Reconstructing supply networks

**Luca Mungo, Alexandra Brintrup,
Diego Garlaschelli and François Lafond**

3 Oct 2023

Reconstructing supply networks[a]

Luca Mungo,[1, 2] Alexandra Brintrup,[3] Diego Garlaschelli,[4, 5] and François Lafond[1, 6]

[1] *Institute for New Economic Thinking at the Oxford Martin School, University of Oxford, UK*
[2] *Mathematical Institute, University of Oxford, UK*
[3] *Institute for Manufacturing, Department of Engineering, University of Cambridge, UK*
[4] *IMT School of Advanced Studies, Lucca, Italy*
[5] *Lorentz Institute for Theoretical Physics, University of Leiden, The Netherlands*
[6] *Smith School of Enterprise and the Environment, University of Oxford, UK*

ABSTRACT

Network reconstruction is a well-developed sub-field of network science, but it has only recently been applied to production networks, where nodes are firms and edges represent customer-supplier relationships. We review the literature that has flourished to infer the topology of these networks by partial, aggregate, or indirect observation of the data. We discuss why this is an important endeavour, what needs to be reconstructed, what makes it different from other network reconstruction problems, and how different researchers have approached the problem. We conclude with a research agenda.
Keywords: link prediction; supply networks

I. INTRODUCTION

Following the 2008 financial crisis, financial networks have been extensively studied by the complex systems community. For example, studying liabilities in banking networks has been key to developing the notion of systemic risk [1, 2], and explaining how certain types of interconnections may amplify the impact of isolated shocks. A key component of this research was the development of methods to reconstruct the network of interdependencies between financial institutions, which are not easily observable.

More recently, systemic failures on of the *supply network* have captured the attention of the complex systems community, as researchers observed the impact of several significant failures, such as disruptions following the Great East Japan Earthquake in 2011, protective equipment shortages during the COVID-19 pandemic, supply shocks after the Suez Canal obstruction by the Ever Given, and the energy supply chain reorganization due to the war in Ukraine.

Production networks, also known as "supply chains" or "supply networks", consist of millions of firms producing and exchanging goods and services. From a mathematical perspective, they can be represented as weighted, directed graphs, where nodes symbolize firms (or establishments), and links may denote a supply-buy relationship with weights denoting transaction volume, such as the monetary value of the goods or services supplied over a given period.

Supply networks share many properties with other economic networks, but also exhibit unique features. Some of their empirical properties include [3]: small-world properties (short average path lengths and high clustering), heavy-tailed degree distributions, heavy-tailed (link and/or node) weight distributions, strong correlations between node strength and degree, and similarly between in- and out-degrees. It is also relatively well documented that, like biological and technological networks but unlike social networks derived from co-affiliation [4], supply networks feature negative degree assortativity.

However, supply networks are in many ways very different from other natural and economic networks. Their properties are deeply influenced by their function. First, the likelihood of a link between any two

firms is driven by what the two firms are producing: for instance, steel manufacturers buy more iron than sugar. In general, each link in a supply network may represent one or more types of products; the diversity of products involved may depend on how the data are collected and may crucially affect network properties such as the reciprocity of connections. Product quality also plays a role, with "high quality" firms usually connecting with other "high quality" firms [5]. Second, supply networks are strongly embedded in geographic space, so that the likelihood of connections and their intensity decreases with distance [6]. Third, in contrast to financial networks, supply networks are less constrained by strict external regulations, and emerge as the result of a decentralized multi-criteria optimization process whereby millions of organizations simultaneously attempt to outsource in a way that minimizes their costs while maintaining acceptable levels of resilience to disruptions, for instance by multi-sourcing.

These characteristics make production networks incredibly complex: in modern economies, a sophisticated product such as an aircraft might involve contracting thousands of firms and sourcing millions of parts that cross national borders multiple times. Organizations in the network choose their dyadic relations and make local decisions, but hardly have visibility over their wider network. No single entity controls, designs and keeps track of the large-scale emergent network. Visibility over the network is, however, increasingly important for several reasons: monitoring of environmental pledges to ensure firms quantify their greenhouse gas emissions, including those from their suppliers and customers; food and pharmaceutical traceability; analysing and improving supply chain resilience; and supply chain due diligence to ensure that actors that violate human rights or engage in environmentally damaging actions are not present in the chain.

In the past decade, researchers in economics and complex systems have worked extensively to better understand supply chains. A key barrier to these studies has been a lack of data, as supply chains compete with one another [7], making information on them highly commercially sensitive. As a result, most studies to date have used firm-centred (e.g. starting with [8]) or sector-specific (e.g. global automotive [9] and aerospace [10], computer and electronics [11]) supply chains. While firm-centric and industry-specific studies have been important to gather insights into how network features shape the operation of supply chains, it remains hard to generalize these findings, due to the sector-specific and incomplete nature of these datasets.

Due to the above challenges, several recent studies have suggested the development of methods to reconstruct or predict the existence of hidden links in supply chain networks, offering a variety of approaches. These range from the use of natural language processing to extract and infer data from the World Wide Web to probabilistic maximum-entropy methods, each with varying success rates.

In this paper, we synthesize recent research on reconstructing supply networks. We start by describing the key problems: what data is available, what data is missing, and how to evaluate reconstruction performance (Section II). We then summarise recent approaches to inferring the network topology (Section III), and to infer the values of transactions when the topology is known (Section IV). We conclude with a discussion (Section V) and two research agendas (Section VI) focusing on macroeconomic and supply chain management applications.

II. THE SUPPLY NETWORK RECONSTRUCTION PROBLEM

Production networks can be modelled at different levels of detail, both for nodes and edges. Naturally, the properties of the network depend on the level of aggregation.

At the most granular level, nodes would represent individual production plants where goods undergo processing and transformation. A more aggregate model would equate nodes with the companies operating these plants. One could further aggregate by either consolidating firms under a common parent company or grouping them by industry sector[1].

Firms exchange various goods and services. In a very detailed approach, each product type could be identified with a specific type of edge, rendering the production network as an edge-labelled multigraph.

---

[1] One could think that the industry level is more aggregated than the firm. While this is mostly true, it is sometimes important to recognize that large firms span many industries. Indeed, industry-level input-output networks produced by National Accounts arise from Supply and Use Tables, which attempt to reallocate the output and inputs of multi-product firms into their appropriate categories.

A simpler model would connect two nodes if they are involved in any type of trade, irrespective of the products' nature. Link weights can also have different definitions, measuring either the flow of goods (in terms, e.g., of the number of items traded) or the monetary value of such flow.

In the context of this paper, we define a *supply network G* as a graph where nodes represent firms while directed, weighted links represent the value of the flow of goods and services in a buyer-customer relation. This definition proves practical when reconstructing real-world supply networks from empirical data, which frequently adopts this format.

A. What data is available?

Almost all countries officially release Input-Output (I-O) tables, which provide the flow of money between industries, typically at the level of 50-500 industries. While we focus on firms here, this data is sometimes useful in the methods below. Besides, I-O tables provide a meso-scale ground truth that could be a good target for reconstruction methods.

Bacilieri *et al.* provides a taxonomy of existing datasets documenting different representations of supply networks. These are mainly: commercial datasets, confidential datasets held by governments, payment data, and industry-specific datasets. We briefly describe these types of data below.

Purchasing data from data providers, such as FactSet, Capital IQ, or Bloomberg is relatively straightforward, but commercial datasets can be very expensive, and cover only a fraction of firms, a very small fraction of links, and do not systematically include the value of the transactions. As commercial data providers typically assemble their data from publicly available information, researchers may also decide to collect this information themselves. An example is the extraction of data from the World Wide Web, after which machine learning algorithms are trained to predict supply-buy relationships [12]. Such an approach enables researchers to successfully gather rudimentary maps of supply chains, although it is limited to publicly available data, hence necessitating reconstruction efforts to identify missing relationships.

The option of using government-held data necessitates datasets to be shared by national authorities, which may not always be feasible. However, where data has been collected by a national authority it tends to be of very high quality. For example, VAT reporting may contain the value of transactions and timestamped data between virtually all firms within a country. Bacilieri *et al.* show that VAT datasets with no reporting thresholds exhibit strikingly similar properties, while incomplete datasets (either because of a reporting threshold or because they are assembled from publicly available information) usually have fewer links, so that many key statistics are likely to be highly biased.

A third option is payment data, which is usually (but not always) limited to individual banks collecting payment flows between their client firms (see, e.g., [13]). Although it is not guaranteed that every transaction corresponds to a business link within a supply network, it can be viewed as a plausible indicator. These datasets are extremely detailed for any subset of firms affiliated with the same bank. However, they do not cover firms served by different banks or accounts held by their clients in different institutions.

Finally, datasets focusing on specific industry verticals are also sometimes gathered by private companies (e.g., MarkLines' automotive dataset used in Brintrup *et al.* [14])) and public regulatory bodies (e.g., the U.S. Drug Enforcement Administration's dataset of controlled substances flow). However, they are usually limited to specific geographies and production sectors.

There are no large-scale publicly available datasets on firm-level production networks, making it impossible at the moment to portray the global supply network. Summing up the number of nodes in the datasets reported in Bacilieri *et al.* [3] gives less than 3m, so less than 1% of the 300m nodes reported earlier. Merging all the available datasets would give only an even smaller portion of the links and weights. This limitation forces researchers to use alternative options to proxy supply networks from smaller-scale, more specific datasets. These methodologies, developed to reconstruct or infer missing information about supply networks, are the main focus of this paper.

B.   A taxonomy of supply network reconstruction approaches

Clearly, what we actually mean by 'reconstructing' a supply network necessarily depends on the data already available to the researchers and on the ultimate use of the (inferred) network, i.e. the goal of the analysis. We discuss these points in what follows and classify the studies we review along four primary axes. We do not see these classifications as having rigid boundaries, but rather as providing continuous dimensions along which models can be placed.

  a.   Predicting network topology and/or weights on transactions.   Consider a matrix Ω where Ω_{ij} shows the amount paid by j to i. We distinguish between methods that focus only on finding the network's *topology*, i.e., the presence or absence of a commercial connection between two firms encoded in the (binary) adjacency matrix $A_{ij} = 1 \leftrightarrow \Omega_{ij} > 0$, and those that assume that the adjacency matrix is known and try to infer the monetary value of the existing connections, i.e. the *link weights* $\Omega_{ij}|A_{ij} = 1$ (see also point c below). Note that some methods try to simultaneously reconstruct both the topology and the weights of the network. Most of the methods we review focus on network topology.

  b.   Predicting individual links or the full network.   Some methods focus on identifying the presence of specific links independently, while others try to reconstruct the entire network at once. The difference is subtle, yet important. Typically, links in real-world production networks are not independent. This happens, for instance, if firms tend to avoid "multi-sourcing": this happens if, when they are connected to supplier j for a key input, they are less likely to be connected to other suppliers for that input. In reconstruction methods, links are sometimes assumed to be mutually dependent, and sometimes assumed to be independent. Generally (but not necessarily), the assumption made is related to the ultimate goal of the reconstruction method. The task of trying to identify the presence of specific links is usually known as *link prediction* [15], while that of inferring the full network architecture is referred to (at least in this paper) as *network inference*. In general, network inference computes the full distribution $P(G)$ over the set $\mathcal{G} = \{G\}$ of all possible networks. Link prediction, instead, computes the marginal probability p_{ij} of an edge between nodes i and j^2. Again, there is no hard boundary between the two methods, which are occasionally equivalent: if one considers link independence as (the result of) a modelling assumption, computing the values $\{p_{ij}\}$ for all pairs of nodes and reconstructing the whole network become two equivalent operations, as the probability $P(G)$ factorizes as

$$P(G) = \prod_{(i,j)\in E(G)} p_{ij} \prod_{(i,j)\notin E(G)} (1 - p_{ij}), \tag{1}$$

where $E(G)$ denotes the set of edges realized in graph G. In this case, link prediction and network inference coincide. On the other hand, whenever the full probability $P(G)$ in a network inference method is available (and irrespective of whether edges are assumed to be independent or not), it is always possible to compute the *marginal* connection probability p_{ij} as $p_{ij} = P(A_{ij} = 1) = \sum_{G\in\mathcal{G}} P(G) A_{ij}$ and use it in a link prediction exercise.

  It is fair to say that the factorization in Eq. (1) is, at most, only approximately true in reality. However, some methods with independent edges can still capture meso- and macro-scale features of the network (see, e.g., [13]) and, by framing the reconstruction problem as a binary classification task, link prediction facilitates easy comparison of methods through standard performance metrics.

  c.   Using topological information or not.   Of course, all reconstruction methods need, at the end of the procedure, the whole empirical network as the 'ground truth' to *test* their predictions. However, while some methods need the full adjacency matrix also in their training, other methods can learn from node-level or pair-level features only. This is important because the methods that do not rely on the adjacency matrix for training can be used in contexts where the detailed network is not observed, as long as certain node-level (and possibly pair-level) features are available.

----

^{2} More generally, link prediction methods produce a *score* s_{ij}, such that $s_{ij} > s_{kl} \implies p_{ij} > p_{kl}$. However, such scores are not necessarily smaller than one, and the ratio between two scores is not necessarily equal to the ratio between links probabilities.

d. Probabilistic or deterministic. Some models produce *deterministic* outputs, usually finding a network configuration by minimizing or maximizing a given loss function. Consequently, their output is a single network realisation that is on one hand optimal according to some score, but on the other hand very unlikely to represent the true network. Other methods provide *probabilities* over possible network realisations. The goal of these methods can then be viewed as finding a 'good' probability distribution, peaked 'around' or 'close' to the true one. Equipped with this probability distribution, researchers can find the typical and most likely realisations of the network and compute, for instance, expected values and confidence intervals for properties of the network.

C. Evaluating the reconstructed networks

In their review paper on network reconstruction, Squartini *et al.* provide a useful taxonomy of performance metrics: *statistical*, *topological*, and *dynamical* indicators.

Statistical indicators evaluate the quality of the reconstructed network on a link-by-link (or weight-by-weight) basis. Different statistical indicators apply to deterministic and probabilistic outcomes.

In the realm of deterministic outcomes, perhaps the most commonly employed indicator is *accuracy* (or *precision*), the proportion of correct predictions. In supply networks, however, there is a strong class imbalance: the number of pairs not linked is much higher than the number of pairs linked. Thus, it is generally easy to make "correct" predictions since predicting that a link does not exist is very likely to be correct. For this reason, a commonly used metric is the *F1-score*, defined as the harmonic mean of precision and recall (how many existing links are predicted as existing), which offers a more balanced performance metric in unbalanced datasets.

For probabilistic reconstructions, the evaluation is often based on the *area under the receiver operating characteristic curve* (AUROC) and the *area under the precision-recall curve*. AUROC, derived from the Receiver Operating Characteristic (ROC) curve, essentially quantifies the ablity of the models to discern between classes at varying threshold levels. The ROC curve plots the true positive rate (recall) against the false positive rate for different decision thresholds (i.e., by considering "true" all the predictions with probability larger than a certain threshold τ, for different values of τ), giving insights into the trade-off between sensitivity (true positive rate) and specificity (true negative rate). The AUROC, being the area under this curve, ranges from 0.5 to 1, with 1 implying an ideal classifier and 0.5 corresponding to no better than random guessing.

Because statistical indicators focus on individual links, they may not adequately evaluate if the reconstructed network replicates complex network structures. *Topological* indicators measure how well the network's macro-level and meso-level features are reproduced. Topological indicators gauge how effectively the reconstruction captures the network 'coarse-grained' features. For instance, Ialongo *et al.*, validate their reconstruction methodology by assessing how accurately it replicates the network degree distribution.

Topological indicators can tell us whether the reconstructed and true networks are "similar". However, ultimately the key question is whether a reconstructed network is good enough to give good answers to substantive economic questions. *Dynamical* (or more generally model-based) indicators assess the similarity in the process' evolution on the real and reconstructed networks. As an example, Diem *et al.* introduced the *Economic Systemic Risk Index* (ESRI) to quantify each firm's importance within an economy. The metric measures the percentage drop in the economy's overall production caused by the removal of a firm from the network. Its computation requires running a dynamical process, wherein the sudden disappearance of a firm first impacts its suppliers and customers and, iteratively, spreads to firms that are further away in the network, until the system reaches an equilibrium. Conceivably, accurately estimating firm-level ESRI may only necessitate identifying a subset of key links, so a good prediction of the other links is not necessarily important for the final economic result.

Armed with these evaluation indicators, we now examine in detail the models employed for reconstructing production networks, starting from methods focusing only on the network topology, and then discussing methods for reconstructing network weights.
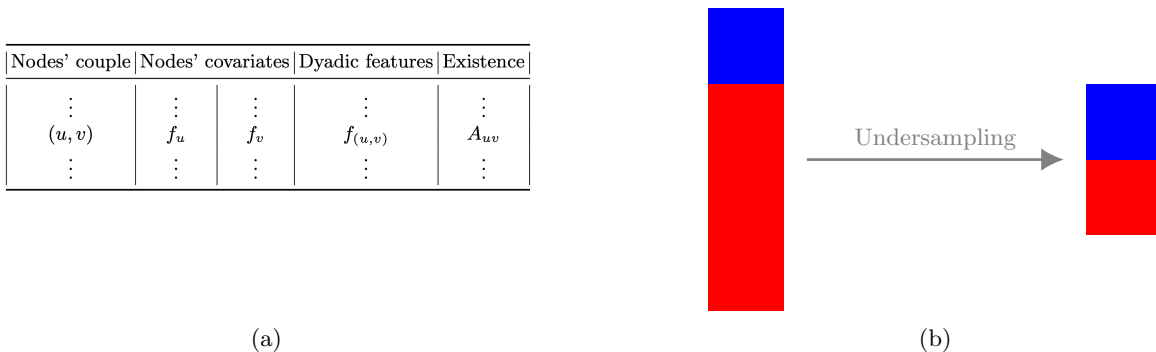
Nodes' couple	Nodes' covariates		Dyadic features	Existence
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
(u,v)	f_u	f_v	$f_{(u,v)}$	A_{uv}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

(a)

Undersampling

(b)

FIG. 1: (a) Datasets for link prediction are usually built by filling rows with two nodes features (f_u, f_v, $f_{u,v}$) and by indicating if there is a link between the two nodes ($A_{u,v}$). (b) These datasets are usually undersampled: in the original dataset, a small minority of the rows will be s.t. $A_{u,v} = 1$ (blue), while most of the rows will be s.t. $A_{u,v} = 0$ (red); undersampling discards a portion of them to generate a more balanced dataset.

III.   RECONSTRUCTING THE NETWORK TOPOLOGY

We start by reviewing studies that reconstruct the network using link prediction, and then those that do so using network inference methods. Table I provides an overall summary of the methods and their differences.

A.   Link prediction

1.   Setting up the problem

An early stream of research employs machine learning for link prediction in production networks. The key idea is to construct a dataset in the form of Fig. 1A, where for each pair (i, j), we collect some features $f_{(i,j)}$ that can be features of each node (e.g., the product it makes, its total sales, etc.) or of the pair (e.g. geographical distance, whether they have a common supplier or client, etc.), and the response A_{ij}, which is equal to 0 or 1.

With such a dataset, one can then train a machine-learning classifier on a set of examples $\left\{f_{(i,j)}, A_{ij}\right\}$. Different papers have then made different choices for the predictors $f_{(i,j)}$ and the predictive algorithm, as we will discuss in detail. But before, let us note another critical element, which is the construction of the dataset. Production networks are very sparse [3], so the ratio between the number of existing ($A_{ij} = 1$) and non-existing ($A_{ij} = 0$) links is very large. Therefore, training a model on the entire set of available examples might simply be computationally intractable (there are $\sim n^2$ pairs). Moreover, sampling a random subset would usually lead to poor predictions, because the scarce number of positive examples hinders the model's ability to effectively discriminate between the two classes. This phenomenon, known as the *class imbalance* problem, can potentially lead to models that are biased toward predicting the majority class, thus failing to accurately identify the existing links.

This problem is commonly addressed by applying *undersampling* (Fig. 1B), a technique that aims to rebalance the class distribution. In the context of production networks, undersampling involves carefully curating the training set to ensure a predetermined ratio between positive ($A_{ij} = 1$) and negative ($A_{ij} = 0$) examples. This controlled selection helps foster a more balanced, discriminative model and was employed in all the machine learning approaches that we are now set to survey.

	Coverage	Dataset	Inputs	Probabilistic
Mori et al. [18]	Regional	Tokyo Area Manufacturing Firms, Source unspecified	Several features regarding firms' activities, balance sheets, management	
Zuo et al. [19]	National	Tokyo Shoko Research	Firms' sales, profits, industrial sector, location, number of employees, network centrality	
Sasaki and Sakata [20]	Regional	Tohoku region, Teikoku Databank	Firms' sales, capital, size, industrial sector, network centrality	X
Lee and Kim [21]	National	Korean Enterprise Data	Description of firms' activities, firms' industrial sector and location, aggregate transaction volumes between industrial sectors	X
Brintrup et al. [14]	Automotive	Markline Automotive Information Platform	Firms' known connections, products, intermediate inputs	X
Kosasih and Brintrup [22]	Automotive	Markline Automotive Information Platform	Firms' known connections	X
Minakawa et al. [23]	Global	Asian bank's transaction data	Firms' known connection, description of firms' activities	X
Mungo et al. [24]	Global, National	Compustat, FactSet, Ecuador VAT	Firms' sales, industrial sector, location	X
Zhang et al. [25]	Global	Specialized Press (Reuters)	Media coverage	X
Wichmann et al. [12]	Global	Specialized Press	Media coverage	
Schaffer P. [26]	Global	Specialized Press	Media coverage	
Reisch et al. [27]	National	Phone calls, survey, Hungary VAT	Firms' phone calls, national IOTs	X
Hooijmaaijers and Buiten [28]	National, 4 commodity groups.	IOTs, Business Register, Structural Business Statistics	Firms' known connections, sales, geographic location, industrial sector	
Hillman et al. [29]	National	IOTs, Business Register, Structural Business Statistics	Firms' known connections, sales, geographic location, industrial sector	
Ialongo et al. [13]	National	Dutch banks' transaction data	Firms' sales, intermediate expenses by sector, network density (for calibration)	X
Mungo and Moran [30]	Global	FactSet	Firms' sales (time series), industrial sector, network sectoral structure	

TABLE I: Overview of the papers that reconstruct the supply network topology.

However, this procedure has implications for model evaluation. Typically, an algorithm is trained on a subsample (the training set), and evaluated on the remaining data (the testing set). If subsampling is done before the split into a testing and training set, the testing set will contain many more positives than a "real-life" testing set, so metrics such as accuracy will be severely biased. [24] found that metrics such as AUC were not substantially affected by the undersampling ratio, so we will tend to report AUCs, which are more comparable across studies. Many studies, however, report the F-score, which is highly dependent on class imbalance [24], so when reporting F-scores we will also report undersampling ratios.

2. Predicting new business partners

Interestingly, link prediction in production networks has not been originally pursued to reconstruct existing networks, but rather to build recommender systems that could suggest new partnerships to companies trying to expand their supplier or customer bases. In this framework, the ability of a model to identify existing (or past) supply-chain links is a target in so far as it is a proxy for their ability to make sensible recommendations, i.e., to identify *candidate* links that firms could turn into existing ones.

Despite aiming for different goals, these studies share several similarities with those on network reconstruction in the problem's layout, framed as a link prediction task, and the tools used, often relying on statistical models and network science.

Mori *et al.* focuses on $\sim$ 30k manufacturing firms in Japan. They build a business partner recommendation system by feeding a Support Vector Machine (SVM) with several companies' features, such as size, industrial sector, and geographic location. On a dataset comprising $\sim$ 34k links and an equal number of negative instances, they achieve an F-score of 0.85. The approach is refined in [19], who still use an SVM but add topological properties in the list of companies' features, such as their degree, betweenness centrality, and closeness centrality. For a network of 180k firms and half a million links assembled through the Tokyo Shoko Research dataset, and again an undersampling ratio of 1:1, they achieve an F-score of 0.81.

Sasaki and Sakata explicitly incorporate the network of second-tier suppliers and their respective industries, providing a more contextual analysis. The authors' intuition is that two firms within the same industry but with different suppliers will have different probabilities to sell to a specific customer. In other words, establishing a relationship between firms A (supplier) and B (customer) does not depend solely on the identity of A and B, but also on who are A's suppliers. Thus, the authors first extract from their network all the triads of firms connected in sequence (i.e., all the motifs $A \rightarrow B \rightarrow C$). Then, they replace each firm with its industrial sector (e.g., if we call S_i the industrial sector of firm i, the triplet $A \rightarrow B \rightarrow C$ becomes $S_A \rightarrow S_B \rightarrow S_C$), and use a Bayesian model called *n-gram* to compute the link probability between B and C given B and C's industrial sectors and the industrial sectors of B's suppliers. Finally, the authors use these probabilities as features in a random model classifier, together with a few firms' attributes (total revenues, number of employees, etc.) and network centralities. The authors focus on $\sim$ 50k links in a network of 130k Japanese firms[3], achieving an F-Score of 0.80 with an undersampling ratio of 1:1.

More recently, Lee and Kim integrated information on firms' geographical position and industrial sector with aggregate trade volumes between sectors and textual information on companies' activities and products. The authors encode this information and use it to train a deep neural network. On a sample of $\sim$ 90k connections between South Korean firms, where 20% of the examples are used as a test set, the authors achieve an AUROC of 0.92[4].

This trajectory of studies reflects a consistent evolution in methodology, with each iteration contributing incremental enhancements in feature integration and model sophistication, partially akin to what we will see now for papers which address supply network reconstruction specifically.

---

[3] The authors test their method on "new" links, missing from their 2010 snapshot of the network and present in the 2011 snapshot. The data is provided by Teikoku Databank Ltd., a business intelligence company.

[4] The authors do not specify the undersampling ratio of their exercise.

3. Can a firm better understand its supply network dependencies?

From a supply chain management perspective, a focal firm is interested in understanding hidden dependencies within its supply network - for instance, two suppliers may rely on a hidden "second tier" supplier, creating a vulnerability for the focal firm that is invisible at first sight. In such a context, the focal firm would typically see a fair part of the network and could use this topological information to make further inferences.

This is the context of the early investigation by Brintrup *et al.*[14], who focuses on the supply networks of three specific major car manufacturers (Jaguar, Saab, and Volvo, using data from the Marklines Automotive Information Platform). Using their domain expertise, the authors create four features for each potential link (i, j): *Outsourcing Association* (the overlap between the goods produced by company i and those bought by company j), *Buyer Association* (how frequently firms that purchase the same inputs as firm i also buy the products of firm j), *Competition Association* (the overlap between the products of firm i and those of firm j.), and *Degrees* (the number of partners of each firm). Training a logistic regression and a Naive Bayes using these features yields an AUROC of around 0.8.

In a subsequent paper [22], the authors refine their approach using Graph Neural Networks (GNNs) [31]. The concept underlying GNNs is that the network's topological information should not be distilled by the researchers through the design of specific features (as was the case with the association measures of the previous paper), but should instead be discovered automatically by the neural network. For production networks, the intuition is that the association measures designed in earlier work [14], while informative, might not convey all the information lying in the network's topology. Instead, a neural network provided with a sufficient amount of examples would identify patterns hidden to the researchers.

Practically, this is accomplished by: 1) for each link $l = (i, j)$, isolating subnetworks G_i, G_j composed by the nodes i and j, along with the set of their neighbours; 2) embedding each node u in the subnetwork $G_l = G_i \cup G_j$ in a vector $f_{u,l}$[5]; 3) feeding the nodes' embeddings $f_{u,l}$ to a series of K *graph convolutional layers*, which are nonlinear functions $f_{ul}^{k+1} = \phi\left(f_{ul}^k, \{k_u\}\right)$, where k_u are the degrees of the nodes in G_u; 4) averaging the final vectors $f_{u,l}^K$ across all the different nodes u, generating an embedding vector f_l' for the subnetwork G_l; 5) feeding the embedding through a sequence of fully connected layers to generate a single prediction for the probability p_{ij}.

The weights in the graph-convolutional and fully connected layers are trained with the usual backpropagation algorithm. The authors find a significant improvement compared to the previous approach, with the GNNs scoring an AUROC value ~ 0.95. While this is an impressive improvement in performance, a downside of this approach is that it becomes very difficult to interpret the predictions made by the neural network and develop novel insights into how firms connect.

A similar approach is proposed in [23], where the authors train a graph neural network with topological information and textual information on firms' activities, encoded via the Doc2Vec algorithm [32]. On a network of 170k firms and 1.2M edges provided by a large Asian bank, the authors AUROC of 0.94-0.95, depending on the respective sizes of the training and the test data. They do not report the undersampling ratio.

4. Predicting the supply networks of entire countries where no network data exist

Mungo *et al.* use similar methods for a different purpose. They observed that in some countries, excellent data is available, while in other countries (including the US), there is no fully reliable information on firm-to-firm transactions, creating a need for methods that predict the supply network using only information available locally (Hooijmaaijers and Buiten [28], reviewed in Section IV B, first developed a method based on data held by the most statistical offices). Based on this observation, they ask whether a model trained on the production network of a country A accurately predicts links within firms in another country B.

---

[5] The embedding usually consists of computing an average distance d between node k and the nodes i and j, and then embedding k in a vector $f_{ij}^k = \delta_{dd'}$. The dimension of this vector is the maximum possible distance, which must be specified as a parameter of the model.

In all countries, there is usually good data available on key features of firms and pairs of firms that could determine link formation. For example, it is well established that large firms have more connections [3], prefer to trade with geographically closer firms [6, 33], and have production recipes that put significant constraints on the inputs they buy. Based on these hypotheses, for each candidate link, the authors build a vector $f_{(i,j)}$ containing information on firms' sales, industrial sector, and geographical distance. They then train a *gradient-boosting* model to predict link probability.

The study is run on three different datasets: two commercial, global datasets (*Compustat* and *FactSet*) and one dataset covering (a subsample of) Ecuador's national production network, assembled by Ecuador's government using VAT data. When tested on the same dataset used to train the model, the approach scores an AUROC similar to that of the previous approach (from ~ 0.91 to ~ 0.95 depending on the dataset), suggesting that indeed, knowing a firm's products, location and size provides sufficient information to make decent predictions.

For making predictions on unobserved countries, they conduct two tests. In the first test, they considered different countries in the same dataset, for instance training their model on FactSet's US and Chinese networks and predicting links in Japan. In this case, the approach still performs relatively well (AUROC > 0.75). In the second test, they predict the links in Ecuador using FactSet and the other way around. Here, the performance deteriorates substantially, which the authors explain by showing that the distribution of features in FactSet (an incomplete, commercial dataset with large firms in rich countries) and Ecuador (a complete administrative dataset, with all firms from a developing economy) are very different.

This partial success suggests that there is potential for further studies, but using multiple administrative datasets. For instance, while it is not possible to predict the Ecuadorian administrative data using the commercial data from FactSet, it might still be possible using similar administrative datasets, given the results from [3] showing that administrative datasets exhibit strikingly similar topological properties. This is a straightforward approach to reconstructing the global firm-level production network, using training data from a few countries, and large-scale firm-level datasets such as ORBIS.

5. *Leveraging alternative data: news and phone calls*

The idea in [25] and [12] is that significant commercial deals might be announced in press releases or covered by the specialized press.

[25] build a system to automate the analysis of articles and investor comments coming from Reuters and identify collaborative[6] and competitive relationships between companies. The authors web-scrape a corpus of $\sim 125k$ documents and manually annotate a sample of $4.5k$, overall identifying 505 relationships. Then, they use a Latent Dirichlet Allocation (LDA) algorithm (a widely used algorithm in text analysis) to examine these examples, finding that the algorithm identifies collaborative relationships with an AUROC of 0.87.

Similarly, [12] automates the analysis of textual data (coming from Reuters Corpora TRC2 and RCV1, NewsIR16, and specific web searches) to find mentions of commercial deals between the firms. First, the authors collect a text corpus describing the relationships between firms. Then, they classify these relationships as either a commercial relationship (e.g., firm i supplies firm j), an ownership relationship (firm i owns firm j), or none of the previous. The annotated examples are embedded into numerical vectors using the word embeddings in the Glove dataset and finally used to train a Natural Language Processing (NLP) classifier with a BiLSTM architecture. 30% of the sentences were left out of the data and used to assess the performance of the model, which scores an F1-score of 0.72 with a class imbalance of 1:7. Unfortunately, the choice of evaluating the score on a binary metric (the F1-Score) does not allow a straightforward comparison with the previous approaches. However, the authors report that a random classifier would get an F1-Score of 0.38. In a follow-up paper [26], the authors improve their results by running the same study using a BERT model, and reach an F1-Score of 0.81.

In [27], instead, the authors use phone calls between companies and survey data to track down supplier-customer relationships in an undisclosed European country. The survey asked companies to list their ten

---

[6] Note that, for the authors, a "collaborative relationship" has a broader meaning than supply relationship.

most important suppliers and customers. On this subsample of the network, the authors find that if the average daily communication time between two firms i and j, denoted τ_{ij}, is greater than 30 seconds, the probability that these two firms are connected is $p_{ij} \approx 0.9$. Equipped with this observation, the authors reconstruct the network by first assuming the presence of a link between i and j if $\tau_{ij} > 30s$ and then assigning a direction to the link stochastically with a probability

$$p\left(i \rightarrow j\right) = \frac{\omega_{a_i b_j}}{\omega_{a_i b_j} + \omega_{b_j a_i}},$$

where a_i and b_j are i and j's respective industrial sector, and ω_{ab} is the total amount of trade (in monetary value) from firms in sector a to firms in sector b, as reported in the country's Input-Output tables[7]. The authors do not provide any 'standard' evaluation metric for their reconstruction. However, they mention that choosing a threshold $\tau_{ij} = 30s/d$ minimizes the Kullback-Leibler divergence between the degree distribution of the reconstructed network and the degree distribution of a well-studied network, the Hungarian production network. The authors' ultimate goal was to compute firms' Economic Systemic Risk Index (ESRI, see Section II C) in the reconstructed network, and they do find a good qualitative agreement between the ESRI sequence of firms in the reconstructed and the Hungarian network.

B. Network Inference

A second stream of research tries to reconstruct the production network as a whole rather than link-by-link. We distinguish three sets of approaches: matching algorithms, maximum entropy methods, and methods based on time series correlations.

1. Matching algorithms

A couple of papers have used matching algorithms to create supply networks. We classify these under "Network Inference" because while they reconstruct the network link-by-link, they typically try to match aggregate constraints, taken from I-O tables and/or from meso-level statistics published independently.

An early study is the one from Hooijmaaijers and Buiten ([28], see [34] for details), who devise an algorithm that matches firms based on commonly observable firm characteristics (industry, size, location) and I-O tables.

Roughly speaking, their method works as follows. First, using a relationship between sales and degrees of $s_i \propto k_i^{1.3}$ [35], they can estimate out-degrees based on total sales. For expenses, using the I-O tables they can estimate expenses of each firm by industry, and assuming that in-degree by industry is a (specific) increasing function of expenses by industry, they can estimate the number of industry-specific suppliers for each firm.

Knowing the degrees of all firms, the next task is to match them. To do this, they create pairwise scores based on assumptions about what determines the likelihood of a match. The final score is a linear combination of three scores: one that increases with firm size, one that decreases with distance, and one that acts as a bonus or penalty if the firms are in industries that trade in I-O tables. The matching algorithm then starts with the buyer that has the highest purchasing volume and goes in descending order. The number of suppliers connected to each buyer is determined by the buyer's in-degree. Among the potential suppliers, those with the highest scores are considered the most likely to trade with the buyer. If any of these top-rated suppliers have no remaining outgoing links, then the next most likely supplier in line is considered instead.

Hillman et $al.$ introduced another algorithm, driven by their need to create a synthetic firm-level network for their agent-based model of the impact of the Covid-19 pandemic. Again, their method makes use of

---

[7] A consequence of the algorithm choosing edge direction is that the reconstructed network has null reciprocity, while we know that real networks exhibit reciprocity of around a few percent [3].

I-O tables and data on sales, although it does not use location information. Their algorithm is less clearly documented, but essentially works by first using I-O tables to determine which industries a firm should sell to, then allocating chunks of its sales to randomly selected firms in the buying industry. They show that their algorithm is able to reproduce a positive strength-degree relationship.

2. Maximum-entropy for network inference

In a sense, matching algorithms try to distribute connections "randomly", while matching some aggregate properties of the network. However, to do so they introduce "plausible" assumptions, such as specific functional forms to create scores. Instead of introducing assumptions, the Maximum Entropy assigns a probability to each possible network in a "maximally non-committal" way. This leads to the question of whether introducing assumptions about what is not fully known is better than just maximizing entropy conditional only on what is fully known. This is the question of Rachkov *et al.*, who showed that the networks obtained from the matching method proposed in Ref. [28] have different properties than those obtained using a simple maximum-entropy model, suggesting possible biases in heuristics-based reconstructions. That being said, simple maximum entropy methods are not well-suited for complete supply networks (i.e., not commodity-specific), because they do not use information on firms' products, which we know is a critical determinant of their probability to link.

Ialongo *et al.* introduced a method that tackles this issue and simultaneously reconstructs the whole network topology and link weights (see Sec. IV for the weights). Following a long-standing tradition in network reconstruction [16], they compute a probability distribution $P(G)$ over the set of all possible graphs $\mathcal{G}$ that maximizes the Shannon Entropy $\mathcal{S}$,

$$\mathcal{S} = -\sum_{G \in \mathcal{G}} P(G) \ln P(G).$$

The maximization is subject to a normalization constraint, $\sum_{G \in \mathcal{G}} P(G) = 1$, and a collection of constraints $\tilde{\boldsymbol{c}}$ representing the macroscopic properties enforced on the system. These constraints are usually enforced in a soft way, that is, by constraining the expected values of the constraints over the set of all possible networks $\mathcal{G}$,

$$\sum_{G \in \mathcal{G}} P(G) c_i(G) = \tilde{c}_i.$$

The authors expand on a pre-existing model [36], constraining the network's density ρ, each firm's total sales ω_i^{out} and the money spent by firm i on inputs from each industrial sector a, $\{\omega_{a \to i}\}$. However, as we have already emphasized, a crucial feature in supply networks is that firms connect to others specifically for the products they make. A method that does not take into account the product or industry of the firm is, in the context of supply networks, doomed to fail.

As a result, the authors design a new model able to handle sector-specific constraints. For instance, in a hypothetical economy with two sectors, a and b, the model enforces three constraints on each firm: one for total sales, $\sum_{G \in \mathcal{G}} P(G) \omega_i^{out} = \tilde{\omega}_i^{out}$ and one for spending on each of the sectors: the money spent on inputs from sector a, $\sum_{G \in \mathcal{G}} P(G) \omega_{a \to i} = \tilde{\omega}_{a \to i}$, and the spending on inputs from sector b, $\sum_{G \in \mathcal{G}} P(G) \omega_{b \to i} = \tilde{\omega}_{b \to i}$ (we use tildas to denote observed quantities). The model accepts an analytical solution for the marginals p_{ij},

$$p_{ij} = \frac{z \tilde{\omega}_i^{out} \tilde{\omega}_{a_i \to j}}{1 + z \tilde{\omega}_i^{out} \tilde{\omega}_{a_i \to j}}, \tag{2}$$

where a_i is the industrial sector of firm i, and z is chosen such that $\sum_i \sum_{j \neq i} p_{ij} = \tilde{\rho}$.

The authors show that their method significantly improves upon the model by [36], where each firm is subject to a single constraint for the overall intermediate expenses. In a maximum-entropy framework, imposing only one constraint on the intermediate expenses would distribute a firm's supplier equally across

all industrial sectors. This is at odds with the reality of supply chains, where firms require only a select range of goods from the basket of products available in an economy.

The authors do not report any standard reconstruction metric, but they show that the in-degree and out-degree distribution of the reconstructed network are, in expectation, in good agreement with the empirical degree distribution. Moreover, the relationship between degrees and strengths of firms is generally well replicated.

A limitation of all the studies discussed so far is that they consider only firm-to-firm links. For macroeconomic applications, it would be useful to reconstruct complete synthetic populations (see Sec. VI), including links between firms (including banks) and consumers/workers. Hazan uses a maximum-entropy approach (more precisely, the fitness-induced configuration model, [38]) for firm-to-firm networks and firm-to-consumer networks, taking average degrees from the literature to estimate z separately in each network.

3. Leveraging the correlation matrix using graph learning

An established literature tackles the problem of reconstructing a network starting from N node-level time series encoded in vectors $x(t) \in \mathbb{R}^N$ [39, 40]. The general philosophy is that the structure of the network $\mathcal{G}$ determines the joint probability distribution of the observations. If one assumes that each observation $x(t)$ is drawn from a probability distribution $p(x|\Theta)$ with a parameter matrix $\Theta \in \mathbb{R}^{N \times N}$, the problem of reconstructing a graph, or *graph learning*, becomes that of finding the correct value of Θ.

Production networks serve as a contagion channel for economic shocks. They spread negative or positive shocks from one firm to its customers and suppliers, generating correlations between firms' fundamentals, such as market valuation and sales [41–43]. Starting from this observation and leveraging the graph learning literature, Mungo and Moran introduce a method to reconstruct the production network from the time series of firm sales, $s_i(t)$. First, the authors show empirically that the correlation between the log-growth rates of firms connected in the production network surpasses the average correlation yielded by randomly sampled firm pairs, and this excess correlation decreases as firms get further apart in the supply chain. Then, the authors harness this observation to design a network reconstruction approach, framed within Gaussian Markov Random Fields [39]. Adapting a modern graph learning strategy [44], the authors assumed that the growth time series data could be modelled as a sequence of draws from a multivariate Gaussian distribution. This distribution's precision matrix (the inverse of the covariance matrix) is, in turn, identified with the network Laplacian $L = D - A$ where $D_{ij} = k_i \delta_{ij}$. To estimate the precision matrix, the authors employed a maximum likelihood approach, constraining the possible Laplacians L to preserve the expected connections' density within and across economic sectors. In addition, a penalization term is included to enforce network sparsity.

Upon assessment against smaller network fragments, their methodology reports an F1-score within the range of $0.2 - 0.3$. Nevertheless, it does not consistently surpass all benchmark tests under consideration. While it is true that, on average, firms that are more closely connected are more correlated, there is a lot of overlap between the distributions of correlations at various distances. In other words, knowing that firms are highly correlated is not very informative of their distance, making the task of network inference based on time series data very challenging.

IV. INFERRING THE VALUE OF TRANSACTIONS

While methods for reconstructing weights have been used extensively on financial and global trade networks [e.g. 16, 45, 46] and aggregate I-O tables [e.g. 47], their application to firm-level networks is relatively novel. A first set of methods uses meso-level information from I-O tables, while another set of papers relies on the maximum entropy principle.

A. Matching I-O tables

Inoue and Todo incorporate aggregate I-O information into their estimates of the weights in the supply network of Japan. They assign to each link between a supplier i and a customer j a weight proportional to the product of firm sales, $\omega_{ij} \propto \tilde{\omega}_i^{\text{out}} \frac{\tilde{\omega}_j^{\text{out}}}{\sum_{j \in \mathcal{N}_i} \tilde{\omega}_j^{\text{out}}}$, where $\sum_{j \in \mathcal{N}_i}$ means that the sum runs only on i's customers. The weights are then rescaled to align with the aggregate transaction amounts within industry sectors $\tilde{\omega}_{ab}$,

$$\omega_{ij} = \tilde{\omega}_i^{\text{out}} \frac{\tilde{\omega}_j^{\text{out}}}{\sum_{j \in \mathcal{N}_i} \tilde{\omega}_j^{\text{out}}} \frac{\tilde{\omega}_{a_i b_j}}{\sum_{k \in a_i, l \in b_j} \tilde{\omega}_k^{\text{out}} \tilde{\omega}_l^{\text{out}}},$$

where a_i and b_j denote the respective industrial sectors of i and j. A similar approach has been used by [29] where, starting from data on firms' sales and inputs, the authors construct individual-firm networks, that, when aggregated, align with the sectoral IO table. The authors rescale firms' input and output to match the IO tables[8], and then allocate links in the network with an iterative algorithm that matches buyers to suppliers, while also imposing that larger firms will have more customers. The weight of each connection is then set to the smallest value between the supplier's maximum capacity and the customer's demand.

Instead of reconstructing the weights, Carvalho *et al.* estimate the *input shares* α_{ij} of each link,

$$\alpha_{ij} = \frac{\omega_{ij}}{\sum_i \omega_{ij}}.$$

For any given customer-supplier pair of firms (i, j) in the data, they assign α_{ij} proportionally to the input-output table entry corresponding to industries i and j belong to, i.e., $\alpha_{ij} \propto \tilde{\omega}_{a_i b_j}$, and renormalize them to ensure $\sum_i \alpha_{ij} = 1$.

Real-world scenarios often present situations where it is unfeasible to find weights that align with aggregate observations. In [49], the authors design an inference strategy that aims to minimize the discrepancy between reconstructed and observed aggregate properties of the network. More specifically, the authors observe that, given a binary network G, it is not always possible to assign weights ω_{ij} that satisfy constraints $\sum_j \omega_{ij} = \tilde{\omega}_i^{\text{out}}$ and $\sum_j \omega_{ji} = \tilde{\omega}_i^{\text{in}}$. Take as an example a firm i who supplies only a single firm j, and assume that i is the only supplier of j. The aggregate constraints will only be satisfied if i's sales match exactly j's expenses, $\tilde{\omega}_i^{\text{out}} = \tilde{\omega}_j^{\text{in}}$, a condition not always respected in the data. The authors solve this issue by introducing a 'residual node' r to capture the portion of the economy that is not covered by the network G. This node accounts for all the firms that are not present in the data. They propose to find the set of weights ω_{ij} that minimize the loss $\mathcal{L} = \sum_i \omega_{i,r} + \sum_i \omega_{r,i}$, where ω_{ij} are subject to the usual constraints.

Finally, Hazan reconstructs the weights for a complete stock-flow consistent economy, with households, firms, banks, and flows of money in the form of consumption, firm-to-firm payments, wages, and interest payments. After reconstructing the network using maximum entropy methods (Sec. III B 2), stock-flow consistency makes it possible to write a linear system for the weights, which can be solved using Non-Negative Least Squares to avoid negative values.

The performance of the methods reviewed in this subsection is unfortunately unknown, as information on the real weights was not available to the authors, who could not compare their reconstructions to the respective ground truths. However, in the future, researchers using these methods could partially validate their results by comparing them to the empirical regularities observed in [3] for weight distributions and the relationships between in- and out-degrees and strengths.

---

[8] More precisely, they match intermediate inputs (roughly, inputs that are neither labour nor investment goods), and gross output (roughly, total sales).

	Coverage	Dataset	Inputs	Probabilistic	MaxEnt
Inoue and Todo [48]	National, Japan	Tokyo Shoko Research	Firm sales, national IOTs		
Carvalho *et al.* [43]	National, Japan	Tokyo Shoko Research	Firm sales, national IOTs		
Welburn *et al.* [49]	National, US	S&P Capital IQ, EDGAR.	Firm sales and inputs (COGS).		
Hazan [37]	National, Czech Republic	Full IOTs	Full IOTs		
Bacilieri and Astudillo-Estevez [50]	National, International	Factset, Ecuador VAT	Firm sales, intermediate expenses, network density	X	X
Ialongo *et al.* [13]	National	Dutch banks' transaction data	Firm sales, intermediate expenses by sector, network density (for calibration)	X	X

TABLE II: Overview of the papers that infer supply network weights.

B. Maximum entropy for weights inference

Another way of predicting weights given some aggregate trade information is to use the maximum entropy principle. The intuition behind this principle is computing a distribution that is *maximally non-committal* with respect to unknown information [51]b or, in simpler words, to build a distribution that minimizes unjustified assumptions about the network. In Sec. III B 2, we saw how maximum entropy can be used to compute probabilities for possible binary networks. We are now going to see how it can be used to predict weights.

If we consider the weights ω_{ij}, subject to the ("hard") constraints $\sum_j \omega_{ij} = \tilde{\omega}_i^{out}$, and $\sum_j \omega_{ji} = \tilde{\omega}_i^{in}$, where $\tilde{\omega}_i^{out}$ and $\tilde{\omega}_i^{in}$ represent the observed total outflow (intermediate sales) and inflow (intermediate expenses) of firm i, we find that the set of weights that maximize the Shannon Entropy

$$\mathcal{S} = -\sum_i^N \sum_j^N \omega_{ij} \ln \omega_{ij},$$

are

$$\omega_{ij} = \frac{\tilde{\omega}_i^{out} \tilde{\omega}_j^{in}}{\tilde{\Omega}}, \tag{3}$$

where $\tilde{\Omega} = \sum_i \tilde{\omega}_i^{out} = \sum_i \tilde{\omega}_i^{in}$. This approach was also used in [27] for an undisclosed European country[9].

A different application of the maximum-entropy principle, where constraints are imposed softly (see Sec. III A), results in the solution used in [50] to reconstruct Ecuador's national production network and in [13] to reconstruct the transaction network between customers of two Dutch banks. Building on [36], these papers first reconstruct the network's topology[10], then sample the (positive) weights ω_{ij} of the existing links from an exponential distribution,

$$P(\omega_{ij} = x) = \beta_{ij} \exp(-\beta_{ij} x),$$

---

[9] Bartolucci *et al.* show that "upstreamness", a classic metric in I-O economics, can be recovered very well from networks reconstructed from maximum entropy, as long as the networks are not too sparse. This is because, under very general conditions for the original network, the first-order approximation of a node's upstreamness is its upstreamness in the maximum entropy-reconstructed network [53].

[10] In the case of [3], the topology is assumed to be known.

where β_{ij} is selected so that the expected value of ω_{ij}, conditional to the existence of a link, is

$$\mathbb{E}_{ij}\left[\omega_{ij}|A_{ij}=1\right] = \frac{\tilde{\omega}_i^{out}\tilde{\omega}_j^{in}}{p_{ij}\sum_i \tilde{\omega}_i^{out}}.$$

In [13], p_{ij} is defined by Eq. (2). In contrast, [50] omits sector-specific constraints for intermediate inputs[11], and defines p_{ij} as

$$p_{ij} = \frac{z\tilde{\omega}_i^{out}\tilde{\omega}_j^{in}}{1 + z\tilde{\omega}_i^{out}\tilde{\omega}_j^{in}}.$$

Ref. [50] reports a cosine similarity of 0.928 between inferred and actual weights, and also compute a few "higher-order" properties of the nodes that describe the propagation of shocks in production networks in an established macroeconomic model [54], which the reconstructed network fails to capture adequately (the cosine similarity for the most relevant property, the *influence vector*, is ~ 0.5).

In [13], visual inspection of the results shows a substantial enhancement in weight reconstruction when applying sector-specific constraints to firms' inputs, further underscoring the crucial role the economy's sectoral structure plays in the accurate reconstruction of production networks.

V. DISCUSSION

In this section, we take stock of what we can learn from existing studies, and provide suggestions on how the field could be further advanced.

A. What have we learned?

A first, clear message from the review is that in the context of supply networks, knowing the kind of product a firm makes is extremely important and substantially improves the reconstruction. This is evident both in the link prediction studies on industry data [14], commercial or country-level data [24], and in the maximum entropy reconstruction on payment data [13]. Unsurprisingly, ongoing research tries to predict the firms' products at a granular level, for instance from websites [55].

Second, the importance of products leads us to ask: to what extent can we, or should we rely on existing (national or inter-country) input-output matrices? While some studies reconstruct weights (conditional on links) using I-O links [29, 43, 48], others refrain from doing so [50], by fear that differences in accounting conventions [3] may create inconsistencies. Here the answer may depend on the goal of the reconstruction (see next section). A useful avenue for further research, however, would be to develop methods that easily make it possible to switch between business- and national accounting conventions. Such methods would necessarily use techniques and assumptions to allocate flows of money based on partially observed data, so that the methods reviewed here may be helpful.

Third, we have seen that more sophisticated machine learning methods do provide substantial boosts in performance. This is clear from the improvement in link prediction performance between the logistic regression and graph neural nets in the automotive dataset [14, 22], and between simpler methods and gradient boosting in Mungo *et al.* [24] [12].

Fourth, there appears to be substantial scope for improving performance using "alternative" data. Zhang *et al.* [25] and Wichmann *et al.* [12] have provided a proof of concept that mining news and websites for supplier-buyer relations can be automated, and we have already mentioned that websites

---

[11] [13] simply assume that the meso-level constraints are observable since they have this in their firm-level data. [29, 43, 48] cannot read this information from the data, so they take meso-level information from the I-O tables. [50] argue that differences in accounting standards between firm- and industry-level networks are large so that the meso-level structure of a firm network should not be constrained to be like the I-O tables. [3] shows that there are indeed some important differences, especially in industries that follow different accounting conventions, such as retail and wholesale trade.

[12] However, in both studies, predictions made by sophisticated models are harder to interpret.

can be an important source of key metadata for link prediction (especially product-related information). While phone data is likely to be difficult to access, it is worth remembering the impressive result in [27] that firms with average daily communication of more than $30s$/day have a 90% probability of being connected.

A related question for further research will be to establish the potential of "dynamical" data. Mungo and Moran [30] showed that while there is information about the network in the sales growth rates correlation matrix, predicting the network remains difficult, as the distribution of pairwise correlation for connected and unconnected pairs overlaps greatly, even though their average is statistically significantly different. Nevertheless, there are interesting developments in this area for networks generally, with only one application to supply networks. One limitation has been that very few supply networks' datasets have a reasonable time-series dimension, but as these become more common it will perhaps become possible to find other firm-level dynamical features that contain fingerprints of their network.

Finally, many studies have shown that baking sensible economic intuition into the models usually improves predictions. To sum up, we have learned (or confirmed from existing literature) that link formation is likely driven by the kind of products firms make, their geographical distance, and their size. We have seen that firms who communicate a lot are likely to be in a supply-buy relationship and that firms that are in a relationship are likely to have a substantial co-movement in sales. While prediction is in some cases the ultimate goal, making methods that prioritize performance over interpretability appropriate [22], the quest for better reconstruction models has also prompted a deeper investigation into the behavioural and economic principles influencing how firms make and unmake their connections [14, 24]. Currently, no fully realistic supply network formation model has been developed (however, see [56] for an early example); we anticipate that reconstruction methods and the development of null models will, at least partly, go hand in hand.

B. How can we learn more?

What method works best for which task? We are not yet able to properly answer this question because the literature uses different datasets, takes different features of the data to make predictions, and uses different evaluation metrics. While this is warranted by the diversity of goals and applications, we think it would be valuable to organize "horse races", as has been done for financial networks [45], and provide standard datasets, as is common in the machine learning community.

Let us first discuss the lack of comparability between studies. The methods proposed are very diverse and usually require distinct data to operate. The diversity of datasets and features used is understandable and valuable. For example, Kosasih and Brintrup [22] use topological features because one of their realistic use cases is to augment an existing "observed" network dataset, while Mungo *et al.* [24] avoid using topological information because their envisioned use case is to port a trained model to a context where no such features are available. As another example, while phone data is very hard to access, the study using this data made it possible to evaluate the systemic risk of each firm in an entire European country.

A slightly less justified "diversity of approaches" is the lack of standardized assessment metrics, as it is in principle relatively easy to report several metrics.

Traditional statistical indicators (accuracy, AUROC, PR-AUC) provide an easy, well-known benchmark, and have already been functional in, e.g., propelling the development of computer-vision models [57]. Yet, the question remains as to whether they are sufficient to evaluate the reconstruction of a network, and what additional metrics should be adopted to supplement them. Some metrics, initially conceived for balanced datasets, may not hold up as reliably when applied to sparse networks, where non-existing links greatly outnumber the existing ones, further complicating the comparison between methods. Overall, the area under the Receiving Operator Characteristic Curve (AUROC) seems robust in the face of class imbalance: if one makes the imbalance more and more severe, its value does not change substantially (see Supplementary Material [24]). Consequently, AUROC is a sensible metric to compare results. The area under the Precision-Recall curve (PR-AUC), which is more sensitive to the performance of the model on the minority class, is also very sensitive to the level of imbalance in the data; PR-AUC and imbalance should always be reported jointly.

Reporting basic topology metrics of the reconstructed network is also a sensible approach, as there is

substantial evidence [3] that some topological properties are universally shared by all production networks. For instance, Bacilieri *et al.* [3] showed that the tail exponents for the in- and out-degree distributions are remarkably similar in national, VAT-assembled datasets.

Ultimately, as we plug reconstructed networks into economic models, the optimal metric will be the one that best correlates with accurate economic predictions. Identifying these proper "dynamical" indicators needs to go hand-in-hand with the development of economic models that are carefully validated on real-world data and can become legitimate standards for evaluating reconstruction performance.

While agreeing on a set of metrics and features appears relatively easy, the key challenge ahead is data availability. To follow our previous analogy, in computer vision, researchers can access standard, large-scale datasets [58] of annotated images to train and evaluate their models. Similar datasets for production network reconstruction are not currently available and, due to the confidential or proprietary nature of such data, its assembly seems unlikely in the near future. The research community should unite to devise strategies to circumvent this issue, possibly by considering the use of synthetic data [59] as an alternative to real data. While synthetic data generation is currently an active and exciting area of research, it is less well-developed for networks than for tabular data and still suffers from either a lack of privacy guarantees (for traditional methods) or a lack of interpretability of the privacy guarantees (for differential privacy).

VI. TWO RESEARCH AGENDAS

For many practical applications, it is necessary to know much more than the value of transactions between firms. We lay out two research programs - one that aims to reconstruct supply networks to allow for real-time monitoring of disruptions and logistics optimization; and one that aims to reconstruct a granular version of global macroeconomic datasets.

A. Towards supply chain visibility for risk management

Past decades have been characterized by supply chain cost optimization objectives, which have led to just-in-time initiatives that stripped buffer inventories from supply lines, that have already become geographically longer with offshoring practices.

While high-impact, rare events such as COVID-19 highlighted the vulnerability of these global, highly complex modes of operation, organisations often struggle with increased volatility in their day-to-day procurement. Supply chain researchers are increasingly seeking methods to build resilience in what is now frequently termed a "shortage economy" [60]. However, these efforts are often hindered by a lack of visibility on supply chain dependencies as companies do not disclose commercially sensitive information such as whom they buy goods and services from.

As link prediction and reconstruction methods presented in this paper do not rely on companies' willingness to share data, they have the potential to become a primary toolset in supply chain risk management. Our review shows that buyer-supplier link prediction is possible with various differing methodologies and success rates. Recently proposed methods for reconstructing knowledge graphs go beyond who-supplies-whom, but also enable prediction of other types of relevant information such as where firms are located, and what they produce, paving the way for a new era of "digital supply chain surveillance" [61].

Much further work is needed in this context. For instance, use cases that evaluate how the identification of risky supplier locations and production dependencies might help with effective mitigation strategies such as multi-sourcing, supply chain reconfiguration, insurance, or inventory buffers. Beyond addressing supply disruption risk, an understanding of supply chain structure could be informative for the detection of supply chain fraud and counterfeit products. Improved visibility may help improve regulatory compliance on the Environmental, Social and Governance (ESG) practice. Methods that help detect transaction volumes could improve supply chain financing in which lenders often struggle with identifying financial risk exposure. To achieve these, different ontologies are needed to be built and integrated into existing knowledge graph completion methods. New methods for inferring compliance, fraud, and other areas of interest from knowledge graphs need to be developed. Lastly, any resulting graph will be limited by

underlying assumptions and incomplete data, which, in turn, may be shaped by the observable data at hand. Hence data imputation and uncertainty quantification will need to inform the resulting graphs.

B. Towards granular global economic accounts

For macroeconomic applications, our interest extends beyond the mere flow of money between firms. Macroeconomics concerns quantities such as GDP, which represents at the same time the total income, total expenditures, and total "value added" of an economy. Firm-to-firm transactions are not sufficient to truly understand how economic agents create value, redistribute income, and spend on goods and services.

As a result, to support the development of large-scale realistic agent-based models, we need an ambitious agenda to develop semi-synthetic populations, which would include all the available micro information and supplement it by synthetic micro data in a way that results in meso- and macro-level aggregates compatible with observed meso- and macro-level quantities typically observable from national accounts. We elaborate briefly on three strands of research within this agenda.

First, it will be important to ensure compatibility between micro and meso-level data, which are usually compiled using different accounting rules. National accounting principles provide a solid conceptual framework, so developing reconstructed datasets that respect these principles would have many advantages, including making it easier to use this data in macro models, easier to improve this data using macro-level information, and easier to match this data with other relevant datasets. However, firm-level data is usually compiled using business accounting rules, so that simply "summing up" firm-level data does not necessarily lead to the supposedly equivalent concept in national accounts. As we have highlighted, this is a potential to, for instance, use IOTs as additional information to reconstruct firm-level networks.

Second, modern work in economics shows that employer-employee-matched data and datasets on consumer baskets' heterogeneity are crucial to understanding inequality, long-run growth, or carbon emissions. As a result, a straightforward extension of the "reconstruction of economic networks" program would be to predict employer-employee relations and consumer-firm relations (See [37] for a first attempt). Existing efforts to develop data-driven agent-based models rely on such synthetic populations. While there exists a lot of work on recommender systems for suggesting products to consumers, and more recently some work on career suggestions, these efforts have not been leveraged to create reliable synthetic populations.

Third, many of the studies presented here worked with money flows, omitting a distinction between prices and quantities. This is driven by the fact that firm-level supply networks with both price and quantity information are very rare, but this is a serious issue for economic modelling, where prices obviously play a key role. To model inflation, understand growth and business cycles, we need measures of quantities produced (or inflation-adjusted values). New methods for inferring prices, perhaps based on companies' websites and other features, would be very extremely helpful in this context.

VII. CONCLUSION

The reconstruction of supply networks through mathematical methods is a young field. This paper offers a review of methodologies that researchers have proposed to grapple with this challenge.

Good proof-of-concept studies exist, but much remains to be done. A striking feature of the literature is the diversity of methods, datasets and evaluation metrics. While this is justified by the different backgrounds and motivations of the researchers, we think that progress in this area would benefit from the availability of open datasets and the definition of standard metrics, so that horse races could be organised.

We were able to propose some guidelines to standardize performance metrics, but the path to open datasets is more complicated and will require international cooperation that either facilitates researchers' access, or fosters the creation of high-fidelity synthetic datasets.

Despite this difficulty, we think that reconstructing supply networks is an excellent playing ground for the complex systems community, as it requires a deep understanding of networks, statistics, and dynamical systems, together with an appreciation that these networks emerge from the decentralized interactions of

millions of highly heterogenous, bounded-rational agents operating with different objectives at different time scales.

[1] A. G. Haldane and R. M. May, Systemic risk in banking ecosystems, Nature **469**, 351 (2011).
[2] M. Bardoscia, P. Barucca, S. Battiston, F. Caccioli, G. Cimini, D. Garlaschelli, F. Saracco, T. Squartini, and G. Caldarelli, The physics of financial networks, Nature Reviews Physics **3**, 490 (2021).
[3] A. Bacilieri, A. Borsos, P. Astudillo-Estevez, and F. Lafond, *Firm-level production networks: what do we (really) know?*, Tech. Rep. 2023-08 (Institute for New Economic Thinking, 2023).
[4] M. E. J. Newman, Assortative mixing in networks, Phys. Rev. Lett. **89**, 208701 (2002).
[5] B. Demir, A. C. Fieler, D. Y. Xu, and K. K. Yang, O-ring production networks, Journal of Political Economy **0**, null (0), https://doi.org/10.1086/725703.
[6] A. B. Bernard, A. Moxnes, and Y. U. Saito, Production networks, geography, and firm performance, Journal of Political Economy **127**, 639 (2019), https://doi.org/10.1086/700764.
[7] M. Christopher and M. Holweg, "supply chain 2.0": Managing supply chains in the era of turbulence, International journal of physical distribution & logistics management **41**, 63 (2011).
[8] T. Y. Choi and Y. Hong, Unveiling the structure of supply networks: case studies in Honda, Acura, and DaimlerChrysler, Journal of Operations Management **20**, 469 (2002).
[9] A. Brintrup, A. Ledwoch, and J. Barros, Topological robustness of the global automotive industry, Logistics Research **9**, 1 (2016).
[10] A. Brintrup, Y. Wang, and A. Tiwari, Supply networks as complex systems: a network-science-based characterization, IEEE Systems Journal **11**, 2170 (2015).
[11] S. Perera, M. G. Bell, and M. C. Bliemer, Network science approach to modelling the topology and robustness of supply chain networks: a review and perspective, Applied network science **2**, 1 (2017).
[12] P. Wichmann, A. Brintrup, S. Baker, P. Woodall, and D. McFarlane, Extracting supply chain maps from news articles using deep neural networks, International Journal of Production Research **58**, 5320 (2020), https://doi.org/10.1080/00207543.2020.1720925.
[13] L. N. Ialongo, C. de Valk, E. Marchese, F. Jansen, H. Zmarrou, T. Squartini, and D. Garlaschelli, Reconstructing firm-level interactions in the dutch input–output network from production constraints, Scientific Reports **12**, 11847 (2022), number: 1 Publisher: Nature Publishing Group.
[14] A. Brintrup, P. Wichmann, P. Woodall, D. McFarlane, E. Nicks, and W. Krechel, Predicting hidden links in supply networks, Complexity **2018**, 9104387 (2018).
[15] L. Lü and T. Zhou, Link prediction in complex networks: A survey, Physica A: Statistical Mechanics and its Applications **390**, 1150 (2011).
[16] T. Squartini, G. Caldarelli, G. Cimini, A. Gabrielli, and D. Garlaschelli, Reconstruction methods for networks: The case of economic and financial systems, Physics Reports **757**, 1 (2018).
[17] C. Diem, A. Borsos, T. Reisch, J. Kertész, and S. Thurner, Quantifying firm-level economic systemic risk from nation-wide supply networks, Scientific Reports **12**, 7719 (2022).
[18] J. Mori, Y. Kajikawa, H. Kashima, and I. Sakata, Machine learning approach for finding business partners and building reciprocal relationships, Expert Systems with Applications **39**, 10402 (2012).
[19] Y. Zuo, Y. Kajikawa, and J. Mori, Extraction of business relationships in supply networks using statistical learning theory, Heliyon **2**, e00123 (2016).
[20] H. Sasaki and I. Sakata, Prediction of business partners using an n-gram-based approach that combines a network model and linear model of a supply chain, in *2017 Portland International Conference on Management of Engineering and Technology (PICMET)* (2017) pp. 1–8.
[21] D. Lee and K. Kim, Business transaction recommendation for discovering potential business partners using deep learning, Expert Systems with Applications **201**, 117222 (2022).
[22] E. E. Kosasih and A. Brintrup, A machine learning approach for predicting hidden links in supply chain with graph neural networks, International Journal of Production Research **60**, 5380 (2022), https://doi.org/10.1080/00207543.2021.1956697.
[23] N. Minakawa, K. Izumi, H. Sakaji, and H. Sano, Transaction prediction by using graph neural network and textual industry information, in *New Frontiers in Artificial Intelligence*, edited by Y. Takama, K. Yada, K. Satoh, and S. Arai (Springer Nature Switzerland, Cham, 2023) pp. 251–266.
[24] L. Mungo, F. Lafond, P. Astudillo-Estévez, and J. Doyne Farmer, Reconstructing production networks using machine learning, Journal of Economic Dynamics and Control , 104607 (2023).
[25] W. Zhang, R. Y. Lau, Y. Xia, C. Li, and W. M. Li, Latent business networks mining: A probabilistic generative model, in *2012 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent*

Agent Technology, Vol. 1 (2012) pp. 558–562.

[26] B. A. Schaffer P., Kosasih E., Extracting supply chain knowledge graphs from natural language text using artificial intelligence, under review (2023).

[27] T. Reisch, G. Heiler, C. Diem, P. Klimek, and S. Thurner, Monitoring supply networks from mobile phone data for estimating the systemic risk of an economy, Scientific Reports **12**, 13347 (2022), number: 1, Publisher: Nature Publishing Group.

[28] S. Hooijmaaijers and G. Buiten, *A methodology for estimating the Dutch interfirm trade network, including a breakdown by commodity*, Tech. Rep. (Technical report, Statistics Netherlands, 2019).

[29] R. Hillman, S. Barnes, G. Wharf, and D. MacDonald, *A new firm-level model of corporate sector interactions and fragility: The Corporate Agent-Based (CAB) model*, Tech. Rep. 1675 (OECD, 2021).

[30] L. Mungo and J. Moran, Revealing production networks from firm growth dynamics (2023), arXiv:2302.09906 [q-fin.ST].

[31] W. L. Hamilton, *Graph Representation Learning*, Synthesis Lectures on Artificial Intelligence and Machine Learning (Springer Nature Switzerland, 2020) p. 141.

[32] Q. Le and T. Mikolov, Distributed representations of sentences and documents, in *Proceedings of the 31st International Conference on Machine Learning*, Proceedings of Machine Learning Research, Vol. 32, edited by E. P. Xing and T. Jebara (PMLR, Bejing, China, 2014) pp. 1188–1196.

[33] A. B. Bernard, E. Dhyne, G. Magerman, K. Manova, and A. Moxnes, The origins of firm heterogeneity: A production network approach, Journal of Political Economy **130**, 1765 (2022), https://doi.org/10.1086/719759.

[34] A. Rachkov, F. Pijpers, and D. Garlaschelli, *Potential Biases in Network Reconstruction Methods Not Maximizing Entropy*, Tech. Rep. (Discussion Papers, Central Bureau of Statistics, Netherlands, 2021).

[35] H. Watanabe, H. Takayasu, and M. Takayasu, Relations between allometric scalings and fluctuations in complex systems: The case of japanese firms, Physica A: Statistical Mechanics and its Applications **392**, 741 (2013).

[36] F. Parisi, T. Squartini, and D. Garlaschelli, A faster horse on a safer trail: generalized inference for the efficient reconstruction of weighted networks, New Journal of Physics **22**, 053053 (2020).

[37] A. Hazan, A maximum entropy network reconstruction of macroeconomic models, Physica A: Statistical Mechanics and its Applications **519**, 1 (2019).

[38] D. Garlaschelli and M. I. Loffredo, Fitness-dependent topological properties of the world trade web, Phys. Rev. Lett. **93**, 188701 (2004).

[39] X. Dong, D. Thanou, M. Rabbat, and P. Frossard, Learning graphs from data: A signal representation perspective, IEEE Signal Processing Magazine **36**, 44 (2019).

[40] L. Peel, T. P. Peixoto, and M. De Domenico, Statistical inference links data and theory in network science, Nature Communications (2022).

[41] J.-N. Barrot and J. Sauvagnat, Input Specificity and the Propagation of Idiosyncratic Shocks in Production Networks *, The Quarterly Journal of Economics **131**, 1543 (2016), https://academic.oup.com/qje/article-pdf/131/3/1543/30636709/qjw018.pdf.

[42] V. M. Carvalho and A. Tahbaz-Salehi, Production networks: A primer, Annual Review of Economics **11**, 635 (2019).

[43] V. M. Carvalho, M. Nirei, Y. U. Saito, and A. Tahbaz-Salehi, Supply Chain Disruptions: Evidence from the Great East Japan Earthquake, The Quarterly Journal of Economics **136**, 1255 (2021), https://academic.oup.com/qje/article-pdf/136/2/1255/36725306/qjaa044.pdf.

[44] S. Kumar, J. Ying, J. V. de Miranda Cardoso, and D. Palomar, Structured graph learning via laplacian spectral constraints, in *Advances in Neural Information Processing Systems*, Vol. 32, edited by H. Wallach, H. Larochelle, A. Beygelzimer, F. Alché-Buc, E. Fox, and R. Garnett (Curran Associates, Inc., 2019).

[45] K. Anand, I. van Lelyveld, Ádám Banai, S. Friedrich, R. Garratt, G. Hałaj, J. Fique, I. Hansen, S. M. Jaramillo, H. Lee, J. L. Molina-Borboa, S. Nobili, S. Rajan, D. Salakhova, T. C. Silva, L. Silvestri, and S. R. S. de Souza, The missing links: A global study on uncovering financial network structures from partial data, Journal of Financial Stability **35**, 107 (2018), network models, stress testing and other tools for financial stability monitoring and macroprudential policy design and implementation.

[46] G. Cimini, R. Mastrandrea, and T. Squartini, *Reconstructing Networks*, Elements in the Structure and Dynamics of Complex Networks (Cambridge University Press, 2021).

[47] A. Golan, G. Judge, and S. Robinson, Recovering information from incomplete or partial multisectoral economic data, The Review of Economics and Statistics **76**, 541 (1994).

[48] H. Inoue and Y. Todo, Firm-level propagation of shocks through supply-chain networks, Nature Sustainability **2**, 841 (2019).

[49] J. W. Welburn, A. Strong, F. E. Nekoul, J. Grana, K. Marcinek, O. A. Osoba, N. Koirala, and C. M. Setodji, *Systemic Risk in the Broad Economy: Interfirm Networks and Shocks in the U.S. Economy* (RAND Corporation, Santa Monica, CA, 2020).

[50] A. Bacilieri and P. Astudillo-Estevez, Reconstructing firm-level input-output networks from partial information (2023), arXiv:2304.00081 [econ.GN].

[51] E. T. Jaynes, Information theory and statistical mechanics, Phys. Rev. **106**, 620 (1957).

[52] S. Bartolucci, F. Caccioli, F. Caravelli, and P. Vivo, Ranking influential nodes in networks from aggregate local information, Physical Review Research **5**, 033123 (2023).

[53] S. Bartolucci, F. Caccioli, F. Caravelli, and P. Vivo, Inversion-free leontief inverse: statistical regularities in input-output analysis from partial information (2020).

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

[55] G. Occhini, E. Tranos, and L. Wolf, Measuring a country's digital industrial structure: commercial websites and weakly supervised classification to the rescue, SocArXiv (2023).

[56] E. Atalay, A. Hortacsu, J. Roberts, and C. Syverson, Network structure of production, Proceedings of the National Academy of Sciences **108**, 5199 (2011).

[57] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, ImageNet Large Scale Visual Recognition Challenge, International Journal of Computer Vision **115**, 211 (2015).

[58] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, Imagenet: A large-scale hierarchical image database, in *2009 IEEE Conference on Computer Vision and Pattern Recognition* (2009) pp. 248–255.

[59] J. Jordon, L. Szpruch, F. Houssiau, M. Bottarelli, G. Cherubin, C. Maple, S. N. Cohen, and A. Weller, Synthetic data – what, why and how? (2022), arXiv:2205.03257 [cs.LG].

[60] D. Ivanov and A. Dolgui, The shortage economy and its implications for supply chain and operations management, International Journal of Production Research **60**, 7141 (2022).

[61] A. Brintrup, E. E. Kosasih, B. L. MacCarthy, and G. Demirel, Digital supply chain surveillance: concepts, challenges, and frameworks, in *The Digital Supply Chain* (Elsevier, 2022) pp. 379–396.