposite sign.

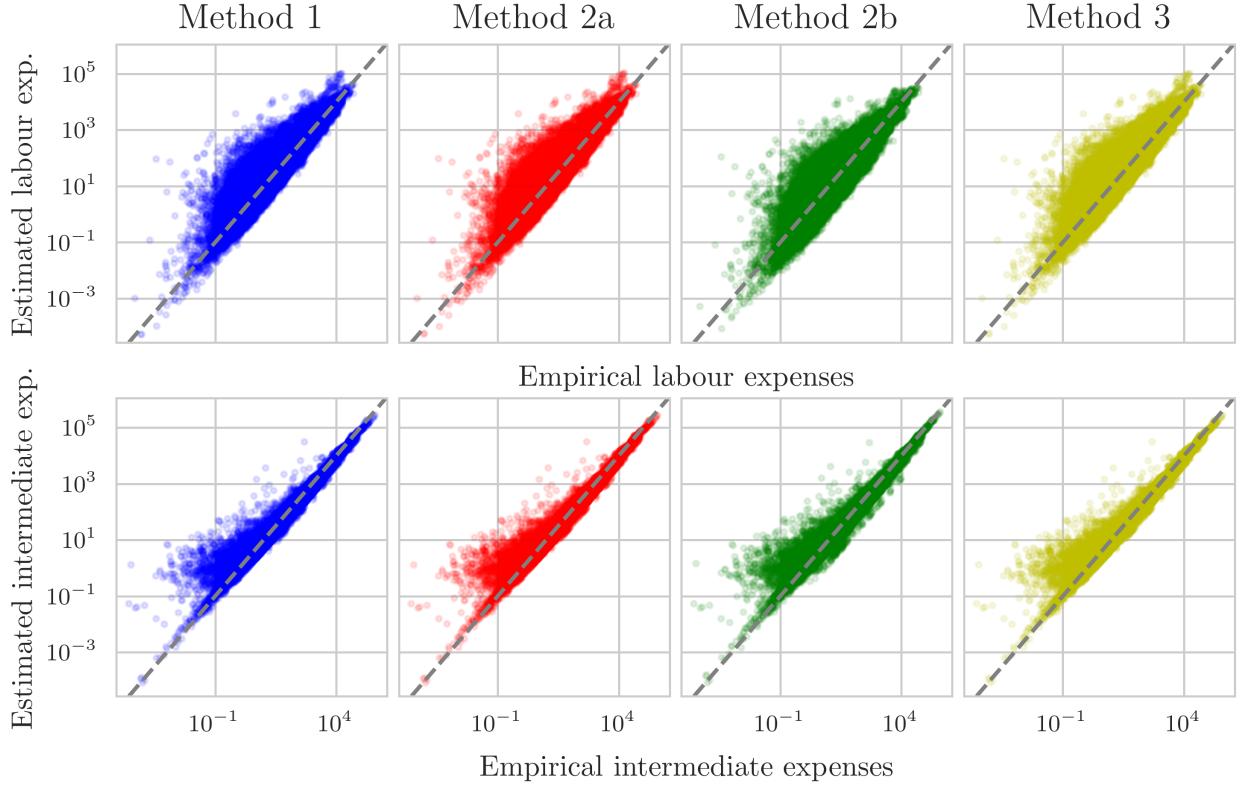


Figure A.3: Empirical (x -axis) and predicted (y -axis) labour expenses (top) and intermediate expenses (bottom) for our four estimation methods.

	Labour exp.			Intermediate exp.		
	RMSE	MAE	MedAE	RMSE	MAE	MedAE
Method 1	1,153.6	126.0	8.5	1,153.6	126.0	8.5
Method 2a	1,163.0	126.1	8.5	1,163.0	126.1	8.5
Method 2b	497.0	90.2	8.4	497.0	90.2	8.4
Method 3	1,162.0	126.1	8.5	1,162.0	126.1	8.5

Table A.3: Error metrics for labour and intermediate expenses for the four estimation methods we devised. RMSE denotes the root mean squared error, MAE the mean absolute error and MedAE the median absolute error.

A.1.3 Value-added

Value-added can be calculated using income statements. It is defined in different ways in the literature. For instance, [Yang et al. \(2019\)](#) calculate it as the sum of deflated wages and deflated earnings before interest and taxes. [Magerman et al. \(2016\)](#) calculate value-added as the difference between sales and material costs, which is standard in the production network literature (at the firm level). We adopt a third strategy based on [Miller and Blair's \(2009\)](#) book on input-output analysis, which also corresponds to the definition in the WIOD (which we use to construct the proxy node).

Following [Miller and Blair \(2009\)](#), valued-added is the wage bill (labour costs) plus EBITDA (earnings before interest and taxes). Their definition, which follows national and international accounting standards, also includes amortisation and depreciation. We do not observe the wage bill for all companies since labour expenses are disclosed depending on the method followed in writing the income statement. Therefore, for some companies, we estimate labour costs as explained in Section A.1.2.

A.1.4 Cleaning financial statements for the construction of the I-O table

We kept firms with positive sales, intermediate expenses and value-added, and non-negative labour costs (all firms with zero labour costs have disclosed them), EBITDA and amortisation and depreciation.

Sanity check. We checked that firms in the dataset respect the accounting identity:

$$\text{sales} = \text{intermediate_expenditure} + \text{value_added} + \text{other_costs}.$$

Since we did not account for other costs, we expected the residual (other_costs) to be non-negative. However, in 2014, for 30% of the firms, the residual is negative, meaning that sales were smaller than the sum of the variables on the RHS (excluding other_costs).

Most firms with a negative residual disclosed their labour expenses: is it possible that we double-counted labour expenses? This would mean that a firm adopts a *by function* method to write the income statement but reports labour costs in a footnote, which would lead FactSet to record the labour expenses. If this were the case, labour costs would already be in the cost of goods sold, leading to counting this expense twice. Thus, we subtract the disclosed labour costs from intermediate expenses for these firms. Is this reasonable? We checked whether this subtraction caused any firm to have negative intermediate expenses. It did, but only for 0.5% of the observations for which we did the procedure.

A positive consequence of the exercise just described is that it allows us to identify firms that likely adopt a *by function* method to write the income statement. We propagate this information back in our estimation of the cost of goods sold share of labour expenses in order to adjust the denominator in Equation A.1, which for these firms would otherwise be double counting labour costs. After doing this procedure only 1.2% of the firms in 2014 do not respect the accounting identity. We drop these firms. The cleaning procedure led to all firms in the sample having a strictly positive value-added.

A.2 FactSet: evaluation of coverage

First, we show what types of firms are in FactSet. Second, we assess the coverage of global economic activity in FactSet by comparing it to the WIOD. We compare the countries covered and the sectoral composition in FactSet against that of the WIOD. Subsequently, we quantify the percentage of world gross output captured by FactSet and then evaluate the growth rates of gross output.

A.2.1 Firm types

Figure A.4 shows the percentage of firms for each firm type as defined by FactSet. We plot data for the year 2014 (there are 5,442 firms). In the dataset, there are some big private companies, which also file financial statements, and companies that are extinct by April 2020, when the dataset was downloaded.

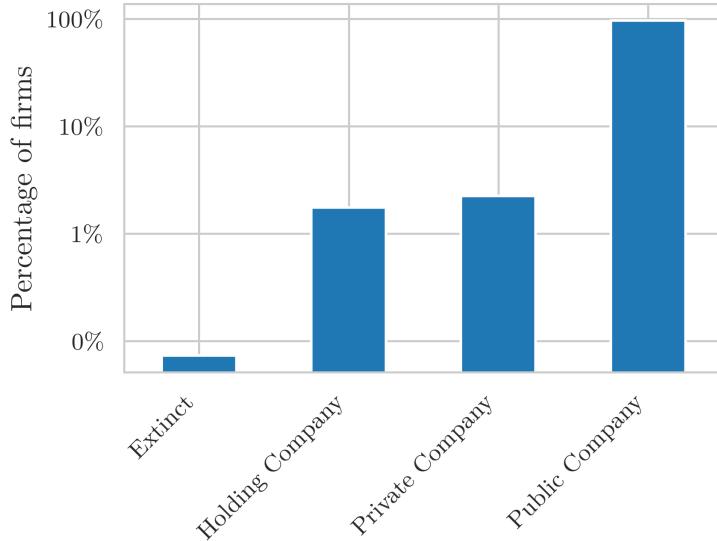


Figure A.4: Percentage of firms per type in FactSet in 2014. The firm type is defined by FactSet.

A.2.2 Countries covered in WIOD and FactSet

Table A.4 shows the countries that are in the WIOD and the countries where firms in FactSet have their headquarters for the period 2013–2019. The WIOD has a node named “rest of the world” (RoW), which captures all the other countries not in the list. The number of overlapping countries is 41. The Czech Republic and Estonia are covered in the WIOD but not in FactSet.

WIOD (43 countries)	FactSet (87 countries)
Australia, Austria, Belgium, Brazil, Bulgaria, Canada, China, Croatia, Cyprus, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, India, Indonesia, Ireland, Italy, Japan, Korea (the Republic of), Latvia, Lithuania, Luxembourg, Malta, Mexico, Netherlands (the), Norway, Poland, Portugal, Romania, Russian Federation (the), Slovakia, Slovenia, Spain, Sweden, Switzerland, Taiwan (Province of China), Turkey, United Kingdom of Great Britain and Northern Ireland, United States of America (the).	Argentina, Australia, Austria, Bahamas (the), Bahrain, Bangladesh, Belgium, Bermuda, Brazil, Bulgaria, Canada, Cayman Islands (the), Chile, China, Colombia, Costa Rica, Croatia, Cyprus, Côte d'Ivoire, Denmark, Egypt, Faroe Islands (the), Finland, France, Germany, Greece, Hong Kong, Hungary, Iceland, India, Indonesia, Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kenya, Korea (the Republic of), Kuwait, Latvia, Lithuania, Luxembourg, Macao, Malaysia, Malta, Marshall Islands (the), Mauritius, Mexico, Monaco, Morocco, Netherlands (the), New Zealand, Nigeria, Norway, Oman, Pakistan, Panama, Peru, Philippines (the), Poland, Portugal, Qatar, Republic of North Macedonia, Romania, Russian Federation (the), Saudi Arabia, Singapore, Slovakia, Slovenia, South Africa, Spain, Sri Lanka, Sweden, Switzerland, Taiwan (Province of China), Thailand, Trinidad and Tobago, Tunisia, Turkey, Ukraine, United Arab Emirates (the), United Kingdom of Great Britain and Northern Ireland (the), United States of America (the), Venezuela (Bolivarian Republic of), Vietnam, Zambia

Table A.4: Countries covered by the WIOD and FactSet.

A.2.3 Sectoral composition

Figure A.5 shows the sectoral composition in the WIOD (black bars) and in FactSet aggregated at the sector level (green bars) using firms’ revenues. We assess the sectoral composition using the sectors’ gross output shares (or revenues for firms). We group ISIC Rev.4 codes at the 1-digit level into eight higher-level classes shown in Table A.5.

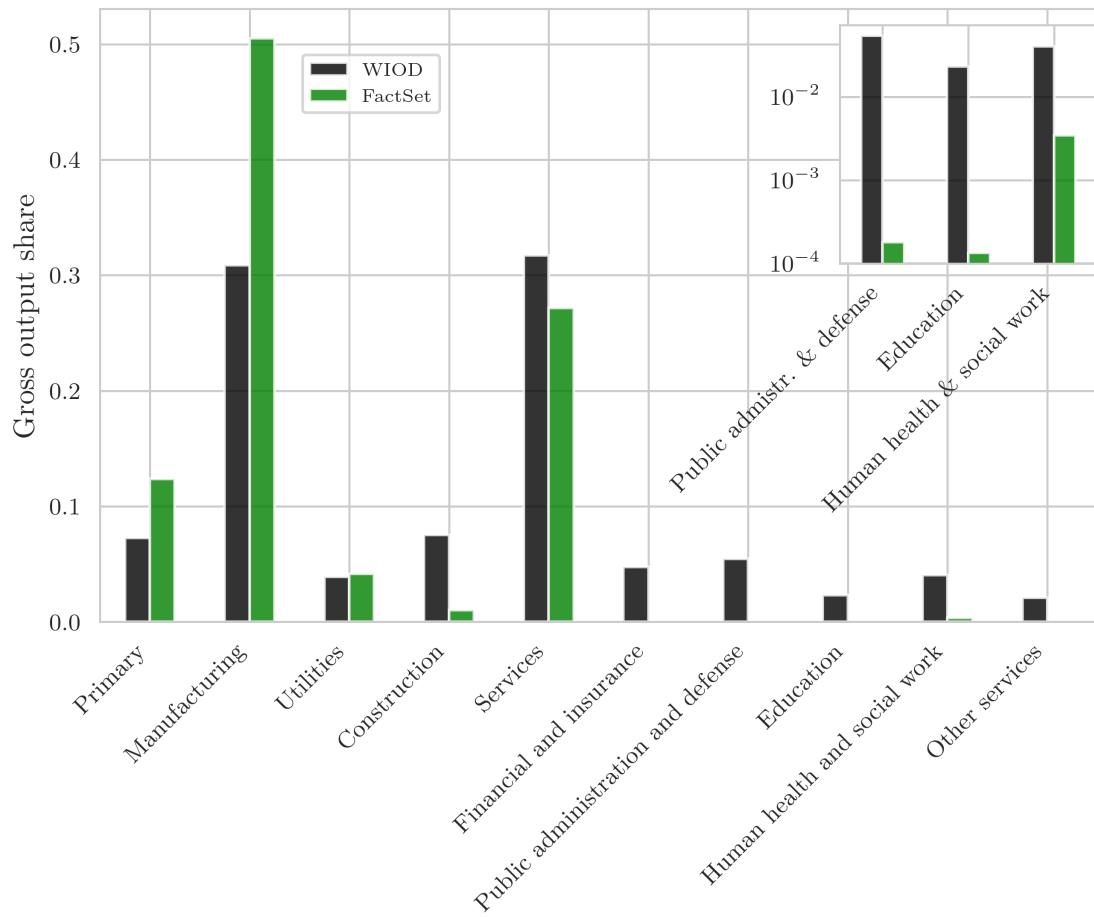


Figure A.5: Sectoral composition (gross output shares) of the WIOD (black) and FactSet (green) in 2014. For FactSet, we group firms into sectors and use firms' revenues. Sectors are defined as in Table A.5, where we aggregate ISIC Rev. 4 codes at the 1-digit level into macro classes.

ISIC	Description	Macro class
A	Agriculture, forestry and fishing	Primary
B	Mining and quarrying	Primary
C	Manufacturing	Manufacturing
D	Electricity, gas, steam and air conditioning supply	Utility
E	Water supply; sewerage, waste management and remediation activities	Utility
F	Construction	Construction
G	Wholesale and retail trade; repair of motor vehicles and motorcycles	Services
H	Transportation and storage	Services
I	Accommodation and food service activities	Services
J	Information and communication	Services
K	Financial and insurance activities	Financial and insurance
L	Real estate activities	Services
M	Professional, scientific and technical activities	Services
N	Administrative and support service activities	Services
O	Public administration and defence; compulsory social security	Public administration and defence
P	Education	Education
Q	Human health and social work activities	Human health and social work
R	Arts, entertainment and recreation	Other services
S	Other service activities	Other services
T	Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use	Not included
U	Activities of extraterritorial organisations and bodies	Not included

Table A.5: Description of ISIC Rev. 4 codes at the 1-digit (left column) and our macro-level classes (right column).

A.2.4 Evaluation of gross output

To understand how well FactSet captures global economic activity, we compare gross output levels and growth rates in FactSet to national accounting data (WIOD). We show the evaluation from 2013 to 2018, although we only use 2014.

To compare with world economic activity, we use gross output from the WIOD. The last year for which the WIOD is available is 2014; therefore, we forecast world gross output from 2015 to 2018.¹⁴ We assess both the levels and growth rates of gross output.

Levels. A comparison of the time series of aggregate firms' revenues and world gross output is shown in Figure A.6a. FactSet captures, on average, 16.4% of world gross output over time. The yearly percentage of world gross output captured by FactSet is given by

$$\phi_t = \frac{\sum_i q_{i,t}}{q_t}, \quad (\text{A.2})$$

where q_t is world gross output at time t and $q_{i,t}$ is firm i 's sales at time t . When forecasting world gross output, besides our central estimate, we also calculate a best and worst case; these yield a lower bound of 15.1% and an upper bound of 18.3% on the central estimate.

¹⁴ The WIOD is available from 2003 to 2014 but our firm-level dataset is available from 2014 to 2020. We forecast world gross output from 2015 until 2018 using GDP from the World Bank as follows. We take the ratio of gross output to GDP, which is known to be fairly stable over time, and assume that after 2014 this ratio stays constant. This gives us gross output q_t as a function of GDP y_t and the gross output/GDP ratio ζ_t , thus $q_t = \zeta_{2014} \cdot y_t$ for all $t = 2015, \dots, 2020$.

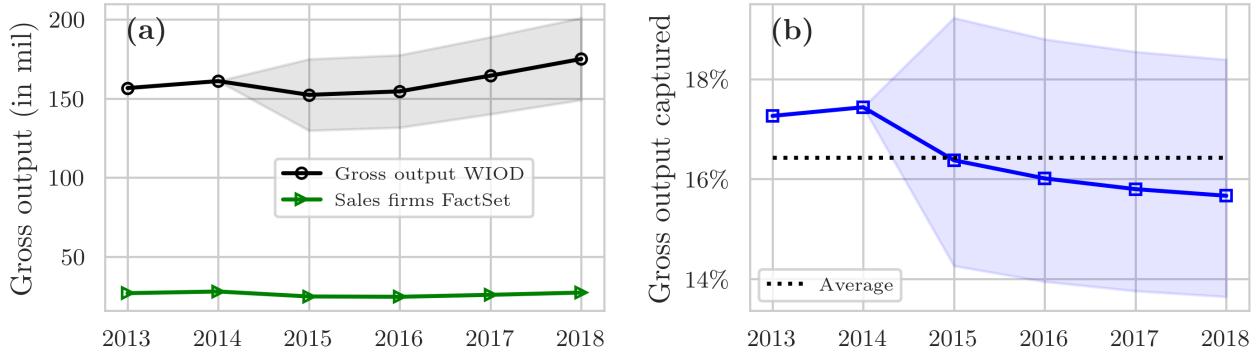


Figure A.6: (a) World gross output in the WIOD (black dotted line) and cumulative revenues of firms in FactSet (green line with triangles). Values in millions of US dollars. We show error bars for the years for which we forecast gross output (2015–2018); see Footnote 14 for a description of the forecasting methodology. (b) Percentage captured by FactSet of world gross output (WIOD, Equation A.2). The shaded blue area shows error bounds related to the gross output forecasts. The dashed black line marks the time average.

Growth rates. For the growth rate of gross output, we carry out the same assessment described above for the levels of gross output. Figure A.7a shows the growth rate of gross output in our aggregated FactSet dataset (green line with triangles) and in the WIOD (black line with circles). Over time, FactSet's growth rates differ from world growth rates by 2.3 percentage points on average. To assess how much the growth rates differ, we used the following metric

$$\phi_t = |g(\tilde{q}_t) - g(q_t)|, \quad (\text{A.3})$$

where $g(q_t)$ denotes the growth rate of world gross output at time t and the tilde indicates our aggregate variable constructed from firms' revenues.

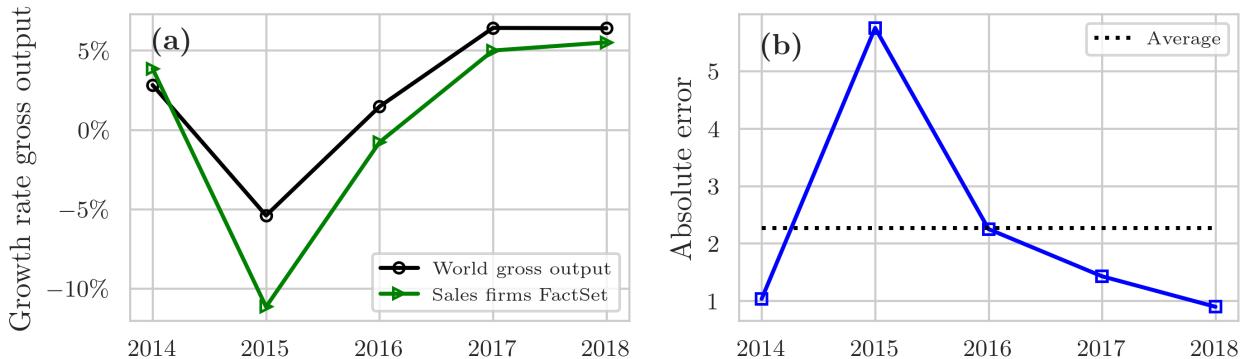


Figure A.7: (a) Growth rate of world gross output (WIOD, black line with circles) and the growth rates of firms' revenues in FactSet (green line with triangles). As in Figure A.6, we forecast world gross output from 2015 to 2018. (b) Absolute error between the growth rate of world gross output (WIOD) and the growth rate of firms' revenues (FactSet). The y -axis shows percentage points and the dotted black line shows the time average.

A.3 Merging FactSet with sector-level I-O tables (the WIOD)

Table A.6 describes the variables and associated notation used.

Notation	Description
q_i	firm i 's total sales
q_s	sector s ' gross output
d_i	firm i 's intermediate sales
d_s	sector s ' intermediate sales
x_i	firm i 's intermediate expenditure
x_s	sector s ' intermediate expenditure
f_i	firm i 's final demand
f_s	sector s ' final demand
k_s	sector s ' gross fixed capital formation
Δn_s	changes in inventories and valuables of sector s
w_i	firm i 's labour costs
y_i	firm i 's value-added
y_s	sector s ' value-added
τ	taxes minus subsidies on products of sector s
e_s	CIF and FOB adjustments on exports of sector s
p_s	direct purchases abroad by residents and purchases on the domestic territory by non-residents for sector s
u_s	international transport margins of sector s
$Z_{r,s}$	the amount of intermediate inputs sector s buys from sector r

Table A.6: Notation and terminology.

Since FactSet does not cover the entire economy, we supplement it with one proxy node constructed from input-output data at the sector level (we use the WIOD). I-O tables at the sector level capture economic linkages between industries in a country and sometimes also across countries – clearly, these data are at a more aggregated level than supply chain networks at the firm level. I-O tables at the sector level provide information on the monetary flows related to inputs or consumption expenditure from each industry to each of the other industries, itself and other agents in the economic system such as households and the government sector.

Figure A.8 shows the structure of the WIOD (the dataset we used), which is standard in aggregate I-O datasets based mostly on national accounts. On the rows, there are the supplying industries grouped by country and on the columns are the customers. Thus, each column portrays an industry production “recipe” and the row its customer base. The central block, which involves only trades among industries, is called the inter-industry transaction matrix or intermediate consumption. On its right, there are final uses or final demand, which in the WIOD comprises both consumption and investment demand. At the bottom, there is value-added. The 2016 version of the WIOD covers 43 countries and 56 sectors from 2000 to 2014.

			Use by country-industries						Final use by countries			Total use	
			Country 1		...	Country M		Country 1	...	Country M			
			Industry 1	Industry N	...	Industry 1	...	Industry N		
Supply from country-industries	Country 1	Industry 1											
		...											
		Industry N											
												
	Country M	Industry 1											
		...											
		Industry N											
Value added by labour and capital													
Gross output													

Figure A.8: Example of a multi-country I-O table; here we show the World Input-Output Database. *Source: Timmer et al. (2015, p.577).*

We construct the I-O table using FactSet and supplement it with a synthetic node constructed using the WIOD. The synthetic, or proxy, node is connected to all the firms (both incoming and outgoing links). The synthetic node accounts for the production, sales and value-added of firms that are not in our network (i.e., the rest of the global economy). While the WIOD is at the country-sector level, we aggregate it into only one node for two reasons. Firstly, the firms in our supply chain data are multinational companies aggregated at the group level, as such their production is scattered around the world, making it difficult to attribute their sales, expenditure and value-added to a specific country. Secondly, while we tried to keep the sector dimension – which would have allowed us to have one proxy node per sector – differences between national accounting standards (WIOD) and financial accounting standards (FactSet) prevented us from accurately constructing an I-O table at the firm-sector level. We briefly outline the differences between national accounts and firms' financial accounts in what follows; for a more in-depth discussion, we refer to [Bacilieri et al. \(2023\)](#).

As shown in Figure A.8, an I-O table is composed of three main parts: the central block (inter-firm or inter-industry transaction matrix), final uses (or final demand) and value-added. The structure of the I-O table implies that there are four main variables to construct and harmonise with our firm-level data when introducing the proxy sector (we discuss this in more detail in the following section, Appendix A.3.1). These four variables are total and intermediate sales and costs, value-added and final demand. We build these variables for the proxy sector by subtracting from the WIOD totals the observed firms' cumulative values (for the firms in the cleaned FactSet dataset). While these four variables are relatively easy to identify in sector-level I-O tables, labour expenses and final demand pose challenges in their exact quantification at the firm level.

A.3.1 Accounting identities in the WIOD

Within a year, the WIOD table is at the country-sector level. Even though we aggregate over countries and sectors, it is helpful to understand the accounting identities underlying it. We show accounting identities by omitting the country and time index, but identities hold for country-sector tables and for one aggregated sector in the same way. Although for one sector, the inter-industry matrix becomes somewhat meaningless.

From the use side, the accounting identity used in the WIOD is

$$q_s = \sum_r Z_{sr} + f_s + k_s + \Delta n_s , \quad (\text{A.4})$$

which states that gross output of sector s is the sum of intermediate sales of sector s , its final demand f_s , gross fixed capital formation k_s (investment) and changes in inventories Δn_s .

From the expenditure side, the accounting identity is

$$q_s = \sum_r Z_{rs} + \tau_s + e_s + p_s + y_s + u_s , \quad (\text{A.5})$$

which states that gross output of sector s (from the expenditure side) equals total expenditure on intermediate inputs, plus taxes minus subsidies τ_s , plus cost, insurance and freight (CIF) and free on board (FOB) adjustments on exports e_s , value-added y_s and international transport margins u_s .

To construct the proxy sector, the variable definitions in the sector-level I-O table need to match the definition of the variables at the firm level taken from firms' financial statements. We discuss how we harmonise the two datasets in the following two sections.

A.3.2 Expenditure-side harmonisation

In this section, we define and describe how we harmonised two main variables that make up the I-O table from the expenditure side: intermediate expenditure and value-added. We do not account for direct purchases abroad by residents and purchases on the domestic territory by non-residents since these are zero for industries.

Expenditure on intermediate inputs. From the expenditure side, we define *sectoral intermediate expenditure* to include also CIF and FOB costs, and international transport margins since these are

costs that if firms incur into are in the cost of goods sold:

$$x_s = \sum_r Z_{rs} + e_s + u_s . \quad (\text{A.6})$$

Firms' intermediate consumption is defined as expenditure on intermediate inputs and services as stated in their invoices and reported as a cumulative figure in their income statements in the cost of goods sold. The cost of goods sold may include labour costs and thus needs to be cleaned for some firms depending on the method they use to write their income statement (see Appendix A.1.1 and A.1.2).

Value-added. Sectoral *gross value-added* is defined by the System of National Accounts (SNA) as labour and capital compensation, consumption of fixed capital and taxes net of subsidies ([United Nations, 2010](#)). The WIOD provides taxes net of subsidies as a separate variable, which thus needs to be added to the definition of value-added.¹⁵ Sectoral gross value-added is given by

$$y_s^{\text{gross}} = \tau_s + y_s . \quad (\text{A.7})$$

Firms do not disclose the subsidies they received as a separate line item; therefore, we did not include subsidies and simply added taxes (a definition of firms' value-added is provided in Appendix A.1.3).

A.3.3 Use-side harmonisation

In this section, we define and describe how we harmonise the four variables that make up the I-O table from the use side, namely final demand, intermediate sales, gross fixed capital formation and changes in inventories.

While it is more straightforward to reconcile prices on the expenditure side, it is less so on the use side. On the use side, prices are at the purchaser's price; that is, the amount paid by the purchaser plus trade margins (wholesale and retail), transport margins (if invoiced by the producer) and non-deductible VAT minus deductible VAT. However, at the firm level output (i.e., revenues) is the amount received by the producer for the good or service sold minus VAT (deductible and non) and subsidies. While transport and trade margins may be in the firm's revenues, VAT and subsidies are not. We cannot take care of these mismatches because we lack the data to do so.

Final demand. Sectoral *final demand* is defined as in the WIOD and it comprises the consumption of households, the government and non-profit organisations.

Firms do not disclose sales to final demand nor sales to other firms; hence revenues include sales to other firms in the network and sales to final demand. To infer how much a firm sells to final demand, we use the share of final demand satisfied by the sector the firm is in. More formally, let q_s be gross output of sector s , f_s be final demand of sector s and q_i be firm i 's revenues (or total sales), then final demand of firm i is given by

$$f_i = q_i \frac{f_s}{q_s} . \quad (\text{A.8})$$

Gross fixed capital formation. Gross fixed capital formation (GFCF) is defined in the SNA as "the value of [...] acquisitions less disposals of fixed assets" ([United Nations, 2010](#)); it is a measure of investment. In the WIOD, GFCF corresponds to the part of a sector's output that ends up as an investment. It also includes some intangible assets when they are part of the SNA, but not all intangibles are covered.

In firm-level data, firms making capital goods disclose their customers and sales of capital goods are accounted for in the disclosed revenues. However, we cannot distinguish these types of transactions.

¹⁵The variable capital compensation that makes up value-added in the WIOD comprises both profit and consumption of fixed capital. It is a residual variable after subtracting labour compensation from value-added. Moreover "it is the gross compensation for capital, including profits and depreciation allowances. Because of its derivation as a residual, it reflects the remuneration for capital in the broadest sense. This does not include only traditional reproducible assets such as machinery and buildings but also includes non-reproducible assets. Examples are mineral resources and land, intangible assets (such as R&D knowledge stocks, software, databases, brand names and organisational capital) and financial capital" ([Timmer et al., 2015](#), p. 601).

Therefore, we apply a correction factor to a firm's revenues similar to Magerman et al. (2016). As for final demand, we estimated a firm's GFCG as follows

$$k_i = q_i \frac{k_s}{q_s} , \quad (\text{A.9})$$

where k is gross fixed capital formation either of firm i or sector s . A firm's expenditure on capital goods is not included in the cost of goods sold, thus no adjustment is needed on the expenditure side.

In the WIOD, the ratio k_s/q_s is negative for ISIC code E37-E39 “sewerage; waste collection, treatment and disposal activities; materials recovery; remediation activities and other waste management services”. We assume that firms in this sector do not provide investment goods and hence have zero GFCF.

Intermediate sales. Sectoral *intermediate sales* are sales of intermediate inputs and services to other sectors in the economy. We use the same definition for firms, with the caveat that to get intermediate sales, we have to subtract from revenues the sales to final demand and GFCF.

Changes in inventories. We exclude changes in inventory since, in the WIOD, it is a column used for adjustments. Suppose we were to include changes in inventories. In that case, it should go with intermediate expenditure since the variable to which inventory changes correspond to at the firm level is the cost of goods sold, which is defined as beginning inventory minus ending inventory plus purchases during the period.

A.4 Sectoral codes

Table A.7 and A.8 describe, respectively, NACE Rev.2 codes at the section level (1-digit) and ISIC Rev. 4 codes at the 2-digit level.

Section (1-digit)	Description	Divisions (2-digit)
A	Agriculture, forestry and fishing	01 – 03
B	Mining and quarrying	05 – 09
C	Manufacturing	10 – 33
D	Electricity, gas, steam and air conditioning supply	35
E	Water supply; sewerage, waste management and remediation activities	36 – 39
F	Construction	41 – 43
G	Wholesale and retail trade; repair of motor vehicles and motorcycles	45 – 47
H	Transportation and storage	49 – 53
I	Accommodation and food service activities	55 – 56
J	Information and communication	58 – 63
K	Financial and insurance activities	64 – 66
L	Real estate activities	68
M	Professional, scientific and technical activities	69 – 75
N	Administrative and support service activities	77 – 82
O	Public administration and defence; compulsory social security	84
P	Education	85
Q	Human health and social work activities	86 – 88
R	Arts, entertainment and recreation	90 – 93
S	Other service activities	94 – 96
T	Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use	97 – 98
U	Activities of extraterritorial organisations and bodies	99

Table A.7: Description of NACE Rev.2 at the 1-digit (first column) and 2-digit level (last column).

ISIC code	Description
A01	Crop and animal production, hunting and related service activities
A02	Forestry and logging
A03	Fishing and aquaculture
B	Mining and quarrying
C10-C12	Manufacture of food products, beverages and tobacco products
C13-C15	Manufacture of textiles, wearing apparel and leather products
C16	Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials
C17	Manufacture of paper and paper products
C18	Printing and reproduction of recorded media
C19	Manufacture of coke and refined petroleum products
C20	Manufacture of chemicals and chemical products
C21	Manufacture of basic pharmaceutical products and pharmaceutical preparations
C22	Manufacture of rubber and plastic products
C23	Manufacture of other non-metallic mineral products
C24	Manufacture of basic metals
C25	Manufacture of fabricated metal products, except machinery and equipment
C26	Manufacture of computer, electronic and optical products
C27	Manufacture of electrical equipment
C28	Manufacture of machinery and equipment n.e.c.
C29	Manufacture of motor vehicles, trailers and semi-trailers
C30	Manufacture of other transport equipment
C31-C32	Manufacture of furniture; other manufacturing
C33	Repair and installation of machinery and equipment
D35	Electricity, gas, steam and air conditioning supply
E36	Water collection, treatment and supply
E37-E39	Sewerage; waste collection, treatment and disposal activities; materials recovery; remediation activities and other waste management services
F	Construction
G45	Wholesale and retail trade and repair of motor vehicles and motorcycles
G46	Wholesale trade, except of motor vehicles and motorcycles
G47	Retail trade, except of motor vehicles and motorcycles
H49	Land transport and transport via pipelines
H50	Water transport
H51	Air transport
H52	Warehousing and support activities for transportation
H53	Postal and courier activities
I	Accommodation and food service activities
J58	Publishing activities
J59-J60	Motion picture, video and television programme production, sound recording and music publishing activities; programming and broadcasting activities
J61	Telecommunications
J62-J63	Computer programming, consultancy and related activities; information service activities
K64	Financial service activities, except insurance and pension funding
K65	Insurance, reinsurance and pension funding, except compulsory social security
K66	Activities auxiliary to financial services and insurance activities
L68	Real estate activities
M69-M70	Legal and accounting activities; activities of head offices; management consultancy activities
M71	Architectural and engineering activities; technical testing and analysis
M72	Scientific research and development
M73	Advertising and market research
M74-M75	Other professional, scientific and technical activities; veterinary activities
N	Administrative and support service activities
O84	Public administration and defence; compulsory social security
P85	Education
Q	Human health and social work activities
R-S	Other service activities
T	Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use
U	Activities of extraterritorial organizations and bodies

Table A.8: Description of ISIC Rev. 4 at the 2-digit level.

B Derivations

B.1 The conditional maximum entropy reconstruction method of Parisi et al. (2020)

The maximum entropy method developed by [Parisi et al. \(2020\)](#) aims to reconstruct an ensemble of likely weighted networks that are consistent with prior information about the binary topology and that satisfy (in expectation) a set of constraints on observed network quantities – in our case, the in- and out-strengths. To reconstruct the ensemble of weighted networks, the method infers the probability distribution over the likely weighted network configurations by maximising the conditional (on the binary topology) Shannon entropy subject to imposed constraints.

The reconstruction procedure consists of two steps. The first step involves a deterministic maximum entropy procedure, known as *MaxEnt* in the literature, from which we derived a value for each link weight. Since MaxEnt has been found to be the best performing in reconstructing weights ([Anand et al., 2015; Lebacher et al., 2021](#)), the weights found in the first step are then used in the second step to constraint the expected value of each link weight.¹⁶ This procedure yields a computational benefit because, as we will see below, we have to solve m (the number of links) decoupled equations. Instead, had we imposed constraints on the in- and out-strengths, we would have had to solve $2N$ coupled equations, where N is the number of nodes (see [Parisi et al. \(2020\)](#) for the derivation).

First step. We derive a value for the expected weights using the MaxEnt method. MaxEnt is derived by solving a deterministic maximum entropy problem, which maximises an entropy-like functional subject to constraints on intermediate sales and costs of each firm:

$$\begin{aligned} \text{maximise}_{\{W_{ij}\}} \quad & S(\mathbf{W}) = - \sum_{ij} W_{ij} \log W_{ij} \\ \text{subject to} \quad & \sum_j W_{ij} = s_i^{\text{out}*} \quad i = 1, \dots, N \\ & \sum_j W_{ji} = s_i^{\text{in}*} \quad i = 1, \dots, N, \end{aligned} \quad (\text{B.1})$$

which yields

$$W_{ij}^{\text{ME}} = \frac{s_i^{\text{out}*} s_j^{\text{in}*}}{W^{\text{tot}*}}, \quad (\text{B.2})$$

where W_{ij} is the money flow from firm j to firm i , $s_i^{\text{out}*}$ are the observed total intermediate sales (out-strength) of firm i , $s_i^{\text{in}*}$ are the observed total intermediate costs (in-strength) of firm i and $W^{\text{tot}*} = \sum_i s_i^{\text{out}*} = \sum_i s_i^{\text{in}*}$ is the total weight of the empirical network. Note that MaxEnt generates a fully connected network.

Second step. We solve the conditional maximum entropy problem to derive a value for the expected weights given that there is a link between i and j . Let $\mathbf{W} \in \mathbb{W}$ be a weighted adjacency matrix, which is a realisation of the random variable \mathcal{W} . Similarly, let $\mathbf{A} \in \mathbb{A}$ be a binary adjacency matrix, which is a realisation of the random variable \mathcal{A} . The method maximises the conditional entropy $S(\mathcal{W} | \mathcal{A})$ defined over the probability density function of the weighted network configurations compatible with the binary topology and satisfying the constraints on the expected weights given by the MaxEnt prescription (Equation B.2):

$$\begin{aligned} \text{maximise}_{\{Q(\mathbf{W}|\mathbf{A})\}} \quad & S(\mathcal{W} | \mathcal{A}) = - \sum_{\mathbf{A} \in \mathbb{A}} P(\mathbf{A}) \int_{\mathbb{W}_A} Q(\mathbf{W} | \mathbf{A}) \log Q(\mathbf{W} | \mathbf{A}) d\mathbf{W} \\ \text{subject to} \quad & \langle W_{ij} \rangle = \sum_{\mathbf{A} \in \mathbb{A}} P(\mathbf{A}) \int_{\mathbb{W}_A} W_{ij}(\mathbf{W}) Q(\mathbf{W} | \mathbf{A}) d\mathbf{W} = W_{ij}^{\text{ME}} \quad \forall (i, j) \in \mathcal{E} \\ & \int_{\mathbb{W}_A} Q(\mathbf{W} | \mathbf{A}) d\mathbf{W} = 1, \quad \forall \mathbf{A} \in \mathbb{A}, \end{aligned} \quad (\text{B.3})$$

¹⁶While we use the MaxEnt prescription, one could choose any value, including an observed value, when available. This value then needs to be subtracted from the in- and out-strengths.

where \mathbb{W}_A is the set of weighted configurations compatible with the binary topology, $Q(\mathbf{W} \mid \mathbf{A})$ is the conditional probability of generating a weighted network \mathbf{W} consistent with adjacency matrix \mathbf{A} and \mathcal{E} is the edge set. The information on the binary network enters in probabilistic terms through $P(\mathbf{A})$, the probability of observing the binary topology \mathbf{A} . $P(\mathbf{A})$ could be estimated, in a prior step using any suitable method. However, we supply the observed topological structure $\mathcal{A} = \mathbf{A}^*$, thus $P(\mathbf{A}^*) = 1$ and $P(\mathbf{A}) = 0, \forall \mathbf{A} \in \mathbb{A}, \mathbf{A} \neq \mathbf{A}^*$.

Solving the above maximisation problem yields the Hamiltonian $H(\mathbf{W}) = \sum_{i \neq j} \lambda_{ij} W_{ij}$ and the graph probability $Q(\mathbf{W} | \mathbf{A}) = \prod_{i \neq j} Q_{ij}(W_{ij} | A_{ij})$ with

$$Q_{ij}(W_{ij} | A_{ij}) = \begin{cases} \lambda_{ij} e^{-\lambda_{ij} W_{ij}} & W_{ij} > 0, \text{ if } A_{ij} = 1, \\ 0 & \text{otherwise.} \end{cases} \quad (\text{B.4})$$

Therefore, $Q_{ij}(W_{ij} | A_{ij} = 1)$ is of exponential form with parameter λ_{ij} , the Lagrange multiplier associated to the constraint on $\langle W_{ij} \rangle$. It follows that the generalised likelihood for the given set of constraints is

$$g(\boldsymbol{\lambda}) = - \sum_{i \neq j} \lambda_{ij} W_{ij}^{\text{ME}} + \sum_{i \neq j} \sum_{\mathbf{A} \in \mathbb{A}} P(\mathbf{A}) A_{ij} \log \lambda_{ij}, \quad (\text{B.5})$$

with first order conditions given by

$$\langle W_{ij} \rangle = \frac{p_{ij}}{\lambda_{ij}} = W_{ij}^{\text{ME}}, \quad \forall i \neq j, \quad (\text{B.6})$$

where $p_{ij} = \sum_{\mathbf{A} \in \mathbb{A}} P(\mathbf{A}) A_{ij}$ is the probability that a link going from i to j exists. Substituting the MaxEnt prescription (Equation B.2), the Lagrange multipliers are given by

$$\lambda_{ij}^* = p_{ij} \frac{W^{\text{tot}^*}}{s_i^{\text{out}^*} s_j^{\text{in}^*}}. \quad (\text{B.7})$$

Since we consider a deterministic binary topology, the Lagrange multipliers simplify to

$$\lambda_{ij}^* = \frac{W^{\text{tot}^*}}{s_i^{\text{out}^*} s_j^{\text{in}^*}} \quad \text{for } (i, j) \in \mathcal{E}. \quad (\text{B.8})$$

The MexEnt procedure assumes that the network is fully connected; however, the generalised maximum entropy method we employ assumes no self-loops. Additionally, we know where the links are. Therefore, we need to redistribute all the weights corresponding to $A_{ij} = 0, \forall (i, j) \notin \mathcal{E}$. To do so, we employ the IPF algorithm that redistributes the weights in an iterative procedure such that at the n -th iteration the weight is given by

$$W_{ij}^n = s_i^{\text{out}^*} \frac{W_{ij}^{(n-1)}}{\sum_{k \neq i} W_{ki}^{(n-1)}} \quad \text{and} \quad W_{ij}^{(n+1)} = s_j^{\text{in}^*} \frac{W_{ij}^n}{\sum_{k \neq j} W_{kj}^n}. \quad (\text{B.9})$$

Confidence interval on the expected link weight. One of the benefits of Parisi et al.'s method is that one can calculate a confidence interval $[w^-, w^+]$ for each of the expected weights. The lower bound is given by

$$w^- = - \frac{\ln[e^{-1} + q^-]}{\lambda_{ij}^*}, \quad (\text{B.10})$$

where q^- is a desired confidence level. The upper bound is given by

$$w^+ = - \frac{\ln[e^{-1} - q^+]}{\lambda_{ij}^*}. \quad (\text{B.11})$$

We refer to Appendix E in Parisi et al. (2020) for the whole derivation.

C Additional results

C.1 Weights

Ecuador. Figure C.1a shows the CCDF of the empirical weights of the full network (blue squares), the test network (green diamonds), the trimmed test networks (black dots) and that of the reconstructed networks (red triangles). All three weight distributions display heavy tails. Figure C.1b shows the power-law fit for one of the randomised networks. The distribution of the empirical weights has a power-law exponent of 1.1, while that of the reconstruction is higher and equal to 1.3. The cut-off point is also higher for the reconstructed weight distribution.

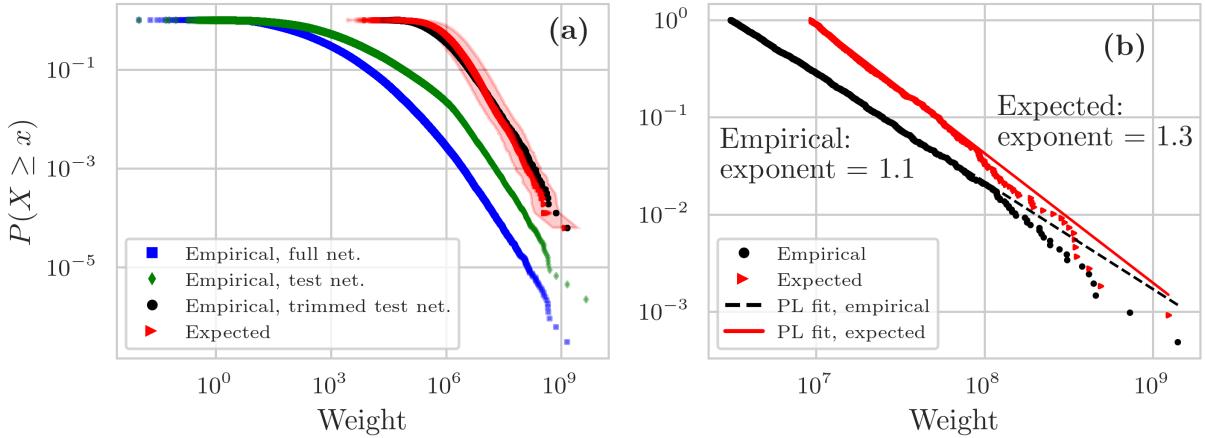


Figure C.1: (a) CCDF of the weights of the empirical full network (all firms in the Ecuadorian economy; blue squares), the empirical test network (green diamonds), the empirical trimmed test networks (black dots) and the reconstructed networks (red triangles); the shaded area indicates the 50% confidence interval. We show all 50 randomised test networks. (b) Power-law fit to the empirical (black dots) and expected (red triangles) weight distribution. The black dashed line shows the power-law fit to the empirical distribution, while the red solid line the power-law fit of the reconstructed distribution; we show results for one of the reconstructions.

The panels in the top row of Figure C.2 show the histograms of the relative prediction errors of Ecuador's weights for one of the randomised reconstructions. The top left panel shows the histogram of the non-positive error terms and the panel on the right the positive errors. The relative prediction error is defined as $\epsilon_i = (\mathbf{W}_{ij} - \mathbf{W}_{ij}^*) / \mathbf{W}_{ij}^*$, where \mathbf{W}_{ij}^* is the empirical value and \mathbf{W}_{ij} is the reconstructed weight. The mean relative prediction error is 285.7%. Such a high mean is driven by a big outlier, which is 674,065%; the median is 40%.

The panels in the bottom row of Figure C.2 show the empirical weights on the x -axis and the relative prediction errors on the y -axis; we plot non-positive errors on the left and positive errors on the right. The reconstruction method consistently under-predicts weights with values above $\sim 10^8$.

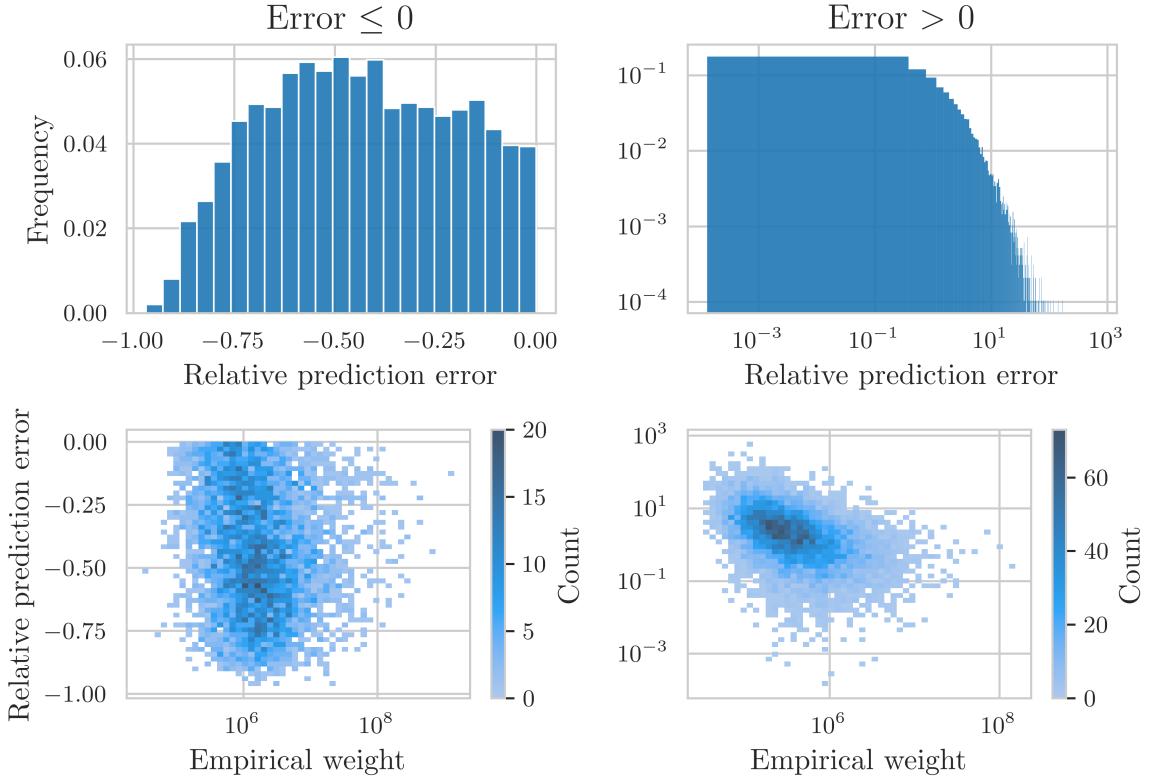


Figure C.2: **Top:** Histogram of the relative prediction errors of the weights for Ecuador. We plot separately the errors that are non-positive (left) and the errors that are positive (right). **Bottom:** Empirical weights (x -axis) against the relative prediction error (y -axis) for Ecuador. We plot the error terms that are non-positive on the left and those that are positive on the right. We divide both axes into 50 log-spaced bins and then count the number of data points falling in each square.

FactSet. Figure C.3 shows the CCDF of the expected weight distribution (blue squares) and its power-law fit (solid blue line) for FactSet. The estimated exponent equals 1.

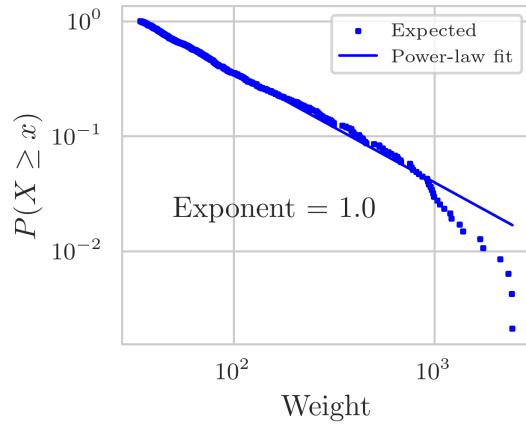


Figure C.3: CCDF of the expected weights (blue squares) and its power-law fit (solid blue line) for FactSet.

C.2 Technical and allocation coefficients

Ecuador. Figure C.4 shows the distribution of the empirical technical (left) and allocation coefficients (right) of the full network (blue squares), the test network (green diamonds), the trimmed test networks (black dots) and reconstructed networks (red triangles). The three distributions are right-skewed but do not display heavy tails.

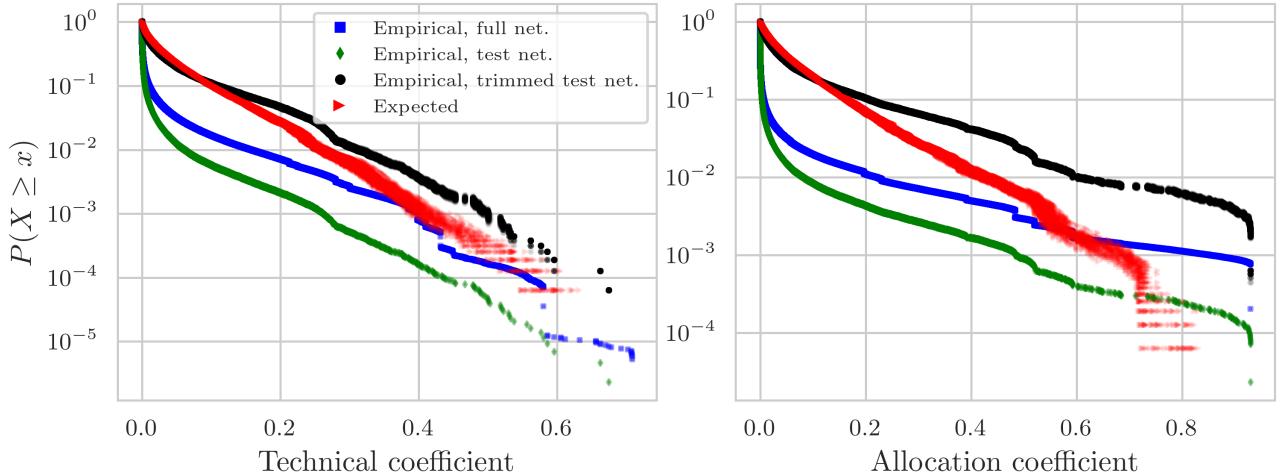


Figure C.4: CCDF of the technical (left) and allocation coefficients (right) of the empirical full network (blue squares), the empirical test network (green diamonds), the empirical trimmed test networks (black dots) and of the reconstructed networks (red triangles). We show all 50 randomised networks.

C.3 Multipliers

C.3.1 Influence vector

Ecuador. Figure C.5a shows the distribution of the empirical influence vector of the full network (blue squares), our test networks comprising 5,440 firms (black dots) and the reconstructed networks (red triangles). We compute the empirical influence vector for the firms in the (trimmed) test network using the network comprising all firms and then plot the distribution of only the 5,440 firms included in the (trimmed) test network. Therefore, the firms in the trimmed test network and the test network have the same influence. Figure C.5b shows the power-law fit for one of the reconstructions. The empirical distribution has a power-law exponent equal to 1.3, while the expected has a higher exponent, equal to 2. The CCDF of the expected influence vector also has a higher cut-off.

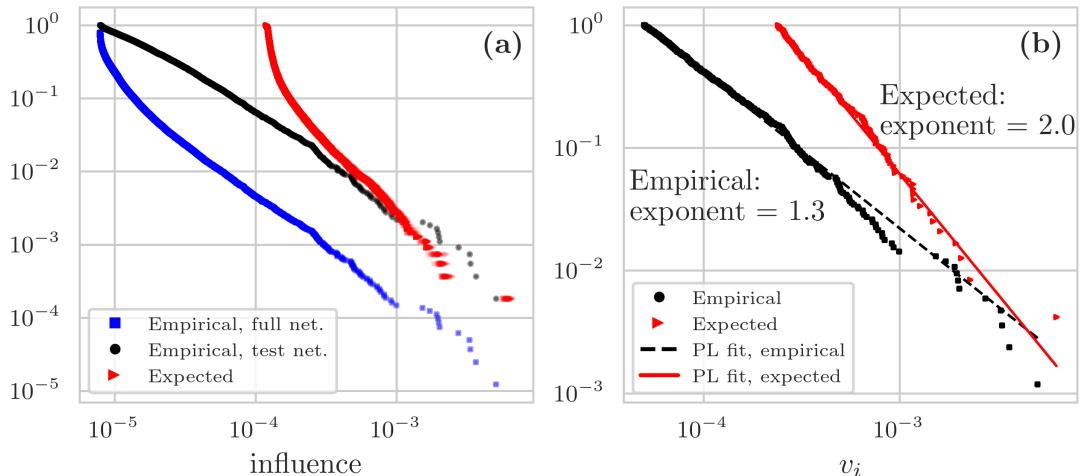


Figure C.5: (a) CCDF of the influence vector of the empirical full network (all firms in the Ecuadorian economy; blue squares), the empirical (trimmed) test network (black dots) and the reconstructed networks (red triangles). We show all 50 randomised test networks. (b) Power-law fit to the empirical (black dots) and expected (red triangles) influence vector distribution. The black dashed line shows the power-law fit to the empirical distribution, while the red solid line the power-law fit of the reconstructed distribution; we show results for one of the reconstructions.

FactSet. Figure C.6 shows the distribution of the expected influence vector (blue squares) and its power-law fit (solid blue line) for FactSet. The power-law exponent equals 1.9.

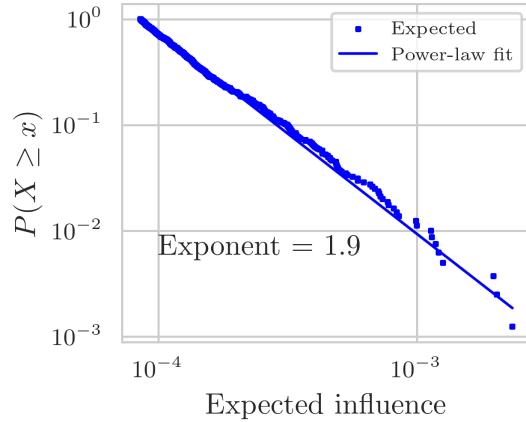


Figure C.6: CCDF of the expected influence vector (blue squares) and its power-law fit (solid blue line) for FactSet.

C.3.2 Prediction errors

Figure C.7a shows the histogram of the relative prediction errors for the output multipliers, while Figure C.8a for the influence vector, both are for Ecuador and for one of the 50 randomised networks. For the output multipliers, the mean relative prediction error is -0.7% and the median is -0.4%. For the influence vector, the mean relative prediction error is 735.1% and the median is 679.4%.

Figure C.7b shows the empirical output multipliers against the relative prediction errors for Ecuador, while Figure C.8b and C.8c for the influence vector. Figure C.8b shows observations for which the error term is non-positive, while Figure C.8c for observations for which the error term is positive.

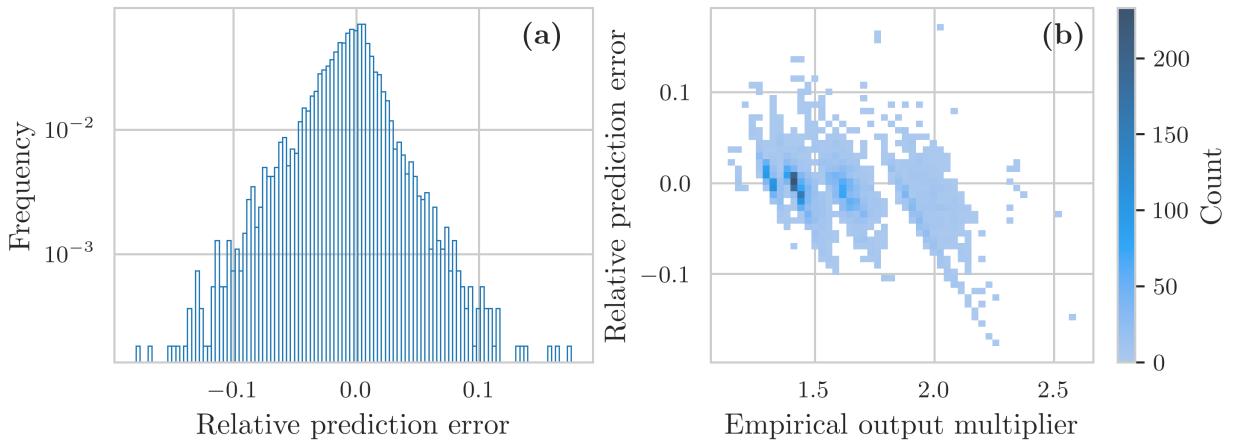


Figure C.7: (a) Histogram of the relative prediction errors of the output multipliers for Ecuador. (b) 2D histogram of the empirical output multipliers (x -axis) and the prediction errors (y -axis). We use 50 log-spaced bins for both axes and count the member of points falling into each square.

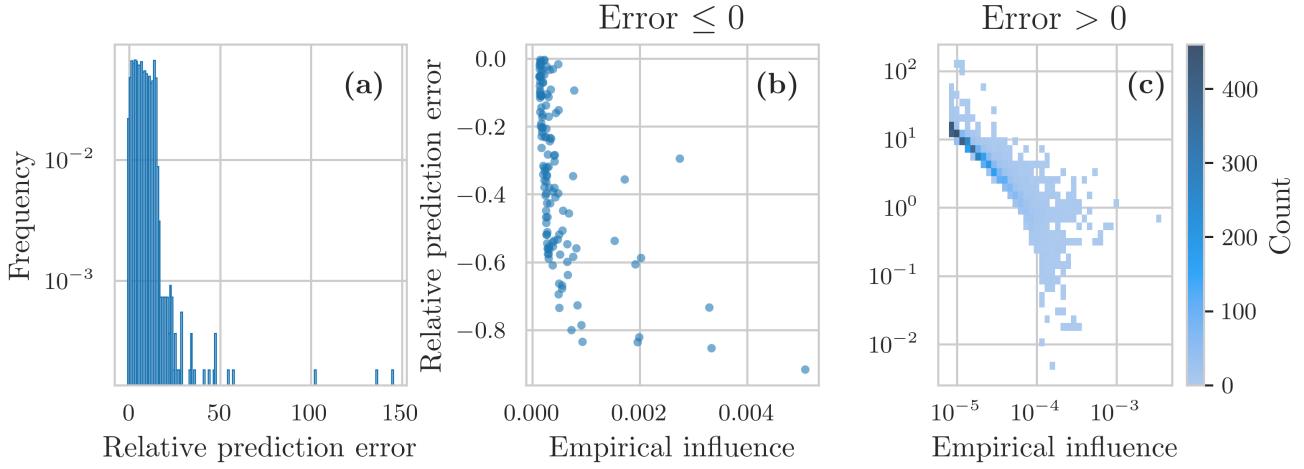


Figure C.8: (a) Histogram of the relative prediction errors of the influence vector for Ecuador. (b), (c) 2D histogram of the empirical influence vector (x -axis) and the relative prediction errors (y -axis) for errors that are non-positive and positive, respectively. We use 50 log-spaced bins for both axes and count the number of points falling into each square.

C.3.3 Empirical and expected multipliers by sector

Output multipliers. Figure C.9 shows the empirical output multipliers on the x -axis against the expected output multipliers on the y -axis for Ecuador. Points are coloured based on the sector firms are in. Data points tend to organise in clusters, which form because the calculation of the technical coefficients requires knowing the value-added of each firm. Since we did not have access to firms' value-added, we used sector-level I-O tables to infer firms' value-added. The value-added of firm i is $y_i = \nu_s \cdot s_i^{\text{in}}$, where $\nu_s = y_s / s_s^{\text{in}}$ is the ratio of value-added and total intermediate expenses of sector s . We show ν_s in Figure C.9b, also colour-coded by sector.

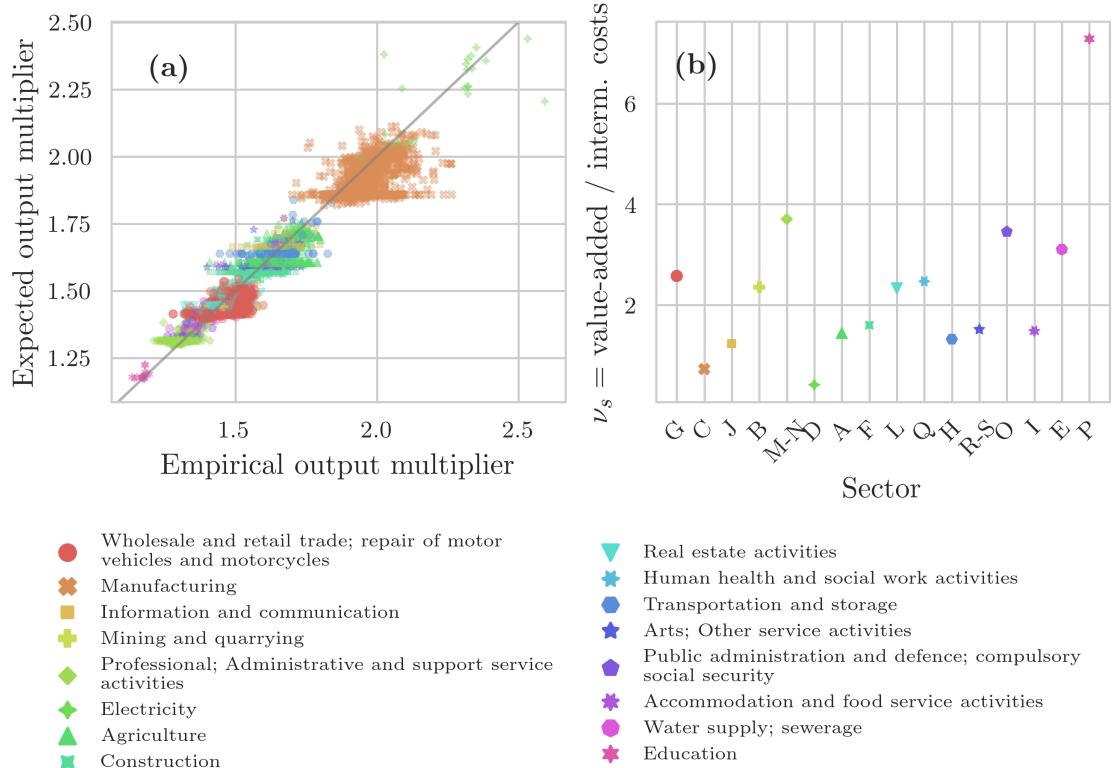


Figure C.9: (a) Empirical (x -axis) and expected (y -axis) output multipliers for Ecuador. Perfect prediction is achieved when all points lie on the 45-degree line (solid grey line). (b) $\nu_s = y_s / s_s^{\text{in}}$, i.e., the ratio of value-added and total intermediate expenses by sector in Ecuador's I-O table. Different colours and shapes correspond to the sectors in the I-O table in which firms are.

Influence vector. Figure C.10 shows the empirical influence vector on the x -axis against the expected influence vector on the y -axis for Ecuador. Points are coloured based on the sectors in the I-O table in which firms are.

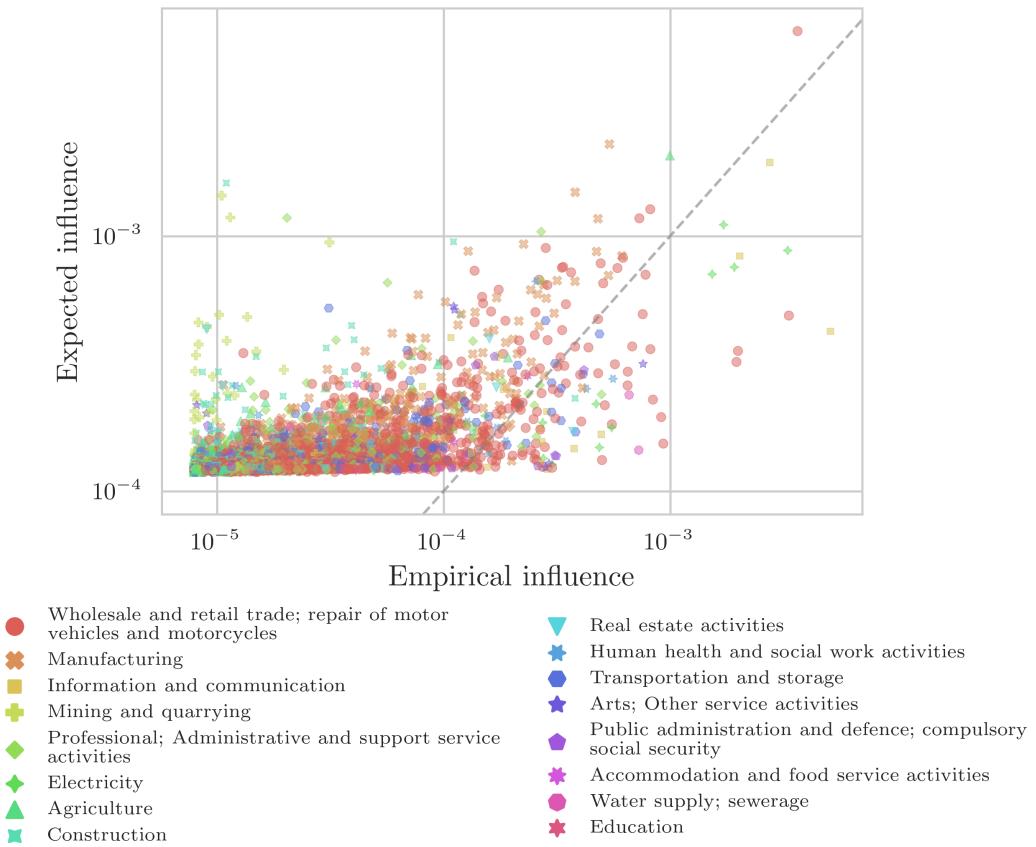


Figure C.10: Empirical (x -axis) and expected (y -axis) influence vector for Ecuador. Perfect prediction is achieved when all points lie on the 45-degree line (dashed grey line). The colours and shapes correspond to different sectors.

C.3.4 The effect of the proxy node on the centrality measures

Figure C.11a and Figure C.11d show the empirical multipliers against the reconstructed multipliers including the proxy node, while Figure C.11b and Figure C.11e excluding the proxy node from the calculations. Figure C.11c and Figure C.11f compare the expected multipliers that include to proxy node to those that exclude the proxy node from the calculations.

The influence vector is not that much affected by the inclusion, or not, of the proxy node (Figure C.11d and Figure C.11e). The inclusion of the proxy node increases the influence of all the nodes compared to when it is not included (Figure C.11f). While the error metrics are only slightly lower when the proxy node is excluded, the cosine similarity is slightly higher (0.56 vs 0.55) and the power-law exponent slightly smaller (2.0 vs 2.1). The mean and median are off for both and we can only recover the variance if we include the proxy node, otherwise it is smaller. The proxy node is always the one with the higher influence.

On the contrary, the output multipliers are highly affected by the inclusion, or not, of the proxy node (Figure C.11a and Figure C.11b). When the proxy node is not included the output multipliers are underestimated (Figure C.11b). The proxy node is never the one with the highest centrality. When the proxy node is included, we can recover the mean, median and standard deviation, while when it is excluded, these are all underestimated (especially the mean and median).

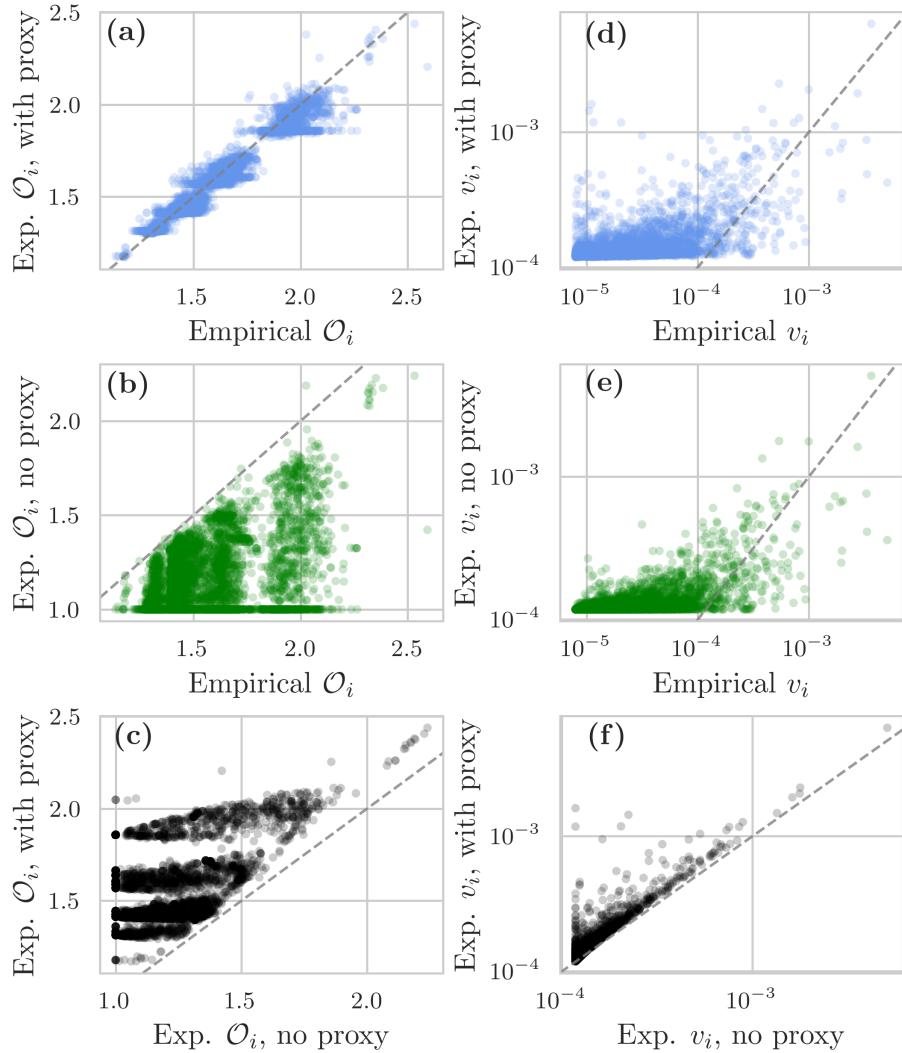


Figure C.11: (a), (d) Empirical centrality measure (x -axis) and the expected calculated including the proxy node (y -axis) for the output multipliers (\mathcal{O}_i) and the influence vector (v_i), respectively. (b), (e) Empirical centrality measure (x -axis) and the expected calculated excluding the proxy node (y -axis) for the output multipliers (\mathcal{O}_i) and the influence vector (v_i), respectively. (c), (f) Expected centrality measure calculated excluding the proxy node (x -axis) and the expected centrality measure calculated including the proxy node (y -axis) for the output multipliers (\mathcal{O}_i) and the influence vector (v_i), respectively. The dashed grey line marks the identity line.

Before delving into the details of why this is happening, it is useful to define again these two centrality measures. The output multiplier is defined as

$$\mathcal{O} \equiv (\mathbf{I} - \mathbf{T}^\top)^{-1} \mathbf{1},$$

or in scalar form

$$\mathcal{O}_j = \sum_i \frac{W_{ij}}{q_j} \mathcal{O}_i + 1.$$

The output multipliers measure the centrality of a node by assigning a centrality score for each incoming edge – of course, these edges are weighted and normalised by the total cost of the node for which we are measuring the multiplier. The influence vector is defined as

$$\mathbf{v} \equiv \frac{\alpha}{N} [\mathbf{I} - (1 - \alpha) \mathbf{\Omega}]^{-1} \mathbf{1},$$

or in scalar form

$$v_i = \sum_j (1 - \alpha) \frac{W_{ij}}{s_j^{in}} v_j + \frac{\alpha}{N}.$$

The influence vector, instead, assigns a centrality score for each node the focal node points to. So nodes that point to more nodes are more central – again, edges are weighted and normalised by the in-strength of the nodes to which the focal node points to. In both centrality measures, a node can be very influential because it has many incoming or outgoing links with nodes of mild influence, for the output multiplier and the influence vector, respectively, or because the node has few incoming or outgoing links with highly influential nodes, for the output multiplier and the influence vector.

Since the only thing that changes in the calculated multipliers is the inclusion or not of the proxy node, let us focus on the proxy node's role in the multipliers:

$$\mathcal{O}_j \sim 1 + \frac{W_{pj}}{q_j} \mathcal{O}_p + \dots, \text{ and}$$

$$v_i \sim \frac{\alpha}{N} + (1 - \alpha) \frac{W_{ip}}{s_p^{in}} v_p + \dots,$$

where p indexes the proxy node. Each link of the proxy node matters more in the output multipliers since, empirically, $W_{pj}/q_j > (1 - \alpha)W_{ip}/s_p^{in}$, for all but the link the proxy node has with itself and with one another node.

C.3.5 Reconstructed and benchmark influence vector

Figure C.12 shows the “uniform” influence, which we use to benchmark aggregate volatility, against the reconstructed influence vector for the 50 randomised networks. It can be seen that the reconstructed influences tend to be bigger than those of the benchmark. However, the error metrics of the two influence vectors have very similar values, except for the cosine similarity, which is lower for the uniform influence vector (Table C.1).

Type	RMSE	MAE	MedAE	Cosine similarity
Influence vector	2×10^{-4} (9×10^{-7})	1×10^{-4} (1×10^{-7})	1×10^{-4} (5×10^{-8})	0.560 (3×10^{-3})
Uniform influence vector	2×10^{-4} (4×10^{-7})	1×10^{-4} (1×10^{-7})	1×10^{-4} (4×10^{-8})	0.478 (4×10^{-3})

Table C.1: Comparison metrics for the reconstructed and uniform influence vector for Ecuador. RMSE denotes the root mean squared error, MAE the mean absolute error and MedAE the median absolute error. For each metric, we show its mean value across the 50 randomised reconstructions. Below the mean value, the standard deviation in parenthesis. We excluded the proxy node from the calculations.

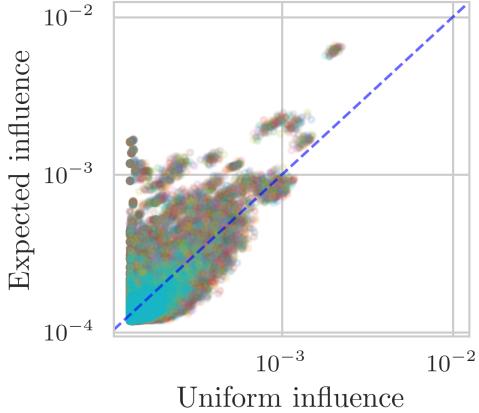


Figure C.12: “Uniform” influence on the x -axis and expected influence on the y -axis for the 50 randomised networks. The “uniform” influence is calculated by assuming that each firm buys inputs from its suppliers in the same proportion; this is the influence vector we use in our benchmark for aggregate fluctuations. The dashed blue line marks the identity line.

C.4 Different numbers of unknown links

C.4.1 In- and out-degrees

Figure C.13 shows the number of suppliers (in-degree) against the number of customers (out-degree) for different networks. The first panel on the top left-most corner shows the in- and out-degrees for the full network, the second panel for the test network and the consecutive panels for test networks with different numbers of unknown links. The last panel, at the bottom right-most corner, shows results for the trimmed test network that we discuss throughout the paper, which has the same average degree as FactSet. It can be noticed that in the full network firms tend to have more customers than suppliers, something also observed by [Bacilieri et al. \(2023\)](#). In creating our test network, firms lose more customers than suppliers. The maximum out-degree is 27,030 in the full network and decreases almost 7 folds in the test network (96% of missing links), reaching a value of 4,081, while the maximum in-degree is 2,543 in the full network and lowers to 781 in the test network (96% of missing links), which is a 3-fold decrease.

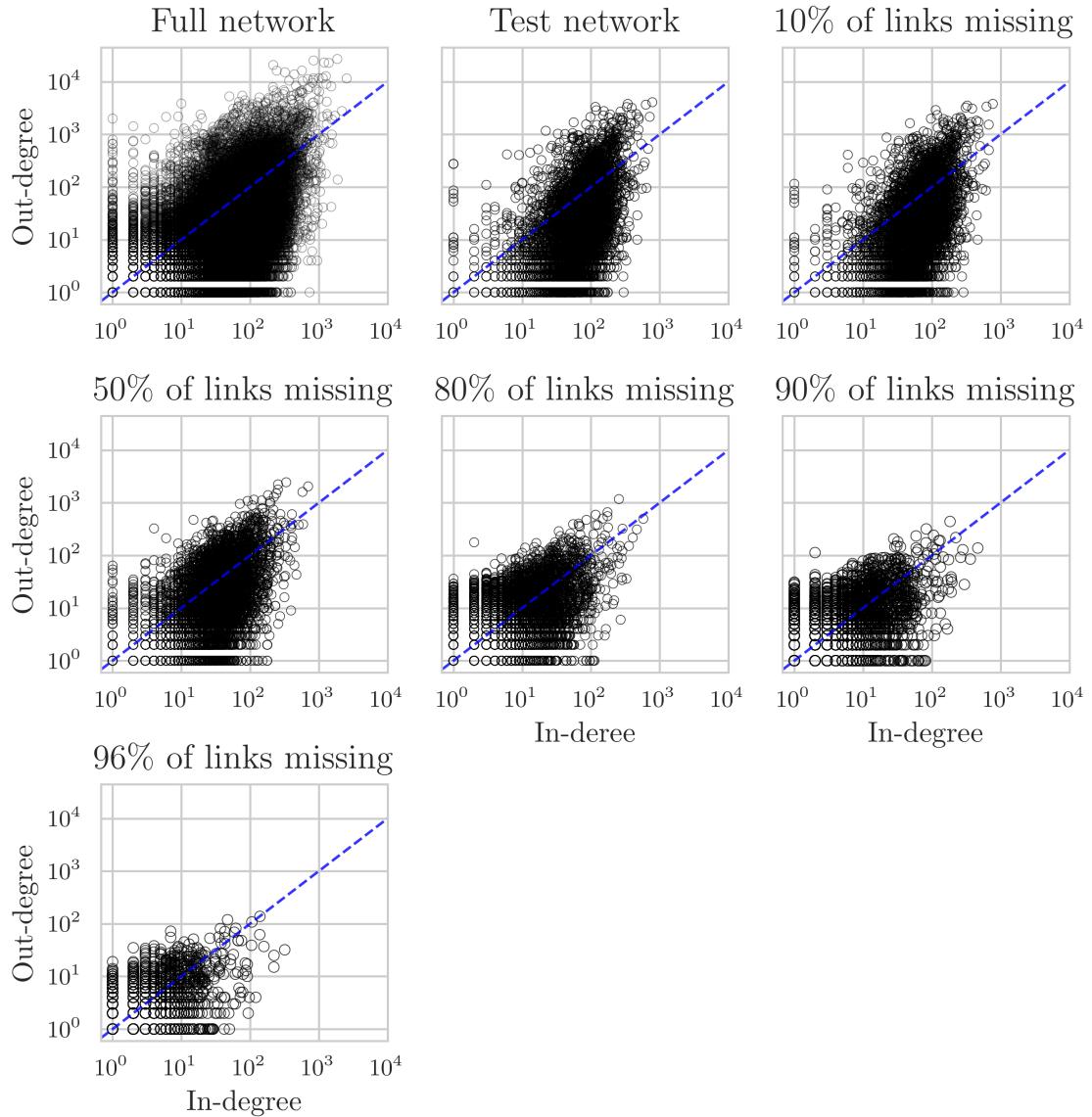


Figure C.13: Number of suppliers (in-degree) on x -axis and number of customers (out-degree) on y -axis for the full network, the test network and for test networks with different numbers of unknown links. The blue dashed line marks the identity line.

C.4.2 Comparison of empirical and expected quantities

Microscale quantities. Figure C.14 shows the empirical values against the expected values for the allocation coefficients (top panels), the technical coefficients (2^{nd} row), and the weights (bottom panels) for different numbers of unknown links. The allocation coefficients display a transition when the number of unknown links rises from 90% to 96%, something that the other two quantities do not experience. It also highlights the sampling process used for the deletion of links: as the number of unknown links increases, higher and higher weights are eliminated. However, the shape of the joint density does not vary, it is only re-scaled.

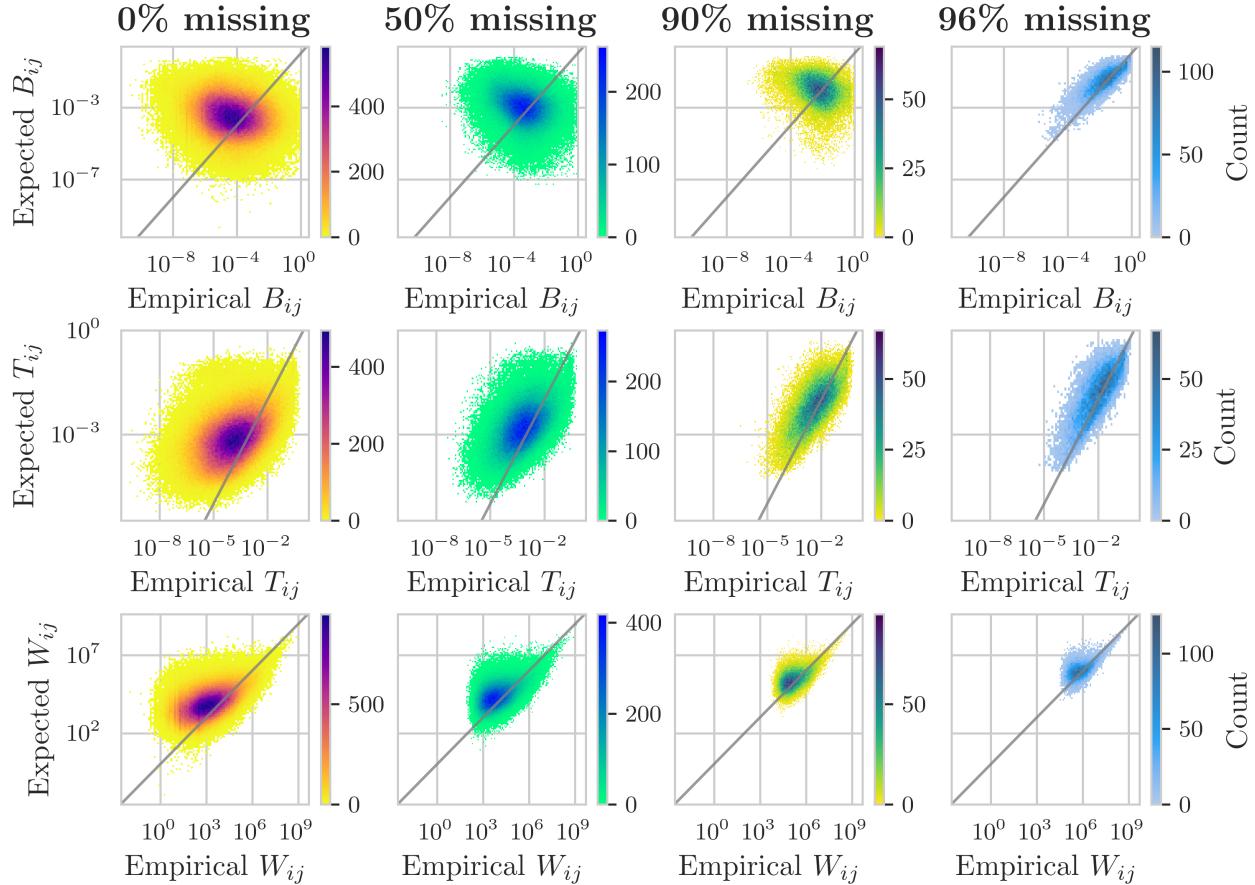


Figure C.14: 2D histograms for the empirical values on the x -axis and the expected values on the y -axis for different numbers of unknown links (0%, 50%, 90% and 96%) for the allocation coefficients (B_{ij}), the technical coefficients (T_{ij}) and the weights (W_{ij}). We bin each axis into 100 log-spaced bins and count the number of data points that fall in each square.

Higher-order quantities. Figure C.15 shows the empirical values against the expected values for the output multipliers (left column) and the influence vector (right column) for different numbers of unknown links (0%, 20%, 50%, 90% and 96%). Figure C.16 shows the empirical CCDF (black dots) for the output multipliers (left) and the influence vector (right). It also shows the CCDF of the reconstructed multipliers for different numbers of unknown links and for the 50 randomised networks.

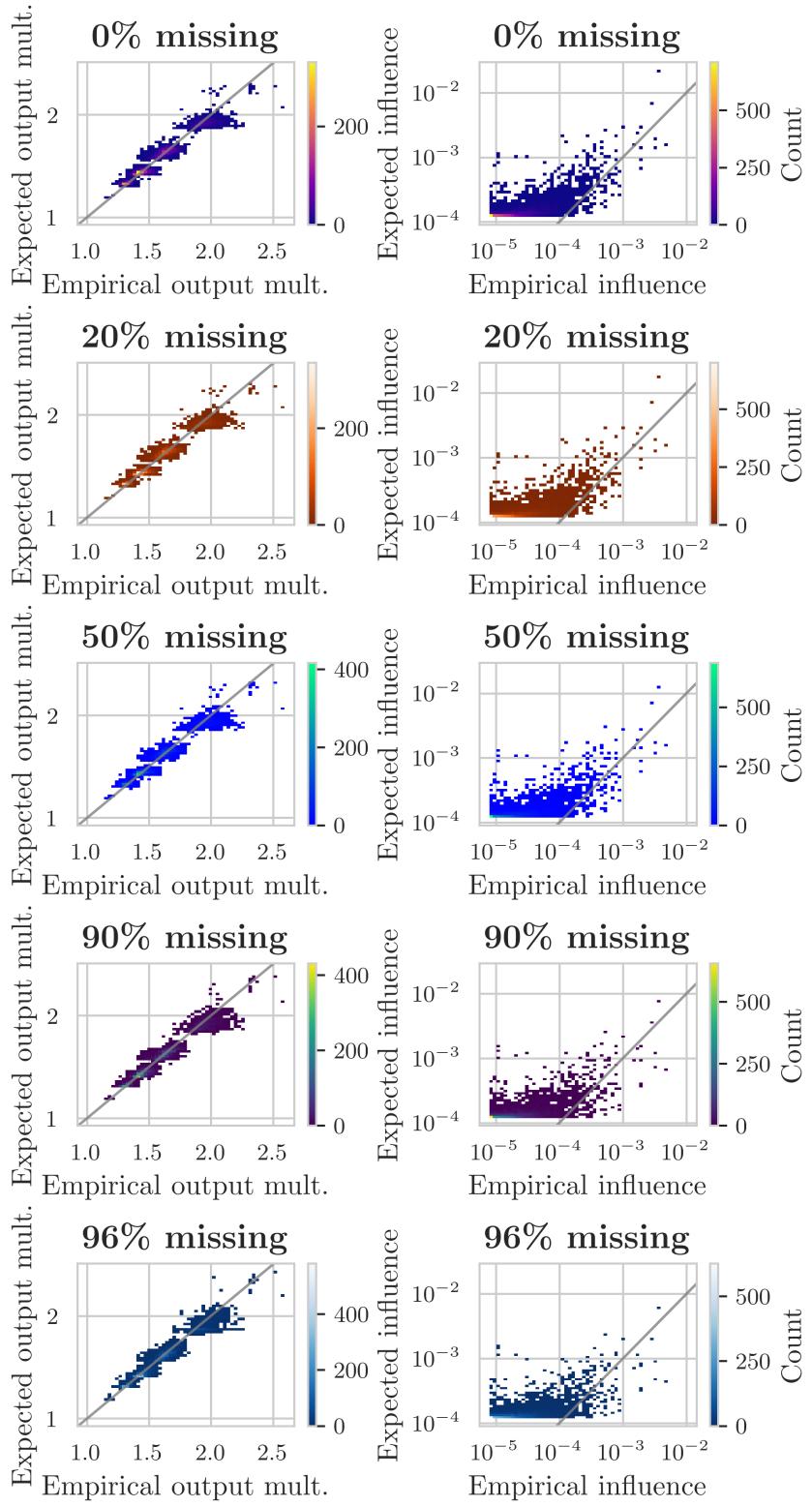


Figure C.15: 2D histograms for the empirical values on the x -axis and the expected values on the y -axis for different numbers of unknown links (0%, 20%, 50%, 90%, and 96%) for the output multipliers (left column) and the influence vector (right column). We bin each axis into 50 log-spaced bins and count the number of data points that fall in each square.

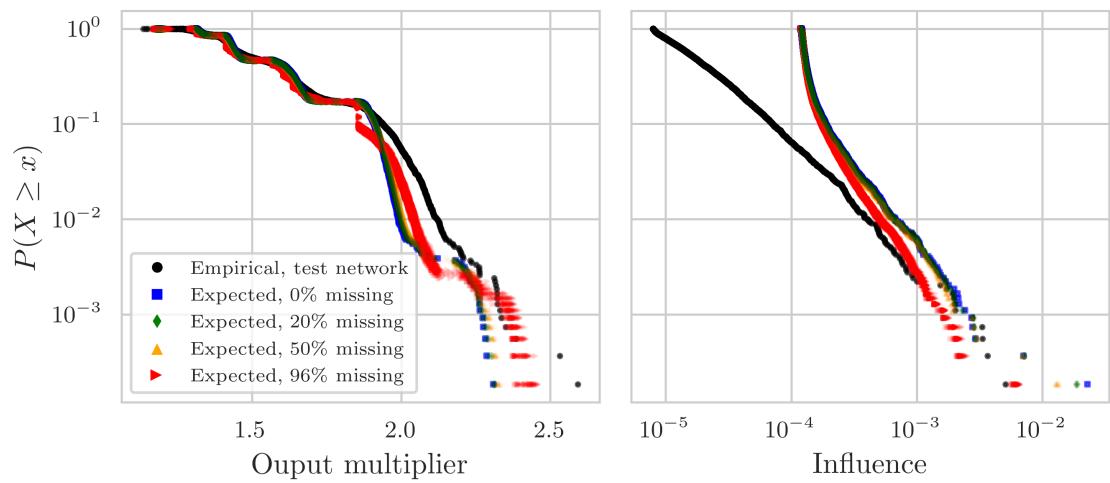


Figure C.16: CCDF of the output multipliers (left) and influence vector (right) for different numbers of unknown links (0% blue squares, 20% green diamonds, 50% yellow triangles and 96% red triangles) and the empirical (black dots).