Measuring the Input Rank in Global Supply Networks[*]

Armando Rungi[†] Loredana Fattorini[‡] Kenan Huremović[§]

Abstract

We introduce the Input Rank as a network measure of relevance of direct and indirect suppliers in Global Value Chains. We conceive an intermediate input to be more relevant for a downstream buyer if a negative shock in that input's productivity affects that buyer more. In particular, in our framework, the relevance of any input depends on: i) the network position of the supplier relative to the buyer, ii) the patterns of intermediate inputs intensities connecting the buyer and the supplier, and iii) the competitive pressures along supply chains. After we compute the Input Rank from both the U.S. and the world Input-Output tables, we provide useful insights into the crucial role of services inputs as well as on the relatively higher relevance of domestic suppliers, and suppliers coming from regionally integrated partners. Finally, we test that the Input Rank is a good predictor of vertical integration choices made by 20,489 U.S. parent companies controlling 154,836 subsidiaries worldwide.

Keywords: global value chains; Input Output; production networks; vertical integration; multinational enterprises

JEL Codes: F23; L23; D23; C63; C67

Data availability. Input-output data are available from the following publicly available resources: US Census Bureau, World Input Output Database. Measures of Upstreamness/Downstreamness are kindly made available from authors' original publications. Firm-level financial accounts for applications on vertical integrations are available from Orbis by Bureau Van Dijk. Restrictions apply only to the availability of firm-level data as they were used under a commercial license.

[*]We are grateful to participants at the conference on 'Global Value Chains - Recent Advances, Developments, Impacts' in Halle (Germany), the American Economic Association Annual Meeting 2018 in Philadelphia (USA), the Sardinian Empirical Trade Conference 2018 in Cagliari (Italy), the European Trade Study Group 2016 in Helsinki (Finland), the Italian Trade Study Group 2016 in Lucca (Italy), the seminars at the University of Alicante and the Vienna Institute for International Economic Studies. We want to thank Pol Antràs, Matteo Bugamelli, Italo Colantone, Davide Del Prete, Raja Kali, Kalina Manova, Malwina Mejer, Gianmarco Ottaviano, Frank Pisch, Massimo Riccaboni, Beata Smarzynska, and Lucia Tajoli for valuable comments at different stages of the project. We owe special thanks to Zhen Zhu, who participated to discussions at the onset of our project.

[†]corresponding author. Mail to: armando.rungi@imtlucca.it. Laboratory for the Analysis of Complex Economic Systems, IMT School for Advanced Studies, piazza San Francesco 19 - 55100 Lucca, Italy.

[‡]Mail to: loredana.fattorini@alumni.imtlucca.it. Laboratory for the Analysis of Complex Economic Systems, IMT School for Advanced Studies, piazza San Francesco 19 - 55100 Lucca, Italy.

[§]Mail to: kenan.huremovic@imtlucca.it. Laboratory for the Analysis of Complex Economic Systems, IMT School for Advanced Studies, piazza San Francesco 19 - 55100 Lucca, Italy.

1 Introduction

Modern economies organize as webs of specialized producers within and across national borders. Production networks have emerged as a major organization of multinational production since mid '80s, and they are now under considerable strain after first a pandemic crisis and then a conflict in Europe are together reshaping global economic relations. In a production network, each company plunges into complex production networks whose configuration is often recursive because the contribution of some intermediate inputs is needed at different stages of completion. Take the case of Audi, a German automobile manufacturer with productive plants around the world. According to the Automotive Industry Portal (2020), it relies upon the deliveries of $1,535$ direct suppliers located in 45 countries. Among its suppliers, we find Brembo S.p.A., a well-known producer of brakes in Italy, and Bosch Corporation, a subsidiary of the Bosch Group producing valves and other components in Japan. Brembo S.p.A., in turn, reports 33 direct suppliers of parts and components, among which we detect Garrett Motion, a Swiss producer of stamping and aluminum casting. Yet, Garrett Motion is also a direct supplier of Audi, and it delivers components to Bosch Corporation in Japan. After first explorations, we encounter Garrett Motion at least three times in Audi's global supply network, once as a direct and twice as an indirect supplier.

Similar interlocking patterns are systematic in the global industry, not only in the automotive sector. Consequently, a global fragmentation of production entails a combination of both spider-like and snake-like configurations of supply networks (Baldwin and Venables, 2013), both within and across national borders. Yet, the organization of Global Value Chains (GVCs) is mainly modeled assuming a technology of productive tasks over linear sequences, i.e., the 'chain', oriented on upstream-downstream segments (Costinot et al., 2012; Antràs and Chor, 2013; Fally and Hillberry, 2018; Alfaro et al., 2019; Antràs and de Gortari, 2020).

The linear production sequence envisioned in these models allows tractability and enables authors to uncover the primitives of an organization based on mutual contractual interdependence between buyers and suppliers. Final producers meet consumers' demands and realize a surplus distributed with direct and indirect suppliers. Therefore, a blend of contractual environment and market forces shapes the final distribution of the surplus. Inspired by these theoretical models, scholars proposed GVC position metrics, i.e., the Upstreamness and the Downstreamness, which capture a notion of distance of suppliers taking into account the structure of the Input-Output network (Alfaro et al., 2019; Wang et al., 2017; Antràs and Chor, 2018; Antràs and Chor, 2013; Antras et al., 2012).

In this paper, we introduce the Input Rank as a bilateral measure of the relevance of any direct or indirect input in a supply network of a given firm, when technology blends with market forces. We start by modeling the problem of a producer, embedded in a network of buyer-supplier relationships, who plans the delivery of her output taking into account the complex web of input-output relations. In this, we build upon recent contributions on production networks (Acemoglu et al., 2012; Carvalho, 2014; Grassi, 2017; Baqaee, 2018; Baqaee and Farhi, 2020; Grassi and Sauvagnat, 2019; Carvalho and Tahbaz-Salehi, 2019). Crucially, in our framework, the most important (direct or indirect) inputs of a firm are the ones that have the potential to affect the marginal costs of that firm the most.

Notably, apart from a full account of the network structure of the supply side of the economy, we allow for a rich heterogeneity concerning how much firms rely on intermediate inputs, as well as competitive forces within each sector. At each stage of production, a lower intensity in the usage of intermediate inputs (i.e., a higher intensity in labor services) buffers the transmission of a shock from upstream markets. Thus, an upstream supplier will rank relatively higher for a given downstream buyer when there are more supply chains (paths) connecting them, and these supply chains are comprised of firms that rely more on intermediate inputs in their production process. At the same time, a higher markup (i.e., a lower competitive pressure) on upstream markets will make downstream buyers more sensitive to input-specific productivity shocks.

In the second part of our paper, we bring our model to the data. We calibrate it on the U.S. Input-Output tables, sourced from the Bureau of Economic Analysis (BEA, 2002), and on the world Input-Output tables, sourced from WIOD (Timmer et al., 2015). Therefore, we provide interesting insights into the network dimensions of global production. In particular, we uncover the crucial role of services inputs, which are most central in the configuration of production networks. Then, when we look at geography, we show a pecking order in buyers' sensitivity to input shocks: first domestic suppliers and then suppliers from regionally integrated neighbors (e.g., intra-EU, intra-NAFTA, intra-ASEAN input trade) have the potential to hit harder on downstream buyers.

Finally, we contrast the Input Rank with the Downstreamness (Antràs and Chor, 2013), Upstreamness (Alfaro et al., 2019), and measures based on the classic Leontief inverse. First of all, we note that these measures are introduced based on accounting identities provided by I-O tables (see Antràs and Chor (2018) for a review) while our measure is a direct result of a fully-fledged general equilibrium model which explicitly incorporates the network structure of the production side of the economy. Second and more importantly, we argue that, contrary to the measures mentioned above, the Input Rank accounts for competition structure within sectors. We test the correlation of the Input Rank with choices of vertical integration made by 20,489 U.S. parent companies controlling 154,836 subsidiaries worldwide. We find that a higher Input Rank is positively associated with higher odds that a (direct or indirect) input is vertically integrated. Additionally, we find that parent companies preferably integrate suppliers that are not too distant on the network because they report a relatively lower Upstreamness.

The rest of the paper is organized as follows. The next Section 2 discusses related literature. Section 3 introduces a compact theory for the Input Rank. In Section 4, we compute the Input Rank on both the U.S. and world Input-Output tables to describe preliminary evidence. In Section 5, we test the role of the Input Rank in firm-level choices of vertical integration. Concluding remarks are offered in Section 6.

2 Related literature

A fruitful strand of research emerged in the last decade to study how network dimensions are essential in the organization of production, and how much they contribute to explaining aggregate fluctuations as a response to microeconomic shocks (Carvalho and Tahbaz-Salehi, 2019; Baqaee, 2018; Acemoglu et al., 2012; Carvalho, 2014), possibly providing the rationale for empirical investigations concerning firm-level heterogeneity (Mundt, 2021; Baqaee and Farhi, 2020). International economics potentially offers a unique environment to study the network dimensions of cross-country trade, investment, and mobility of workers (Chaney, 2016, 2014). In particular, network metrics have been useful to show how the entire trading system has become intensely interconnected over time (De Benedictis and Tajoli, 2011), and trade policies may have an impact that depends on network dimensions (Orefice and Rocha, 2014; Conconi et al., 2018; Giammetti et al., 2020). Therefore, scholars focused on indicators that could catch the participation of countries and industries into Global Value Chains (GVCs) through trade in intermediate inputs, mainly based on available input-output tables (Koopman et al., 2014; Borin and Mancini, 2021; Imbs and Pauwels, 2022). Although significant progress has been made to understand global interdependence on production, we believe the literature on trade and production networks is still in its infancy, and many questions remain unanswered (Bernard and Moxnes, 2018; Baldwin et al., 2022).

In this contribution, we focus on bridging network dimensions with the theory and empirics of Global Value Chains (GVCs), as the latter have been mainly studied as supposedly linear technological sequences (Del Prete and Rungi, 2020; Alfaro et al., 2019; Fally and Hillberry, 2018; Rungi and Del Prete, 2018; Del Prete and Rungi, 2017; Antràs and Chor, 2018; Wang et al., 2017; Miller and Temurshoev, 2017; Antràs and Chor, 2013; Antras et al., 2012), although the existence of spider-like vs. snake-like configurations had been acknowledged by Baldwin and Venables (2013). The primary purpose of previous literature is to explain how contractual interdependence can shape the organization of GVCs in terms of vertical integration *vis á vis* outsourcing strategies[1]. Berlingieri et al. (2018) highlights how the technological importance of inputs on GVCs can be shaped by a solution to both *ex post* contracting problems (transaction costs) and *ex ante* lack of incentives and under-investment problems (property rights forces). An implicit step in modeling a network dimension in GVCs has been made by Antràs and de Gortari (2020), who assume that a linear technology interacts with the geographic centrality of alternative locations. de Gortari (2019) and Caliendo and Parro (2014) also exploit implicit information on the network configurations of GVCs to build numerical trade policy counterfactuals based on the transmission of value from inputs to outputs across national borders.

Our model is closely related to literature on propagation of microeconomic shocks through the production network (Grassi, 2017; Carvalho and Tahbaz-Salehi, 2019; Huremović et al., 2020). While we are not directly interested in how shocks propagate, we frame the problem of input sourcing from a similar perspective. First, a downstream producer observes the topology of her supply structure made of direct and indirect inputs; then, it can assess the impact of a potential productivity shock of any intermediate input on its production process, based on how the shock propagates through the network. An input (direct or indirect) whose productivity change has the largest effect on the marginal cost of the downstream producer is then more likely to be integrated by that producer.

Therefore, the Input Rank that we propose is a recursive measure of a node's relative position in the network, and as such is similar to classic eigenvector-type centrality measures like the PageRank (Brin and Page, 1998), and the Katz (1953) centrality. See Bloch et al. (2019) for an extensive discussion and axiomatization of centrality measures used in network analysis. Different variants of such centrality measures have been applied in many different domains, from biology and genetics to financial debts, bibliometrics, and road engineering (Gleich, 2015; Newman, 2018). In either case, a node is more important in a network if it is connected to other important nodes. Similarly, in our production framework, a (direct or indirect) input is more relevant if it delivers (receives) to (from) other highly requested inputs.

Finally, we believe it is important to understand how we relate to recent literature in

[1] For a review on the importance of GVCs in shaping international trade flows, see Antras and Chor (Forthcoming), where also the benefits of a microeconomic approach are discussed.

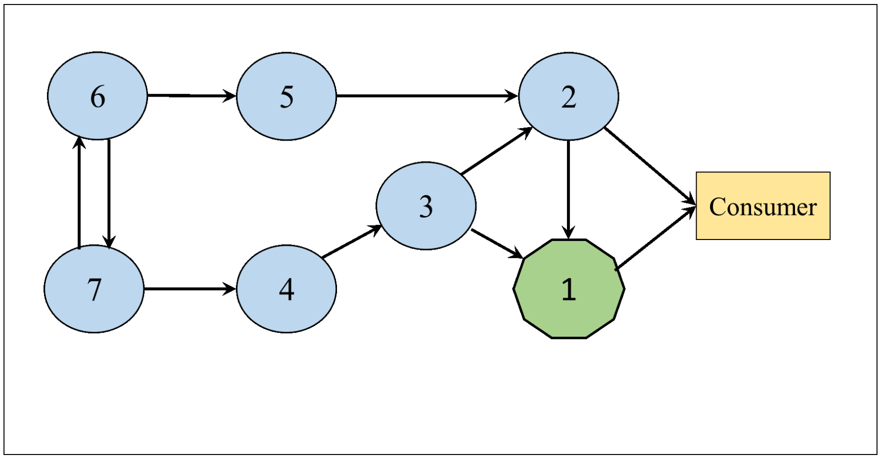
which production networks are studied by looking at firm-to-firm transactions (Bernard et al., 2022; Dhyne et al., 2022). In principle, one could calculate suppliers' Input Rank by looking at firm-to-firm transactions, but the Input Rank that we propose is an industry-level measure. In fact, our purpose is made explicit through our theoretical framework, where we frame the problem of the Input Rank as catching the relevance of an entire input market, not of single suppliers belonging to it. In our framework we want to consider not only a supplier's network position, but also its patterns of intermediate inputs' intensities and the competitive pressures encountered along supply chains on the way to the downstream final producer.

3 A simple model for ranking inputs

In this section, we lay out the theoretical foundations for ranking inputs in supply networks. Assessing the importance of a supplier is not a straightforward task when production processes are fragmented. To illustrate this point, we start with a stylized example depicted in Figure 1, where nodes indicate sectors or representative firms from those sectors (hereafter simply referred to as firms or nodes), whereas directed links indicate deliveries of goods or services.

We focus on the supply chain of firm 1. Consider a scenario in which firm 4 is affected by a distortion or a shock due to which it produces an output of a lower quality/productivity for its customers (firm 3). This affects firm 3's production process since firm 3 uses the output of firm 4 directly in production, and moreover the aforementioned shock is (partially) passed down from firm 3 directly to firm 2, and both directly and indirectly (through firm 2) to firm 1. Thus, the shock hitting firm 4 is transmitted *downstream* along the network of producers, reaching and affecting firm 1 through different routes. Intuitively, we expect the indirect effect of the shock hitting firm 4 on firm 1 to depend on how much firm 1 relies, directly and indirectly, on the output produced by firm 4. This, in turn, will depend on the relative position of firms 4 and 1 in the network of producers, in particular on the number of weighted paths of *any* length stemming from 4 and reaching firm 1, but also, as will be apparent from the model, the competition structure in sectors along the supply chains.

Figure 1: A fictional supply network



In this section, we present a simple general equilibrium model which formalizes the intuition from our simple example. For any given firm operating in sector k in the economy, the model provides a measure of importance of a firm from sector h as a supplier (direct and indirect) of that firm, taking *fully* into account the network structure of the production side of the economy. We call this measure the Input Rank of h relative to k. Our theoretical framework is, in many respects, standard in the literature of production networks, and thus we present it in a quite compact manner. In particular we build directly on Grassi (2017) and Baqaee (2018). Formal proofs of the claims are relegated to the Appendix.

3.1 Consumers

There are two types of agents in the economy: firms and the representative consumer. We denote the set of firms in the economy with N. Firms are grouped in M sectors. Each firm belongs to exactly one sector, and it produces a single differentiated variety of a sector-specific good.

The representative consumer owns all firms in the economy and supplies one unit of labor inelastically. The preferences of the consumer over M goods are defined with the following Cobb-Douglas utility function:

$$U(c_1, c_2, ..., c_M) = C = \theta \prod_{k=1}^{M} c_k^{\gamma_k} \tag{1}$$

where c_k is the consumption of good k, $\sum_k \gamma_k = 1$, and $\theta = \prod_{k=1}^{M} \gamma_k^{-\gamma_k}$ is a normalization constant to simplify computations. The composite consumption good k is defined with:

$$c_k = \left(\sum_{i=1}^{M_k} c(k,i)^{\frac{\epsilon_k - 1}{\epsilon_k}} \right)^{\frac{\epsilon_k}{\epsilon_k - 1}}, \tag{2}$$

where $c(k,i)$ denotes the consumption of variety i of good k, $\epsilon_k > 1$ is the elasticity of substitution across varieties of good k, and M_k denotes the number of firms (varieties) in sector k. The consumer maximizes her utility subject to the following budget constraint:

$$\sum_{k=1}^{M} \sum_{i=1}^{M_k} p(k,i)c(k,i) = w + \sum_{k=1}^{M} \sum_{i=1}^{M_k} \pi(k,i), \tag{3}$$

where $p(k,i)$ is the price of variety i in sector k, $\pi(k,i)$ is the profit of firm $i \in k$, and w is the worker's wage bill.

3.2 Firms

Firms in sector k are symmetric and use the same constant returns to scale technology that combines labor ℓ and intermediate inputs.[2] Each firm in k produces an imperfectly substitutable variety i of good k. Let us denote with $y(k,i)$ the output of firm i from sector k, with $\ell(k,i)$ its labor input, and with $x(k,i,h,j)$ the amount of variety j of good h used in the production of variety i of good k. Thus, the profit of firm i in sector k is defined with:

$$\pi(k,i) = p(k,i)y(k,i) - \sum_{h=1}^{M} \sum_{j=1}^{M_h} p(h,j)x(k,i,h,j) - w\ell(k,i) \tag{4}$$

The production possibilities of a typical firm i from sector k are defined with the following nested production function:

$$y(k,i) = \zeta_k \ell(k,i)^{\beta_k} \left(\prod_{h=1}^{M} \left[\sum_{j=1}^{M_h} \left[\tau_h^{-1} x(k,i,h,j) \right]^{\frac{\varepsilon_h - 1}{\varepsilon_h}} \right]^{\frac{\varepsilon_h}{\varepsilon_h - 1} g_{hk}} \right)^{\alpha_k}. \tag{5}$$

where $\alpha_k + \beta_k = 1$ and $\sum_h g_{hk} = 1$. We normalize $\zeta_k = \beta_k^{-\beta_k} \prod_{h=1}^{N} g_{hk}^{-\alpha_k g_{hk}}$ to simplify the

[2]For simplicity labor is the only factor of production in the model, i.e., it is not delivered by other firms. Clearly, the model can be extended to include more production factors (i.e. capital) in a symmetric way. Including additional factors of production will not affect the resulting Input Rank measure. As it will be apparent, in the empirical application we take into account that other inputs, besides labor and intermediate inputs, figure in the cost structure of a firm.

computations.

Matrix $\mathbf{G} = (g_{hk})_{h,k=1}^{M}$ defines the sector level (technological) production network of the economy, in which the set of nodes is the set of sectors in the economy $\{1, 2, ..., M\}$, and two sectors h and k are connected with directed link $h \to k$ if firms in k use good h as an input in production. Entry g_{hk} of the adjacency matrix $\mathbf{G}$ represents the weight of link $h \to k$, where $g_{hk} = 0$ indicates that firms from sector k do not use input h as a (direct) input in production. Given our functional assumptions, $g_{hk} \geq 0$ will correspond to the cost share associated to intermediate inputs from sector h. Therefore, we can directly construct matrix $\mathbf{G}$ using the data from the Input-Output tables. We discuss this in more detail in Section 4.

Parameter $\tau_h \geq 1$ in (5) captures, in a parsimonious way, the productivity of each variety of good h when used as an intermediate input. Higher values of τ_h imply a lower productivity of intermediate input h when used in production. In this paper we are interested in the effects of marginal changes in τ_h on the equilibrium outcomes. We sometimes refer to these changes as shocks, while we refer to τ_h as a productivity distortion associated to (inputs produced in) sector h.[3] Intuitively, we think of distortion τ_h as the extent to which firms from sector h fail to deliver compatible or productive input to their customers on time. Anticipating our later discussion of the vertical integration, we assume that a firm that uses input h either directly or indirectly in its production process can, at least partially, mitigate this distortion by integrating a producer of good h. This is one of the reasons we are interested in this particular type of distortions.

It is useful to define the composite intermediate input h used in production of variety $i \in k$ (denoted with $x(k, i, h)$) as an aggregate of varieties of input h:

$$x(k,i,h) = \left[\sum_{j=1}^{M_h} \left[x(k,i,h,j) \right]^{\frac{\varepsilon_h - 1}{\varepsilon_h}} \right]^{\frac{\varepsilon_h}{\varepsilon_h - 1}}, \tag{6}$$

in which case we can write (5) as:

$$y(k,i) = \zeta_k \ell(k,i)^{\beta_k} \left(\prod_{h=1}^{M} (\tau_h^{-1} x(k,i,h))^{g_{hk}} \right)^{\alpha_k}. \tag{5A}$$

[3]For simplicity we assume that distortions and shocks are sector specific. One could of course consider firm-specific distortions $\tau(h, j)$ for $j \in h$, or even pair-specific distortions $\tau(h, j, k)$. We abstract from this type of heterogeneity in the paper.

3.3 Equilibrium

We assume that firms in the same sector operate in a monopolistic competition environment, and thus set their price to a constant markup over marginal costs. Following Atkeson and Burstein (2008), we assume that firms set their prices taking as given the other sectors' prices and quantities, the wage, and the aggregate prices and quantities. In Definition 3.1 we define the equilibrium concept we are considering.

Definition 3.1 (Equilibrium)**.** A market equilibrium is a collection of prices $p(k,i)$, wage w, input demands $x(k,i,h,j)$, outputs $y(k,i)$, consumption $c(k,i)$ and labor demands $\ell(k,i)$ such that:

(i) Each firm i minimizes production costs and applies its mark-up μ_i to set its price.

(ii) The representative consumer chooses consumption to maximize utility.

(iii) Markets for each good and labor clear.

The existence and the uniqueness of the equilibrium follows from standard arguments, see for instance (Grassi, 2017; Baqaee, 2018).

3.4 Ranking inputs

Informally, we say that (direct or indirect) input h is more important than input r for firm $i \in k$, or that $i \in k$ relies more on h compared to r if a change τ_h affects firm $i \in k$ more than the same change in τ_r. The first question we ask is if and how the importance of good h for firms in k depends on the underlying structure of sector-level production network. Proposition 1 provides an answer.

Proposition 1. Let $\lambda(k,i)$ denote the marginal cost of production of firm i in sector k. The following holds in the equilibrium:

$$\frac{\partial \log \lambda(k,i)}{\partial \log \tau_h} = \mathbf{e}_k' \left[\mathbf{I} - \mathbf{A}\mathbf{G}'\right]^{-1} \mathbf{e}_h = \mathbf{e}_h' \left[\mathbf{I} - \mathbf{G}\mathbf{A}\right]^{-1} \mathbf{e}_k \tag{7}$$

where $\mathbf{e}_k$ is the k-th unit vector, and $\mathbf{A}$ is the diagonal matrix that collects information about sector specific intermediate inputs' cost shares, $\{\alpha_k\}_{k=1}^{M}$, and $\mathbf{I}$ is the identity matrix.

In our simple framework, a negative shock (distortion) implies an increase in parameter, τ_h. Since the shock cascades through all production paths running towards the downstream buyers, its impact is a function of both the structure of the production network (captured by $\mathbf{G}$), and of the intensities in intermediate inputs, $\{\alpha_k\}_{k=1}^{M}$. Motivated by Proposition 1 we introduce the Input Rank as a measure of the importance of inputs.

Definition 3.2 (Input Rank)**.** The Input Rank of any upstream supplier of input h relative to the downstream producer of an output k is:

$$\upsilon_{hk}(\mathbf{GA}) = \mathbf{e}_h' \left[\mathbf{I} - \mathbf{GA}\right]^{-1} \mathbf{e}_k \tag{8}$$

In other words, the bilateral Input Rank, υ_{hk}, is the (h,k)-th element of the matrix $[\mathbf{I} - \mathbf{GA}]^{-1}$. From the perspective of a producer, vector $\boldsymbol{\upsilon}_k = \{\upsilon_{hk}\}_{h=1}^{M}$ encodes information on the relevance of any direct or indirect input h based on its position in the supply structure of a producer in sector k. Note that the inverse in (8) exists since $\mathbf{G}$ is a column-stochastic matrix with spectral radius equal to 1, and $\alpha_k \leq 1$ for each industry k.

In Section 5, we empirically investigate determinants of a firm's decision to vertically integrate suppliers. Assuming that the integration of an input allows a firm to mitigate[4], at least partly, distortion τ_h associated to good h, we expect that a cost-minimizing firm $i \in k$ is more likely to integrate those inputs h for which a decrease in distortion τ_h will have the largest positive effect on that firm's marginal cost of production $\lambda(k,i)$. The following proposition is useful as it rationalizes empirical results presented in Table 10.

Proposition 2. The marginal effect of a change in τ_h on the marginal cost of firm $i \in k$ is higher for higher values of Input Rank υ_{hk}, and for lower values of elasticity of substitution ϵ_h.

A higher Input Rank, υ_{hk}, implies that suppliers from sector h are relatively more important for firms operating in sector k (in a sense of Proposition 1), and thus it is not surprising that the effect of a decrease of distortion τ_h associated to an upstream sector has a larger effect on the costs of firm $i \in k$ the higher the input rank of that sector relative to k. Moreover, a lower degree of substitutability among varieties in sector h amplifies the effect of an input productivity distortion.

[4]Please note that our purpose is not to characterize a full incentive structure for the relations between buyers and suppliers. Our model is agnostic with respect to the impact of property rights forces, which have the potential to reduce the investment incentives by vertically integrated suppliers, hence reduce productivity.

In an Appendix Section B, we further introduce a variant of the Input Rank where the downstream buyer imperfectly observes her supply network. We show that, in such a case, there is a potential downplay of more distant nodes, and the downstream buyer overestimates the role of more proximate industries. We believe that cases of imperfect information of the supply network can arise when the contractual environment is imperfect and/or transactions are relation specific, for example in the sense introduced by Rauch (1999), Nunn (2007), Nunn and Trefler (2013), and Nunn and Trefler (2014).

4 The Input Rank from Input-Output tables

In this Section, we describe how the Input Rank can be calculated using publicly available Input-Output (I-O) tables. Then, we contrast the Input Rank with well-known GVC positioning metrics used in previous literature i.e., the Downstreamness, the Upstreamness, and the more classical Leontief inverse. Finally, we provide some insights on inputs' centrality in the organization of global production. For consistency with previous studies, we compute the Input Rank on both the U.S. and the world Input-Output tables.

U.S. I-O 2002 tables, compiled by the Bureau of Economic Analysis (BEA), sketch a reasonably fine-grained supply network established among 6-digit industries. The same tables have been extensively used to study production networks (Carvalho, 2014), vertical integration choices (Acemoglu et al., 2009; Alfaro et al., 2016), and to compute Upstreamness/Downstreamness metrics (Alfaro et al., 2019; Antràs and Chor, 2013).

On the other hand, multicountry Input-Output tables like the ones we sourced from WIOD (Timmer et al., 2015) have been extensively used in settings where the geographical dimension of GVCs is important (among others, see World Bank (2020); Fajgelbaum and Khandelwal (2016); Adao et al. (2017); Caliendo and Parro (2014); Costinot and Rodríguez-Clare (2014), and where the main issue has been to track value added generation by country and industry while avoiding a classical problem of multiple accounting from standard gross trade data. For a useful review, see also Johnson (2018). Usefully, WIOD data report information on exchanges among 43 countries and 56 ISIC rev. 4 2-digit industries in the period 2000 - 2014. A model for the rest of the world balances world trade. Additional socio-economic accounts contain the information we need to calculate labor and intermediate shares. For further details, see (Timmer et al., 2015).

4.1 Reconciling with Input-Output tables

To show how the Input Rank can be calculated using data from I-O tables, let us first derive the demand for intermediate inputs at the sector level within our model. Aggregating individual demands $x(h, i, k)$ from Eq. (6), we show that (see Lemma 3 in Appendix A)

$$x(k,h) = \frac{\alpha_k g_{hk} p_k y_k}{\mu_k p_h} = \frac{\epsilon_k - 1}{\epsilon_k} \frac{\alpha_k g_{hk} p_k y_k}{p_h}.$$

Let us define $Z_{hk} \equiv p_h x(k, h)$, which is the value of input deliveries from sector h to sector k and is readily available in the I-O data. From (A4) it directly follows that

$$\frac{p_h x(k,h)}{\sum_\ell p_\ell x(k,\ell)} = \frac{Z_{hk}}{\sum_\ell Z_{\ell k}} = g_{hk}. \tag{9}$$

Therefore, elements g_{hk} of matrix $\mathbf{G}$ can be directly calculated from the I-O data as the ratio of value of input h used directly by sector k, and k's total cost of intermediate inputs. Furthermore, thanks to the Cobb-Douglas assumption, elements $\{\alpha_k\}_{k=1}^{M}$ of the diagonal matrix $\mathbf{A}$ can be calculated as cost shares of intermediate inputs at the sector level. To be more precise, we calculate α_k in the following way:

$$\alpha_k = \frac{\text{Cost of intermediate inputs of sector } k}{\text{Total cost of all inputs of sector } k}.$$[5]

The Input Ranks are then calculated as elements of matrix $(\mathbf{I} - \mathbf{GA})^{-1}$.

4.2 Relation with other Input-Output metrics

The Input-Rank, as defined in (8), bears some similarity to other commonly used measures based on I-O network including the Upstreamness, the Downstreamness, and the total input requirements. In this section we compare the Input Rank to these measures, and argue that the first captures additional information ignored by other measures.

Before doing so, let us introduce some additional notation. Let $Y_k \equiv p_k y_k$ denote the value of gross output of sector k, and let d_{hk} denote the direct requirement coefficient obtained from I-O tables and equal to the ratio of value of sales from sector h to sector k (Z_{hk}), and the value of the gross output of sector k (Y_k). The matrix $(\mathbf{I} - \mathbf{D})^{-1}$ is known as the *Leontief inverse* or *the total requirements matrix*, where $\mathbf{D}$ is the matrix of direct requirements. Antràs and Chor (2018) discusses a number of network statistics based on the Leontief inverse and its square

$(\mathbf{I}-\mathbf{D})^{-2}$, including the Downstreamness and the Upstreamness measures, introduced in Antras et al. (2012); Antràs and Chor (2013); Miller and Temurshoev (2017).

There are at least two important differences between the Input Rank and these measures. First, even though the Input Rank matrix $(\mathbf{I}-\mathbf{GA})^{-1}$ is reminiscent of the Leontief inverse $(\mathbf{I}-\mathbf{D})^{-1}$, matrices $\mathbf{GA}$ and $\mathbf{D}$ are, in general, different. To see this, note that from (A4) it directly follows that:

$$d_{hk} = \frac{Z_{hk}}{Y_k} = \frac{p_h x(k,h)}{p_k y_k} = \frac{\epsilon_k - 1}{\epsilon_k}\alpha_k g_{hk}, \tag{10}$$

and hence $\mathbf{GA} = \mathbf{DM}$, where $\mathbf{M}$ is a diagonal matrix of markups, $\mu_k = \frac{\epsilon_k}{\epsilon_k - 1}$. Thus, one can see the matrix $\mathbf{GA}$ as a *markup-adjusted* matrix of directed requirements $\mathbf{D}$. The two measures coincide only in the special case when there is a perfect competition in *each* sector.[6] Importantly, we do not need to estimate the markups to calculate the Input Rank, as the latter is a function of matrix $\mathbf{GA}$ only, and this matrix can be directly constructed from the I-O data as explained before. Since elements of the diagonal matrix $\mathbf{M}$ are greater than one, the Input Rank matrix implies *stronger* interdependence between sectors than the Leontief inverse.

Second, the above mentioned measures of Downstreamness and Upstreamness are introduced on the basis of accounting identities implied by the I-O tables (see Antràs and Chor (2018) for a review of these measures). The Input Rank measure is a direct result of a fairly standard general equilibrium model that explicitly takes into account the network nature of the production side of the economy.

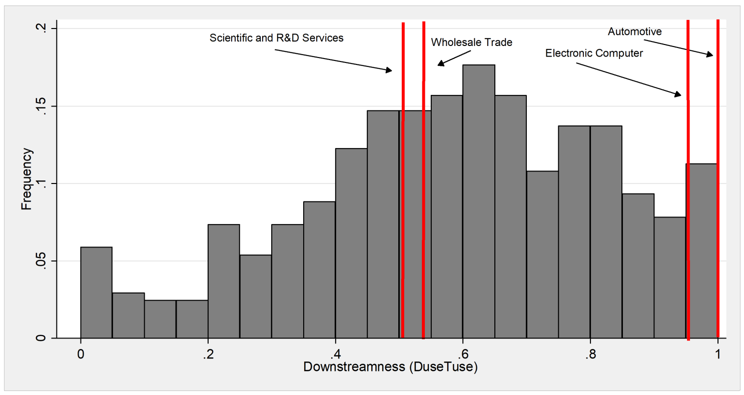
Finally, the Upstreamness and the Downstreamness are built with a purpose to capture the notion of (weighted) distance from final consumption or primary factors of production[7]. Notably, once we compare the positions of selected industries in the network reproduced in Figure 3 with their positions on the Downstreamness segment in Figure 2, we curiously find that most central inputs in the U.S economy, i.e., R & D and Wholesale Trade are to be found in the middle. This is in contrast with the stylized value chain we may have in mind, where a representative business line starts with R & D services and ends with distribution services. The latter evidence is instead consistent with the existence of a production network where both R & D services and wholesale traders have a central position, i.e., they are present

[6] Although we assume monopolistic competition within each sector, our model can be easily extended to different types of within-sector imperfect competition, in which case the exact expression for markups changes. See Baqaee (2018); Baqaee and Farhi (2020) for more details.

[7] For an updated introduction to Upstreamness and Downstreamness metrics, see Antràs and Chor (2018)

as intermediate inputs on different production paths[8]. In this respect, we argue, the Input Rank is of help in complementing metrics like Upstreamness and Downstreamness, whose scope is limited to capturing the relative distance of an input from final demand or factors of production, respectively.

Figure 2: *Downstreamness* from the U.S. BEA 2002 I-O tables



Downstreamness (DuseTuse) is sourced from Antràs and Chor (2013). Frequency indicates how many industries out of a total of 425 from U.S. Input-Output tables occupy that specific position. Selected industries: Scientific Research and Development Services (code 541700, value 0.504); Wholesale Trade (code 541800, value 0.666); Electronic Computer Manufacturing (code 334111, value 0.959); Automobile Manufacturing (code 336111, value 0.999).

To illustrate the empirical difference between the Input Rank and the other measures discussed above, we report rank correlations in Table 1. Eventually, we also decide to include in the comparison a more mundane indicator of an input's capital intensity, as the latter can represent yet another indicator of an input's technological relevance.

Usefully, both the Spearman's and the Kendall (1948)'s correlations assess how well the relationship among I-O metrics can be described by a monotonic function, i.e., when inputs have a similar ordinal position across metrics. Interestingly, we observe that there is a low and weakly significant rank correlation between the Input Rank and the Upstreamness, albeit with a negative sign given the upstream orientation of the latter (-.40 and -.27, respectively). In fact, the rank correlations of Input Rank with Upstreamness are lower than the ones with the direct coefficients, at the bottom of the table (.55 and .30, respectively). In this respect, the Upstreamness has a relatively higher association with original information from I-O

[8]Please note how our argument is confirmed by the most recent Upstreamness by Alfaro et al. (2019), which is a bilateral position metrics that accounts for the heterogeneity of input positions oriented towards different root outputs. Also there, R & D services are still on average located in the middle of the output-specific technological sequences, i.e., the average Upstreamness value is 3.044 for an indicator that ranges approximately from 1 to 8.9.

tables (-.68 and -.39, respectively) than with the Input Rank.

When we compare the Input Rank with a classic Leontief inverse, we observe how Spearman correlation stops at .59, and Kendall's τ at .30. At the end of the day, Upstreamness reveals a relatively higher rank correlation with a Leontief inverse than the Input Rank. We can conclude that all indicators in Table 1 convey different pieces of information on I-O linkages, as introduced in following paragraphs.

Finally, both the Input Rank and the Upstreamness show relatively low rank correlations with an input's capital intensity, as retrieved on the latter rows of Table 1. As for the rank correlation with the Input Rank, this is consistent also with the evidence we will find in the following paragraphs, according to which labor-intensive services are indeed the inputs that are technologically most relevant in modern economies.

Table 1: Rank correlations

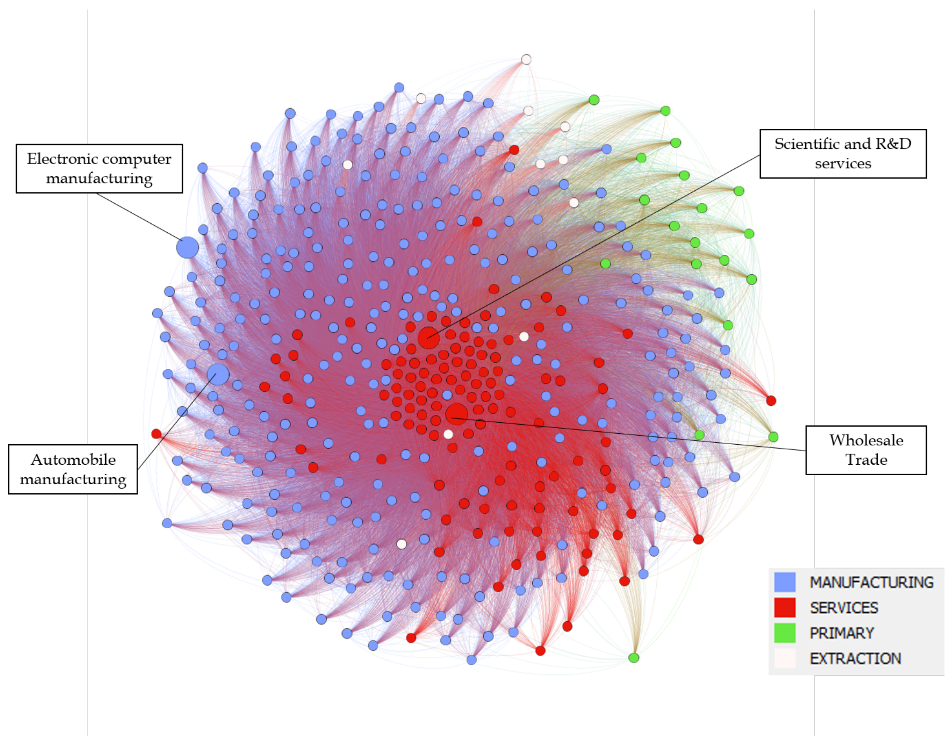
	Input Rank (υ_{hk})	Upstreamness ($upstr_{hk}$)
Upstreamness ($upstr_{hk}$)		
Spearman's ρ	-.3962*	
Kendall's τ	-.2672*	
Leontief Inverse (L_{hk})		
Spearman's ρ	.5904*	-.6845*
Kendall's τ	.4011*	-.4595*
I-O coefficient (d_{hk})		
Spearman's ρ	.5468*	-.6789*
Kendall's τ	.3040*	-.3875*
Capital Intensity (kl_{hk})		
Spearman's ρ	.0140*	-.0251*
Kendall's τ	.0121*	-.0189*

Note: Spearman's and Kendall's rank correlations are computed to underline the different informative value of the Input Rank, The Upstreamness (Alfaro et al., 2019), the Leontief inverse, the direct requirement coefficients, and the inputs' capital intensities. All measures are reported for $88,595$ input-output pairs. * stands for p-value $< .1$.

4.3 Evidence from the U.S. Input Output tables

The U.S. tables are especially useful to grasp our intuition of inputs' centrality. Thanks to them, we can plot the U.S. economy as a production network[9] in Figure 3. According to the BEA (2002) Input-Output tables, we can represent it as a collection of 425 industries (i.e., nodes) linked by 51,768 transactions (i.e., edges). In Figure 3, we organize U.S. industries on a two-dimension space according to their reciprocal connectivity, following a Fruchterman and Reingold (1991) layout, which in our case posits more requested inputs at the center stage. Interestingly, services industries make the core of the U.S. production network because they are used as direct inputs in many other (manufacturing and services) industries.

Figure 3: Input-Output Network from U.S. BEA 2002 I-O tables

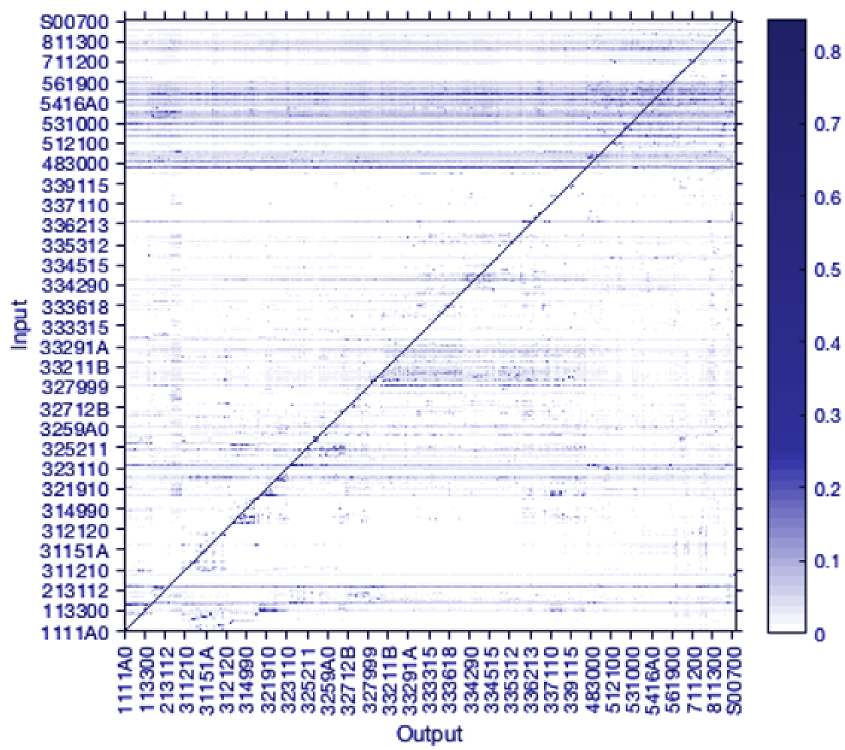


Note: Nodes represent 425 6-digit NAICS industries from the U.S. BEA 2002 Input-Output tables. Edges represent 51,768 industry-pair transactions. Network density: 0.286. Average path length: 1.7 links. The graph is visualized using a Fruchterman and Reingold (1991) layout with the GEPHI software. More connected industries (weighted out-degree) at the center stage. Selected industries in evidence.

[9]A bird's eye view of the U.S. production network represented in Figure 3 gives an idea of connectivity of each industry within a production network. In the U.S. I-O tables, we find a network density of 0.286, i.e., the fraction of actual linkages out of all potential linkages is relatively high. The average path length connecting any two industries is just 1.7 links, pointing to the U.S. economy's small-world nature. Briefly, on average, any producer in an output industry sources inputs from most of the other industries, either directly or indirectly. Indeed, the network of Figure 3 is connected – it is possible to go from an arbitrary node to any other node by following input linkages.

On the other hand, primary industries like agriculture and forestry are rather peripheral and mostly located in the north-west area of the graph because their inputs are not as widely used as other inputs. Among services, let us pick the case of R&D (code 541700) and Wholesale Trade (code 541800), which are among the most connected industries. In fact, wholesalers have a prominent role in professionally distributing many intermediate inputs in different moments of the production process, whereas R&D services are pivotal in fostering innovation across most U.S. sectors. Now, let us contrast them with two consumer goods industries: Electronic Computer Manufacturing (code 334111) and Automobile Manufacturing (code 336111). They appear to be at the periphery of the U.S. production network because they mostly meet final consumers.

Figure 4: Input Rank computed on U.S. 2002 Input-Output tables



Note: Input Rank vectors are computed for each of the 421 using industries classified at the 6-digit in the U.S. BEA 2002 tables. A darker cell implies that an input is more relevant for an output, i.e., a shock on that market has the potential to hit harder the producer of that output.

After a closer look, we register a strong heterogeneity in sourcing at the industry level. For example, in Appendix Figures C1 and C2, we report both the in-degree and out-degree distributions, i.e., the number of inputs received and the deliveries made by each industry in the U.S. production network. On average, the in-degree of an industry is higher than its out-degree. As expected, the industry with the highest number of inputs (296) is the Retail Trade (code 4A0000), because retailers professionally sell physical goods to consumers. On the other hand, the industry with the highest number of purchasing industries (425) is the Wholesale Trade (code 420000), because wholesalers professionally distribute intermediate physical inputs to all industries.

The Input Rank returns the economic relevance of an input market for a downstream producer taking into account all paths from input providers to the producer, and therefore is a global measure of positioning in the network. In Figure 4, we visualize the results from the computation of the Input Rank as a matrix of industry-pair values. A darker cell implies that an input industry is more relevant for the producers in a specific output industry. Interestingly, as from the upper part of the matrix, we find that services industries are the most economic relevant inputs for both manufacturing and services outputs. Please note how the diagonal is always darker, because most producers source an important amount of inputs from within their industry, hence intra-industry suppliers are much relevant.

Table 2: Top 20 inputs

Abs. Rank	I-O code	Industry name	avg. Input Rank	st. dev.
1	531000	Real Estate	.03442	.03390
2	550000	Management of Companies and Enterprises	.02369	.02977
3	420000	Wholesale Trade	.02067	.02903
4	211000	Oil and Gas Extraction	.01799	.01378
5	324110	Petroleum Refineries	.01425	.01342
6	221100	Electric Power Generation, Transmission, and Distribution	.01416	.02026
7	533000	Lessors of Nonfinancial Intangible Assets	.01383	.00770
8	517000	Telecommunications	.01350	.02307
9	52A000	Monetary Authorities and Depository Credit Intermediation	.01263	.02721
10	541800	Advertising and Related Services	.01254	.02084
11	331110	Iron and Steel Mills and Ferroalloy Manufacturing	.01216	.02891
12	484000	Truck Transportation	.01017	.01935
13	325190	Other Basic Organic Chemical Manufacturing	.00922	.01273
14	523000	Securities, Commodity Contracts, Investments, and Related Activities	.00870	.02910
15	221200	Natural Gas Distribution	.00783	.00779
16	325211	Plastic Material and Resin Manufacturing	.00736	.01384
17	522A00	Non-depository Credit Intermediation and Related Activities	.00729	.02284
18	722000	Food Services and Drinking Places	.00707	.02276
19	230301	Nonresidential Maintenance and Repair	.00682	.02646
20	541610	Management, Scientific, and Technical Consulting Services	.00645	.02803

Note: Average values computed for usage across using industries. Input Rank vectors are computed for each using industry among 421 industries classified at the 6-digit in the U.S. BEA 2002 tables.

Table 3: Top 20 manufacturing inputs

Abs. Rank	I-O code	Industry name	avg. Input Rank	st. dev.
5	324110	Petroleum Refineries	.01425	.01342
11	331110	Iron and Steel Mills and Ferroalloy Manufacturing	.01216	.02084
13	325190	Other Basic Organic Chemical Manufacturing	.00922	.01273
16	325211	Plastic Material and Resin Manufacturing	.00736	.01384
21	336300	Motor Vehicle Parts Manufacturing	.00645	.01634
26	334413	Semiconductor and Related Device Manufacturing	.00560	.01530
27	334418	Printed Circuit Assembly Manufacturing	.00525	.01347
28	325110	Petrochemical Manufacturing	.00502	.00708
29	32619A	Other Plastics Product Manufacturing	.00497	.01531
33	322210	Paperboard Container Manufacturing	.00466	.01181
38	321100	Sawmills and Wood Preservation	.00448	.01495
41	323110	Printing	.00410	.01906
43	322120	Paper Mills	.00403	.01360
48	3259A0	All Other Chemical Product and Preparation Manufacturing	.00365	.01077
50	326110	Plastics Packaging Materials and Unlaminated Film and Sheet Manufacturing	.00349	01280
52	332710	Machine Shops	.00345	.02262
54	33131A	Alumina Refining and Primary Aluminium Production	.00331	.014346
55	322130	Paperboard Mills	.00315	.01106
63	332800	Coating, Engraving, Heat Treating and Allied Activities	.00290	.01739
64	331411	Primary Smelting and Refining of Copper	.00283	.01312

Note: Average values computed for usage across using industries. Input Rank vectors are computed for each using industry among 421 industries classified at the 6-digit in the U.S. BEA 2002 tables. Absolute rank indicates the position when considering also services inputs.

In Table 2, we report the average Input Rank with standard deviations for top 20 inputs across all I-O industries. Here, as in Figure 4, we confirm that services industries rank on average higher than manufacturing. Top three inputs are Real Estate (code 531000), Management of Companies and Enterprises[10] (code 550000), and Wholesale Trade (code 420000), although they have a relatively high standard deviation, meaning that the sourcing strategies by using industries are much heterogeneous.

Further down, we find energy inputs including fossil fuels (codes 211000, 324110, and 221200) and electricity (code 221100). Royalties (code 533000), telecommunications (code 517000), the financial sector (codes 52A000, 523000, and 522A00), and R&D services (code 541610) are also relatively more important than many manufacturing industries. In fact, when we look only at top manufacturing inputs in Table 3, we have to go deep further down the rank with lower magnitudes and rising standard deviations. The ratio between the Input Ranks of top first and top 20th input in both Tables 2 and 3 is above 5, indicating that there is a relatively high skewness of the Input Rank distributions.

[10] This sector mainly gathers headquarters services by holding firms. As from the original definition (BLS, 2018): 'This sector comprises: i) companies that hold financial activities (securities or other equity interests) in other companies for the purpose of a corporate control to influence management decisions; ii) companies that professionally administer, oversee, and manage other companies through strategic or organizational planning and decision making.'

Clearly, values of the Input Rank are heterogeneous across different downstream industries. For example, in Appendix Tables C1 and C2, we look from the perspective of producers in Electronic Computer Manufacturing (code 334111) and in Automobile Manufacturing (code 336111). We find that the selection of top ten inputs by economic centrality always includes a combination of: i) manufacturing inputs that are predictably specific for their production processes (e.g., parts and components for cars), ii) and a selection of services (e.g., distribution for computers) that rank higher despite their relatively lower technical requirements in I-O tables.

4.4 Evidence from multi-country Input-Output tables

After computing the Input Rank on WIOD in year 2014, we report averages and standard deviations for top 10 supplying countries in Table 4. On the podium, we find China, Germany, and the USA. Notably, Russia is the fourth country with the highest average Input Rank thanks to its shipments of primary inputs and fuels that are relevant input markets for many countries and industries. Italy ranks also relatively high, on average, thanks to its centrality within many manufacturing supply chains, especially in Europe.

When we report averages and standard deviations for top input industries in Table 5, we find on the podium: i) Mining and Quarrying (code B); ii) Wholesale Trade (code G46); iii) Electricity, Gas, Steam and Air Conditioning Supply (code D35). Similarly to what observed in the case of U.S. I-O tables, we find that manufacturing inputs have a relatively lower market relevance than services and primary industries.

Table 4: Top 10 supplying countries by Input Rank in the World I-O tables (WIOD).

Abs. Rank	Origin country	Avg. Input Rank	st. dev.
1	China	.00352	.02773
2	Germany	.00310	.02409
3	USA	.00275	.02403
4	Russia	.00218	.02389
5	Italy	.00206	.02403
6	UK	.00178	.02395
7	France	.00168	.02390
8	S. Korea	.00150	.02489
9	Poland	.00146	.02386
10	Turkey	.00140	.02545

Note: The full Input Rank matrix on WIOD data has been computed starting from 43 countries *per* 56 industries. Average values and standard deviations reported for every triplet including: destination country, destination industry, sourcing industry.

Table 5: Top 10 input industries by Input Rank in the World I-O tables (WIOD).

Abs. Rank	I-O code	Origin industry	Avg. Input Rank	st. dev.
1	B	Mining and Quarrying	.00343	.02810
2	G46	Wholesale Trade	.00331	.02545
3	D35	Electricity, Gas, Steam and Air Conditioning Supply	.00309	.03349
4	C20	Manufacture of Chemicals and Chemical Products	.00286	.02558
5	N	Administrative and Support Service Activities	.00285	.02620
6	K64	Financial Service Activities	.00275	.02760
7	C19	Manufacture of Coke and Refined Petroleum Products	.00252	.02403
8	M69&M70	Legal and Accounting Activities; Head Offices; Management Consultancy	.00249	.02508
9	C24	Manufacture of Basic Metals	.00242	.02598
10	H49	Land Transport and Transport via Pipelines	.00227	.02501

Note: The full Input Rank matrix on WIOD data has been computed starting from 43 countries *per* 56 industries. Average values and standard deviations reported for every triplet including: destination country, destination industry, sourcing country.

However, please note that ISIC 2-digit industries in WIOD aggregate much more than NAICS 6-digit industries present in U.S. I-O tables, thus confounding several intermediate inputs, on one hand, and intermediate inputs with final goods in the same category, on the other hand. For this reason, we prefer keeping U.S. tables for following exercises on vertical integration (Section 5), when we are interested in measuring the economic relevance of an upstream input from wherever it has been sourced.

Nonetheless, we continue to show here how one can still use the Input Rank with geography, in order to assess the relevance of trading partners. For example, we look in Table 6 at the sourcing strategies of three major producing countries, USA, China, and Germany, and we assess the relevance of partner countries, where providers of intermediate inputs are located. Interestingly, on top of the ranks, we find that domestic producers are by far the most relevant because a domestic shock has the potential to affect domestic producers relatively more than foreign shocks. In second places, China is a crucial partner for both USA and Germany, whereas South Korea is the most central partner country for China. Further down in the ranks, we notice a predominance of regional trading partners, possibly located in the same region/continent, where economic integration is more likely to occur (e.g., factory Europe, factory America, and factory Asia).

Table 6: Top 10 origin countries by Input Rank for USA, China, and Germany

	USA			China			Germany		
Abs. Rank	Origin country	Avg. Input Rank	st. dev	Origin country	Avg. Input Rank	st. dev	Origin country	Avg. Input Rank	st. dev
1	USA	.04166	.14990	China	.05793	.17196	Germany	.03822	.15154
2	China	.00167	.00327	S. Korea	.00068	.00211	China	.00185	.00344
3	Canada	.00123	.00419	USA	.00062	.00096	Netherlands	.00162	.00507
4	Mexico	.00073	.00191	Japan	.00053	.00130	USA	.00149	.00237
5	Germany	.00048	.00089	Australia	.00051	.00166	Russia	.00101	.00319
6	Japan	.00046	.00127	Taiwan	.00040	.00203	France	.0010	.00203
7	Korea	.00037	.00090	Russia	.00039	.00105	Italy	.00094	.00200
8	UK	.00034	.00049	Germany	.00037	.00064	UK	.00081	.00137
9	Russia	.00026	.00067	Brazil	.00026	.00063	Poland	.00066	.00180
10	France	.00022	.00041	Indonesia	.00015	.00041	Belgium	.00065	.00180

Note: Input Rank vectors for USA, China and Germany have been computed starting from 43 countries *per* 56 industries possible origins of sourcing. Average values and standard deviations reported for within-country industry-level Input Ranks.

4.5 A pecking order in the geography of sourcing

To check whether the case studies showed in Table 6 represent regularities that can be extended to the rest of the data, we test simple binary regressions in Table 7, and repeat the exercise first for all industries and then for manufacturing industries only.

We consider the (log of the) Input Rank as a dependent variable, and we regress it against two binary variables. The idea is to catch the premia in relevance of inputs' markets while looking at the geography of sourcing. We argue that understanding the geography of inputs' markets is becoming increasingly relevant, even more after the outburst of a pandemic crisis in 2020, and the current conflict on European soil.

The first variable of interest in Table 7 is equal to one if the (direct or indirect) supplier is *Domestic*, and zero otherwise. Its scope is to catch the preference for domestic vs. foreign inputs. A second binary indicator is equal to one if the (direct or indirect) supplier comes from the same region/continent of the final buyers, i.e., they are *Intraregion*. Two-way fixed effects for supplying and buying industries are included. Standard errors are clustered multi-way following Cameron et al. (2011), based on the quartet of countries and industries both in origin and destination. The exercise is repeated for EU members, former NAFTA members, and for Asian countries[11]

Interestingly, the results clearly indicate a pecking order in the geography of upstream markets. Domestic producers rank relatively higher than intra-regional ones, which in turn rank higher than all other suppliers. In fact, we find that domestic (direct or indirect) suppliers are on average 5 times, 3.2 times, and 4.7 times more central for a representative buyer in the European Union, the former NAFTA[12], and Asian countries, respectively. Despite high trade openness, domestic producers are still by far the most relevant for most producers.

On the other hand, if any shock occurs on foreign upstream markets, our results show that final producers will be hit harder if they operate under the same regional trade agreements. The premium on centrality is, on average, 1.6 and 2 times more in the case of intra-EU and intra-NAFTA shipments. It is relatively less (0.85%) but still positive and statistically significant in the case of Asian producers.

[11] Please note that we included Australia in the group of Asian countries present in WIOD data (China, Indonesia, India, Japan, South Korea, Taiwan) because the country is among the biggest trading partners of all of them. Moreover, all these countries, including Australia, participate in the ASEAN Plus Six, an enlarged version of the Association of the South-East Asian Nations (ASEAN) that promotes regional economic integration.

[12] The North-American Free Trade Agreement (NAFTA) has been substituted by the United States–Mexico–Canada Agreement (USMCA). After renegotiations in the period 2017/2018, the USMCA entered into force in all member states on July 1, 2020.

Table 7: Input Rank: a pecking order on geography

Dependent variable (log of) Input Rank	EU	EU	NAFTA	NAFTA	ASEAN	ASEAN
Domestic	5.053***	4.299***	3.196***	2.078*	4.722***	4.042***
	(.281)	(.302)	(.703)	(.569)	(.383)	(.430)
Intraregion	1.596***	1.642***	1.984***	2.488***	.849***	.739**
	(.193)	(.171)	(.151)	(.075)	(.252)	(.221)
Constant	-13.226***	-12.758***	-11.251***	-10.630***	-11.550***	-10.585***
	(.373)	(.402)	(1.105)	(.806)	(.373)	(.478)
N. obs.	3,486,124	394,856	360,797	42,390	965,888	98,910
Activities	All	Manuf Only	All	Manuf Only	All	Manuf Only
N. origin countries	43+1	43+1	43+1	43+1	43+1	43+1
N. destination countries	28	28	3	3	7	7
Origin industry fe	Yes	Yes	Yes	Yes	Yes	Yes
Destination industry fe	Yes	Yes	Yes	Yes	Yes	Yes
Multiway clustered errors	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R squared	.366	.244	.344	.363	.489	.393

Note: Cross-section data for year 2014 are sourced from WIOD data (Timmer et al., 2015), including 43 countries, 56 industries, and a closure for the 'rest of the world'. Fixed effects by supplying and buying industries. Errors are clustered multiway following Cameron et al. (2011), considering origin country, supplying industry, destination country, and buying industry. ***, **, * stand for $p < .01$ and $p < .05$, and $p < .1$, respectively.

When we restrict our analysis to manufacturing industries only, in columns 2, 4, and 6 of Table 7, we find that there is a lower albeit still high centrality of domestic suppliers in all cases. Yet, in the case of former NAFTA, intra-regional manufacturing suppliers are, on average, about 2.5 times more central than other foreign suppliers. The latter finding is driven by a high degree of regional integration along the manufacturing supply chains developed across the U.S., Canadian, and Mexican borders (e.g., in the automotive industry), making all producers most sensitive to upstream shocks. In fact, the latter is the highest premium on inputs' centrality that we can recover in Table 7.

5 The role of the Input Rank in vertical integration

The decision to *make or buy* an input is the typical situation when a producer needs gathering information on the relevance of both direct and indirect inputs. In this Section, we test whether the Input Rank is a good predictor of vertical integration choices. The intuition is that vertical integration possibly allows mitigating the frictions from upstream markets. More in general, we expect that a parent company can internalize part of the shocks coming from (direct or indirect) suppliers. In this case, inputs that are 'central' in the sense we discussed in Section 3 would be given priority in vertical integration choices. Our empirical exercise takes on the firm-level framework by Alfaro et al. (2019) and Del Prete and Rungi (2017), while augmenting equations with the Input Rank besides the other main indicators for the relative importance of I-O linkages.

We will make use of a sample of U.S. parent companies that have integrated at least one production stage over time up to 2015. Ownership data and firms' accounts are sourced from the Orbis database, compiled by the Bureau van Dijk. For our scope, we collected information on 20,489 U.S. parent companies controlling 154,836 subsidiaries around the world[13]. In Table 8, we provide descriptive statistics of the geographic coverage of the subsidiaries. Both subsidiaries and parent companies can be active in any industry: manufacturing (28.86%), services (69%), primary (0.29%), and extractive (1.85%). About 81% of subsidiaries integrated by U.S. parents are domestic. Not surprisingly, U.S. parent companies are engaged mainly in global supply networks across other OECD economies, where 96% of their subsidiaries are located. The Member States of the European Union host the largest number of foreign subsidiaries in our sample. Among them, Germany, the United Kingdom, and the Netherlands attract a significant share of U.S. foreign subsidiaries active in services industries. Not surprisingly, USMCA (former NAFTA) partners, Canada and Mexico, mainly host manufacturing of final and intermediate goods. However, a non-negligible share of subsidiaries of U.S parent companies is present in Asia, Africa, and the Middle East.

To validate our sample, we compare with official figures reported by the report on the 'Activities of U.S. Multinational Enterprises' (BEA, 2018), and against OECD (2018) statistics. In 2015, BEA (2018) reports 6,880 billion dollars of total sales by foreign affiliates and 12,628 billion dollars of total sales by parent companies. The U.S. multinational enterprises present in our sample account for 94% and 92% of the BEA (2018) values, respectively. The number of foreign affiliates in our sample corresponds to 88.6% on the total of U.S. foreign subsidiaries reported in OECD (2018), although the latter source only reports the values for the year 2014.

[13]We follow international standards for the identification of corporate control structures (OECD, 2005; UNCTAD, 2009, 2016), according to which the unit of observation is the control link between a parent company and each of its subsidiary that is controlled after a concentration of voting rights (50% plus one stake). Similar data on MNEs have been used in Del Prete and Rungi (2017), Alviarez et al. (2017) and Cravino and Levchenko (2016)

Table 8: Geographic coverage

Hosting country	Final goods N.	%	Intermediate inputs N.	%	Services industries N.	%	All industries N.	%
United States	20,571	*16.3*	24,590	*19.5*	80.279	*64.1*	125,890	*100.0*
European Union	1,934	*11.5*	2,084	*12.3*	12,872	*76.2*	16.890	*100.0*
of which:								
Germany	273	*13.2*	306	*14.8*	1,494	*72.1*	2,073	*100.0*
France	171	*11.0*	213	*13.7*	1,167	*75.2*	1,551	*100.0*
United Kingdom	563	*11.4*	624	*12.7*	3,734	*75.9*	4,921	*100.0*
Italy	136	*19.4*	139	*19.8*	427	*60.8*	702	*100.0*
Netherlands	158	*6.8*	171	*7.3*	2,005	*85.9*	2,334	*100.0*
Canada	980	*30.4*	923	*28.6*	1,325	*41.1*	3,228	*100.0*
Russia	18	*11.7*	30	*19.5*	106	*68.8*	154	*100.0*
Asia	251	*15.0*	312	*18.7*	1,109	*66.3*	1,672	*100.0*
of which:								
Japan	87	*11.5*	76	*10.1*	592	*78.4*	755	*100.0*
China	92	*12.1*	66	*8.7*	605	*79.3*	763	*100.0*
India	122	*15.7*	149	*19.1*	508	*65.2*	779	*100.0*
Africa	67	*14.2*	93	*19.7*	313	*66.2*	473	*100.0*
Middle East	82	*18.2*	80	*17.8*	288	*64.0*	450	*100.0*
Other Americas	221	*12.1*	495	*21.6*	1,210	*66.3*	1,926	*100.0*
of which:								
Argentina	24	*8.1*	70	*23.6*	203	*68.4*	297	*100.0*
Brazil	137	*14.6*	219	*23.3*	583	*62.1*	939	*100.0*
Mexico	98	*23.3*	154	*36.6*	169	*40.1*	421	*100.0*
Australia	123	*14.2*	157	*18.1*	586	*67.7*	866	*100.0*
Rest of the world	489	*16.5*	585	*19.7*	1,892	*63.8*	2,966	*100.0*
Total	24,834	*16.0*	29,403	*19.0*	100,599	*65.0*	154,836	*100.0*

Note: firm-level data for year 2015 are sourced from Orbis, by Bureau Van Dijk. A U.S. parent company controls a foreign subsidiary with a direct or indirect equity stake higher than 50%, as from international standards. Classification by intermediate and final products is based on correspondence tables with Broad Economic Categories (BEC) rev. 4 provided by the UN Statistics Division.

For the scope of our analysis, we map industry affiliations of both parent companies and subsidiaries from the NAICS rev. 2012 classification onto the 2002 U.S. BEA I-O Input-Output Tables. The match by industry affiliations allows us combining firm-level data with sector-level metrics, including the Input Rank we computed as from Section 4. The underlying assumption is that the affiliation to the parent company indicates the buying industry (downstreamn), and the affiliation of each subsidiary indicates the (direct or indirect) supplying industry (upstream). In the absence of actual data on firm-to-firm transactions, such a mapping allows us proxying buyer-supplier relationships. For more details, see Section 4. Finally, we complement our data with input-industry estimates of demand elasticity sourced from Broda and Weinstein (2006). In Appendix Table C3, we report descriptive statistics of industry-level variables after the matching with sample firms and baseline estimates.

5.1 Empirical strategy

Let $h = 1, 2, \ldots, N$ denote the set of inputs, as from the I-O tables, and let $r = 1, 2, \ldots, R$ denote the set of parent companies, each active in an output industry, $k = 1, 2, \ldots, K$. The dependent variable, $y_{r(k)}$, takes on a value 1 when at least one subsidiary in the h-th input market has been integrated by a parent r in industry k, and 0 otherwise. Therefore, for each parent company, we have a vector $\boldsymbol{y}_{r(k)} = \left[y_{1r(k)}, \ldots, y_{Nr(k)}\right]$ made of 0s and 1s when an h-th input has been integrated or not, respectively. At this point, we can consider the probability that a generic parent chooses among a set of alternatives such that:

$$Pr\left(\boldsymbol{y}_{r(k)} | \sum_{h=1}^{N} y_{hr(k)}\right) = \frac{exp\left[y_{hr(k)}\boldsymbol{x}_{hr(k)}\boldsymbol{\beta}\right]}{\sum^{N}_{\boldsymbol{s}_{r(k)} \in \mathfrak{S}_{r(k)}} exp\left[\boldsymbol{s}_{r(k)}\boldsymbol{x}_{hr(k)}\boldsymbol{\beta}\right]} \tag{11}$$

where the element $s_{hr(k)}$ of a vector $\boldsymbol{s}_{r(k)}$ is equal to 1 if an input has been integrated, and zero otherwise. Thus, $\boldsymbol{s}_{r(k)}$ indicates the combination of inputs that have been integrated by the r-th parent company after considering all the possible combinations one could pick from $\mathfrak{S}_{r(k)}$. Therefore, we identify a vector of covariates for each input-output pair, $x_h r(k)$, which includes: i) the *Input Rank* of the h-th input with respect to the k-th output; ii) a binary variable *Complements* relative to the h-th input market; the *Upstreamness* sourced from Alfaro et al. (2019); iii) the Leontief inverse estimated by us, and iv) the bilateral direct requirement coefficient sourced directly from I-O tables. As in Antràs and Chor (2013) and Alfaro et al. (2019), the variable *Complements* is equal to 1 when the elasticity of substitution of the output market is above the median ($\rho_k > \rho_{med}$), and 0 otherwise. Errors are clustered by input markets. Fixed effects are included at the parent level. Results from nested specifications are reported in Tables 9.

5.2 Results

The odds ratios for the Input Rank are always higher than one and significant, therefore parent companies preferably integrate (direct or indirect) inputs that are more central in their supply network because a productivity shock on those upstream markets has the potential to hit harder on their downstream production process. The magnitudes of the odds ratios are quite stable across baseline specifications, in a range between 1.26 and 1.41. They are also quite stable after controlling for other indicators of I-O linkages.

Interestingly, odds ratios are lower than one on Upstreamness to indicate that it is less likely that a parent company integrates distant inputs in the production network, i.e., the more upstream they are the less likely they are integrated. From this point of view, it is clear enough that the Input Rank and the Upstreamness convey different information on positions along GVCs. In general, the coefficients on direct requirements are not statistically significant, whereas the coefficients on the Leontief inverse total coefficients are.

Table 9: Input Rank and vertical integration - Fixed-effects conditional logit

Dependent variable: input is integrated = 1	(1)	(2)	(3)	(4)	(5)
$Input\ Rank_{hk}$	1.413***	1.258***	1.195***	1.154***	1.114**
	(.093)	(.064)	(.082)	(.074)	(.101)
$Complements_h$			.554***	.581*	.522**
			(.121)	(.185)	(.158)
$Input\ Rank_{hk} \times\ Complements_h$			1.179	1.205	1.135
			(.115)	(.188)	(.174)
$Upstreamness_{hk}$		.545***	.540***	.566***	.509***
		(.055)	(.073)	(.092)	(.095)
$Leontief\ Inverse_{hk}$		1.055***	1.064***	1.032***	1.042***
		(.025)	(.019)	(.029)	(.025)
$Direct\ Requirement_{hk}$		1.095	1.094	1.085	1.096
		(.257)	(.383)	(.431)	(.721)
N. obs.	7,805,667	7,805,667	1,131,406	586,782	531,002
N. Parent companies	20,489	20,489	4,069	2,110	1,910
Parent fixed effects	Yes	Yes	Yes	Yes	Yes
Clustered errors	Yes	Yes	Yes	Yes	Yes
Activity of parent companies	All	All	All	Final	Intermediate
Pseudo R-squared (McFadden's)	.402	.409	.215	.114	.275
Log pseudo-likelihood	-94,000.003	-92,010.006	-24,185.115	-14,111.221	-10,774.411

Note: Odds ratios after a parent-level fixed-effects conditional logit are reported in the form $\frac{Pr(y_{hr(k)}=1)}{Pr(y_{hr(k)}=0)}$. Errors are clustered by I-O output industries. Variables are standardized at their mean and variance. *, **, *** stand for $p < .10$, $p < .05$, and $p < .01$, respectively. The variable Input Elasticity is available only for traded inputs.

Notably, we also find that parent companies less likely integrate suppliers from highly elastic industries, i.e., from suppliers that produce inputs whose substitutes are easier to find on the market. Please note that the main reason why the sample reduces from column 2 to column 3 onwards is because input elasticities are available only for traded inputs, as they are sourced from an exercise based on U.S. international trade data by Broda and Weinstein (2006).

Results are robust to controls for sample composition in Table 10, when we look only at manufacturing inputs (columns 1 and 2) or manufacturing outputs (columns 3 and 4), as well as when we look only at top 100 inputs with the highest direct requirement coefficients (columns 5 and 6). Finally, main findings are not sensitive to changes in the functional form. In Appendix Table C4, we test alternatively that simple probit, logit, and linear probability models present similar correlations between vertical integration choices and the Input Rank.

Table 10: Input Rank and vertical integration - Robustness to sample composition

Dependent variable: input is integrated = 1	Manuf input	Manuf input	Manuf output	Manuf output	Top 100	Top 100
$Input\ Rank_{hk}$	1.294***	1.305***	1.295***	1.302***	1.411***	1.234***
	(.057)	(.078)	(.063)	(.081)	(.070)	(.083)
$Complements_{h}$		.711		.674		.327***
		(.145)		(.198)		(.034)
$Input\ Rank_{:hk}$ X $Complements_{h}$		1.143		1.160		1.222
		(.190)		(.188)		(.175)
$Upstreamness_{hk}$	.495***	.529***	.555***	.760	.757	.823
	(.070)	(.075)	(.072)	(.155)	(.144)	(.179)
$Leontief\ Inverse_{hk}$	1.009	1.022	1.044**	1.028	1.062***	1.034*
	(.035)	(.025)	(.020)	(.071)	(.019)	(.018)
$Direct\ Requirement_{hk}$	1.009	1.022	1.044**	1.028	1.062***	1.034*
	(.035)	(.025)	(.020)	(.071)	(.019)	(.018)
N. obs.	935,648	893,464	1,396,382	1,103,884	254,201	154,824
N. Parent companies	2,175	2,134	4,166	3,970	256	240
Parent fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Clustered errors	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R-squared (McFadden's)	.237	.243	.189	.223	.264	.334
Log pseudo-likelihood	-20,135.425	-19,554.875	-31,249.311	-21,213.111	-13,744.276	-8,540.827

Note: Odds ratios after a parent-level fixed effects conditional logit are reported in the form $\frac{Pr(y_{hr(k)}=1)}{Pr(int_{hr(k)}=0)}$. Errors are clustered by I-O output industries. Variables are standardized at their mean and variance. *, **, *** stand for $p < .10$, $p < .05$, and $p < .01$, respectively. Please note, the variable Input Elasticity is available only for traded industries.

Please note, once again, that results reported in Tables 9, 10, and C4 cannot be interpreted in a structural way, because we do not provide a full incentive structure of the relations established between buyers and suppliers in a production network. For example, we do not consider the contractual environment that can endogenously determine the positions of suppliers and the decisions to keep them at arms' length. Yet, our intention is to show that network dimensions matter for GVCs, and correlations point to a significant role for input markets' centrality, in the sense introduced in Section 3, once we consider the potential impact of a shock on an upstream market.

6 Conclusions

In this contribution, we introduced the Input Rank as a theoretically sound measure to catch the relevance of both direct and indirect suppliers from the perspective of downstream buyers, whose production process can rely on increasingly sophisticated supply networks in times of a global fragmentation across interdependent markets. In our framework, a supplier ranks higher for a downstream buyer if a productivity shock has a higher impact on that buyer's marginal costs after transmission through the supply network. Besides a full account of the topology of a production network, we allow for a rich heterogeneity concerning how much firms rely on intermediate inputs, as well as on how much strong competitive forces are within each sector. At each stage of production, a lower intensity in the usage of intermediate inputs (i.e., a higher intensity in labor services) can buffer the transmission of a shock from upstream markets. At the same time, a higher markup (i.e., a lower competitive pressure) on an upstream market will make downstream buyers more sensitive to input-specific productivity shocks.

When we bring our model to I-O tables, we discover how central services industries are in production networks, because they connect across most manufacturing and services industries in a modern economy. After we look at international I-O tables, we retrieve a pecking order in the geography of sourcing. Most crucial inputs come first from domestic producers and then from regionally integrated partners.

Finally, we show that the Input Rank is a good predictor of vertical integration choices on a sample of U.S. parent companies controlling subsidiaries on a global scale, possibly because making an input in-house allows mitigating shocks from upstream markets.

Further work is needed to address, among others, the endogenous relationship between the organization of the global supply network and the contractual environment. However, in general, we argue that the Input Rank is useful from the perspective of downstream buyers, to understand how the recursive and complex nature of real-world supply networks can affect the ability to deliver in downstream markets. Ours is a first step that shows that technological non-linearities blend with endogenous market forces to shape global production networks, whose structure affects policies and shocks in globally integrated markets.

References

Acemoglu, D., Carvalho, V. M., Ozdaglar, A., Tahbaz-Salehi, A., 2012. The network origins of aggregate fluctuations. Econometrica 80 (5), 1977–2016.

Acemoglu, D., Johnson, S., Mitton, T., 2009. Determinants of Vertical Integration: Financial Development and Contracting Costs. The Journal of Finance 64 (3), 1251–1290.

Adao, R., Costinot, A., Donaldson, D., 2017. Nonparametric counterfactual predictions in neoclassical models of international trade. American Economic Review 107 (3), 633–89.

Alfaro, L., Antràs, P., Chor, D., Conconi, P., 2019. Internalizing global value chains: A firm-level analysis. Journal of Political Economy 127 (2), 508–559.

Alfaro, L., Conconi, P., Fadinger, H., Newman, A. F., 2016. Do Prices Determine Vertical Integration? Review of Economic Studies 83 (3), 855–888.

Alviarez, V., Cravino, J., Levchenko, A. A., 2017. The growth of multinational firms in the Great Recession. Journal of Monetary Economics 85 (C), 50–64.

Antràs, P., Chor, D., 2018. On the Measurement of Upstreamness and Downstreamness in Global Value Chains. Taylor & Francis Group, Ch. 5, pp. 126–194.

Antras, P., Chor, D., Forthcoming. Global Value Chains. Vol. 5. Elsevier, Ch. 5, pp. 297–366.

Antras, P., Chor, D., Fally, T., Hillberry, R., 2012. Measuring the Upstreamness of Production and Trade Flows. American Economic Review 102 (3), 412–416.

Antràs, P., Chor, D., 2013. Organizing the global value chain. Econometrica 81 (6), 2127–2204.

Antràs, P., de Gortari, A., 2020. On the Geography of Global Value Chains. Econometrica.

Atkeson, A., Burstein, A., 2008. Pricing-to-market, trade costs, and international relative prices. American Economic Review 98 (5), 1998–2031.

Automotive Industry Portal, 2020. Automotive Industry Portal, Marklines. 50,000 Supplier Database.
URL https://www.marklines.com/en/supplier_db

Baldwin, R., Freeman, R., Theodorakopoulos, A., September 2022. Horses for courses: Measuring foreign supply chain exposure. Working Paper 30525, National Bureau of Economic Research.

Baldwin, R., Venables, A. J., 2013. Spiders and snakes: Offshoring and agglomeration in the global economy. Journal of International Economics 90 (2), 245–254.

Baqaee, D. R., 2018. Cascading failures in production networks. Econometrica 86 (5), 1819–1838.

Baqaee, D. R., Farhi, E., 2020. Productivity and misallocation in general equilibrium. The Quarterly Journal of Economics 135 (1), 105–163.

BEA, 2002. Input output accounts data. US Bureau of Economic Analysis.
URL https://www.bea.gov/industry/input-output-accounts-data

BEA, 2018. Activities of U.S. Multinational Enterprises. U.S. Bureau of Economic Analysis.
URL https://www.bea.gov/data/intl-trade-investment/activities-us-multinational-enterprises-mnes

Berlingieri, G., Pisch, F., Steinwender, C., 2018. Organizing global supply chains: Input cost shares and vertical integration. Working Paper 25286, National Bureau of Economic Research.

Bernard, A. B., Dhyne, E., Magerman, G., Manova, K., Moxnes, A., 2022. The origins of firm heterogeneity: A production network approach. Journal of Political Economy 130 (7), 1765–1804.

Bernard, A. B., Moxnes, A., 2018. Networks and trade. Annual Review of Economics 10 (1), 65–85.

Bloch, F., Jackson, M. O., Tebaldi, P., 2019. Centrality measures in networks. Available at SSRN 2749124.

BLS, 2018. Industries at a Glance. Bureau of Labor Statistics.
URL https://www.bls.gov/iag/

Borin, A., Mancini, M., 2021. Measuring what matters in global value chains and value-added trade. Working Paper 8804, World Bank.

Brin, S., Page, L., 1998. The anatomy of a large-scale hypertextual web search engine. Computer Networks and ISDN Systems 30 (1), 107 – 117, proceedings of the Seventh International World Wide Web Conference.

Broda, C., Weinstein, D. E., 2006. Globalization and the Gains from Variety. The Quarterly Journal of Economics 121 (2), 541–585.

Caliendo, L., Parro, F., 2014. Estimates of the Trade and Welfare Effects of NAFTA. The Review of Economic Studies 82 (1), 1–44.

Cameron, A. C., Gelbach, J. B., Miller, D. L., 2011. Robust inference with multiway clustering. Journal of Business and Economic Statistics 29 (2), 238–249.

Carvalho, V. M., 2014. From micro to macro via production networks. Journal of Economic Perspectives 28 (4), 23–48.

Carvalho, V. M., Tahbaz-Salehi, A., 2019. Production networks: A primer. Annual Review of Economics 11 (1), 635–663.

Chaney, T., November 2014. The network structure of international trade. American Economic Review 104 (11), 3600–3634.

Chaney, T., 2016. Networks in international trade. In: Bramoullé, Y., Andrea, G., Rogers, B. (Eds.), The Oxford Handbook of the Economics of Networks. Oxford University Press, pp. 754–775.

Conconi, P., García-Santana, M., Puccio, L., Venturini, R., 2018. From final goods to inputs: The protectionist effect of rules of origin. American Economic Review 108 (8), 2335–65.

Costinot, A., Rodríguez-Clare, A., 2014. Chapter 4 - trade theory with numbers: Quantifying the consequences of globalization. In: Gopinath, G., Helpman, E., Rogoff, K. (Eds.), Handbook of International Economics. Vol. 4 of Handbook of International Economics. Elsevier, pp. 197 – 261.

Costinot, A., Vogel, J., Wang, S., 2012. An Elementary Theory of Global Supply Chains. The Review of Economic Studies 80 (1), 109–144.

Cravino, J., Levchenko, A. A., 2016. Multinational Firms and International Business Cycle Transmission. The Quarterly Journal of Economics 132 (2), 921–962.

De Benedictis, L., Tajoli, L., 2011. The world trade network. The World Economy 34 (8), 1417–1454.

de Gortari, A., 2019. Disentangling global value chains. Working Paper 25868, National Bureau of Economic Research.

Del Prete, D., Rungi, A., 2017. Organizing the global value chain: A firm-level test. Journal of International Economics 109, 16 – 30.

Del Prete, D., Rungi, A., 2020. Backward and forward integration along global value chains. Review of Industrial Organization 57 (2), 263–283.

Dhyne, E., Kikkawa, A. K., Magerman, G., 2022. Imperfect competition in firm-to-firm trade. Journal of the European Economic Association 20 (5), 1933–1970.

Fajgelbaum, P. D., Khandelwal, A. K., 2016. Measuring the Unequal Gains from Trade *. The Quarterly Journal of Economics 131 (3), 1113–1180.

Fally, T., Hillberry, R., 2018. A coasian model of international production chains. Journal of International Economics 114, 299 – 315.

Fruchterman, T. M. J., Reingold, E. M., 1991. Graph drawing by force-directed placement. Software: Practice and Experience 21 (11), 1129–1164.

Giammetti, R., Russo, A., Gallegati, M., 2020. Key sectors in input–output production networks: An application to brexit. The World Economy 43 (4), 840–870.

Gleich, D. F., 2015. Pagerank beyond the web. SIAM Review 57 (3), 321–363.

Grassi, B., 2017. IO in I-O: Size, Industrial Organization, and the Input-Output Network Make a Firm Structurally Important. mimeo.

Grassi, B., Sauvagnat, J., 2019. Production networks and economic policy. Oxford Review of Economic Policy 35 (4), 638–677.

Huremović, K., Jiménez, G., Moral-Benito, E., Peydró, J.-L., Vega-Redondo, F., 2020. Production and financial networks in interplay: Crisis evidence from supplier-customer and credit registers. CEPR Discussion Papers, No. DP15277.

Imbs, J., Pauwels, L., 2022. Measuring openness. Working Paper 17230, CEPR.

Johnson, R. C., 2018. Measuring global value chains. Annual Review of Economics 10 (1), 207–236.

Katz, L., 1953. A new status index derived from sociometric analysis. Psychometrika 18 (1), 39–43.

Koopman, R., Wang, Z., Wei, S.-J., 2014. Tracing value-added and double counting in gross exports. American Economic Review 104 (2), 459–94.

Miller, R. E., Temurshoev, U., 2017. Output upstreamness and input downstreamness of industries/countries in world production. International Regional Science Review 40 (5), 443–475.

Mundt, P., 2021. The formation of input–output architecture: Evidence from the european union. Journal of Economic Behavior & Organization 183, 89–104.

Newman, M. E. J., 2018. Networks, 2nd Edition. Vol. 4 of 10. Oxford University Press, New York.

Nunn, N., 2007. Relationship-specificity, incomplete contracts and the pattern of trade. Quarterly Journal of Economics 122 (2), 569–600.

Nunn, N., Trefler, D., 2013. Incomplete contracts and the boundaries of the multinational firm. Journal of Economic Behavior and Organization 94 (1), 330–344.

Nunn, N., Trefler, D., 2014. Domestic institutions as a source of comparative advantage. In: Gopinath, G., Helpman, E., Rogoff, K. (Eds.), Handbook of International Economics. Vol. 4 of Handbook of International Economics. Elsevier, pp. 263 – 315.

OECD, 2005. Guidelines for Multinational Enterprises. Organization for Co-operation and Economic Development.

OECD, 2018. The Analytical AMNE database - Multinational enterprises and Global Value Chains. Organization for Economic Co-operation and Development.
URL https://www.oecd.org/sti/ind/analytical-amne-database.htm

Orefice, G., Rocha, N., 2014. Deep integration and production networks: An empirical analysis. The World Economy 37 (1), 106–136.

Rauch, J. E., 1999. Networks versus markets in international trade. Journal of International Economics 48 (1), 7 – 35.

Rungi, A., Del Prete, D., 2018. The smile curve at the firm level: Where value is added along supply chains. Economics Letters 164 (C), 38–42.

Timmer, M. P., Dietzenbacher, E., Los, B., Stehrer, R., 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.

UNCTAD, 2009. UNCTAD Training Manual on Statistics for FDI and the Operation of TNCs - Volume I: FDI Flows and Stocks. United Nations Conference on Trade and Development.

UNCTAD, 2016. World investment report 2016. Investor Nationality: Policy Challenges. United Nations Conference on Trade and Development.

Wang, Z., Wei, S.-J., Yu, X., Zhu, K., Mar. 2017. Measures of Participation in Global Value Chains and Global Business Cycles. Nber working papers, National Bureau of Economic Research, Inc.

World Bank, 2020. World development report 2020. trading for development in the age of global value chains. World Bank.

Appendix A: Proofs

Lemma 1. Let N_k^+ denote the set of inputs of firms in sector k. The cost function of firm i is given by

$$c(y(k,i);w,\mathbf{p}) = \lambda(k,i)y(k,i),$$

where $\lambda(k,i) = w^{\beta_k}\prod_{h\in N_k^+} p_h^{\alpha_k g_{hk}}\tau_h^{\alpha_k g_{hk}}$.

Proof of Lemma 1. The Langragian of the cost minimization problem of firm i from sector k:

$$\mathcal{L} = w\ell(k,i) + \sum_{h=1}^{M} p_h x(k,i,h) - \lambda(k,i)\left[\zeta_k \ell(k,i)^{\beta_k}\left(\prod_{h=1}^{M}(\tau_h^{-1}x(k,i,h))^{g_{hk}}\right)^{\alpha_k} - y(k,i)\right]$$

From the first-order necessary conditions (also sufficient, given convexity), we can deliver the following conditional demand functions:

$$x(k,i,h) = \lambda(k,i)\alpha_k g_{hk}\frac{y(k,i)}{p_h};\quad \ell(k,i) = \lambda(k,i)\beta_k\frac{y(k,i)}{w}. \tag{A1}$$

Plugging (A1) into the expression for the cost of firm i, it directly follows that $c(y(k,i),w\mathbf{p}) = \lambda(k,i)y(k,i)$. Hence $\lambda(k,i)$ is the marginal cost of production of firm i.

Substituting (A1) into the production function we obtain:

$$
\begin{aligned}
y(k,i) =& \zeta_k \left(\frac{\lambda(k,i)\beta_k y(k,i)}{w}\right)^{\beta_k} \prod_{h\in N_k^+} \left(\tau_h^{-1}\frac{\lambda(k,i)\alpha_k g_{hk} y(k,i)}{p_h}\right)^{g_{hk}\alpha_k} \\
=& \zeta_k \lambda(k,i) y(k,i) \left(\frac{\beta_k}{w}\right)^{\beta_k} \prod_{h\in N_k^+} \left(\frac{g_{hk}\alpha_k}{\tau_h p_h}\right)^{g_{hk}\alpha_k} \\
=& \lambda(k,i) w^{-\beta_k} \prod_{h\in N_k^+} (p_h\tau_h)^{-g_{hk}\alpha_k}\, y(k,i),
\end{aligned}
$$

where for the last equality we used the fact that $\zeta_k = \beta_k^{-\beta_k}\prod_{h\in N_k^+}(\alpha_k g_{hk})^{-\alpha_k g_{hk}}$.

Solving for $\lambda(k,i)$, we get:

$$
\lambda(k,i) = w^{\beta_k} \prod_{h\in N_k^+} p_h^{\alpha_k g_{hk}} \tau_h^{-\alpha_k g_{hk}}. \tag{A2}
$$

as desired. □

Lemma 2. The following relations between firm-level marginal cost $\lambda(k,i)$ and sector-level marginal cost λ_k, and firm level markup $\mu(k,i)$ and sector level markup μ_k hold:

$$
\begin{aligned}
\lambda_k &= M_k^{\frac{1}{1-\epsilon_k}} \lambda(k,i), \\
\mu_k &= \mu(k,i) = \frac{\epsilon_k}{\epsilon_k - 1}
\end{aligned}
$$

Proof of Lemma 2. Using results from the theory of monopolistic competition, the sector-level price of good k, p_k, and the sector-level output, y_k, are given by:

$$
\begin{aligned}
p_k &= \left(\sum_{i=1}^{M_k} p(k,i)^{1-\epsilon_k}\right)^{\frac{1}{1-\epsilon_k}}, \\
y_k &= \left(\sum_{i=1}^{M_k} y(k,i)^{\frac{\epsilon_k - 1}{\epsilon_k}}\right)^{\frac{\epsilon_k}{\epsilon_k - 1}}.
\end{aligned} \tag{A3}
$$

In the symmetric equilibrium, $p(k,i) = p(k,j)$, and $y(k,i) = y(k,j)$. Hence:

$$p_k = M_k^{\frac{1}{1-\epsilon_k}} p(k,i),$$
$$y_k = M_k^{\frac{\epsilon_k}{\epsilon_k - 1}} y(k,i).$$

Thanks to the assumption on constant returns to scale, and using the expression for y_k, we can write the sector-level marginal cost of production as:

$$\lambda_k = \sum_{i=1}^{M_k} \lambda(k,i) \frac{y(k,i)}{y_k} = \frac{M_k \lambda(k,i) y(k,i)}{M_k^{\frac{\epsilon_k}{\epsilon_k - 1}} y(k,i)} = M_k^{\frac{1}{1-\epsilon_k}} \lambda(k,i)$$

Finally, from the firm pricing rule we have $p(k,i) = \mu(k,i)\lambda(k,i)$. Plugging the pricing rule in the expression for p_k from (A3) we get:

$$\begin{aligned} p_k &= \left(\sum_{i=1}^{M_k} (\mu(k,i)\lambda(k,i))^{1-\epsilon_k} \right)^{\frac{1}{1-\epsilon_k}} \\ &= \left[M_k \mu(k,i)^{1-\epsilon_k} \lambda(k,i)^{1-\epsilon_k} \right]^{\frac{1}{1-\epsilon_k}} = M_k^{\frac{1}{1-\epsilon_k}} \mu(k,i)\lambda(k,i) = \mu(k,i)\lambda_k \end{aligned}$$

which completes the proof. □

Lemma 3. The aggregate demand of sector k for deliveries of input h is given by:

$$x(k,h) = \frac{\alpha_k g_{hk} p_k y_k}{\mu_k p_h} = \frac{\epsilon_k - 1}{\epsilon_k} \frac{\alpha_k g_{hk} p_k y_k}{p_h}. \tag{A4}$$

***Proof of Lemma** 3.* Symmetry implies that $x(k,h) = M_k x(k,i,h)$. From (A1) and Lemma 2, we get:

$$x(k,h) = M_k \lambda(k,i) \alpha_k g_{hk} \frac{y(k,i)}{p_h} = M_k M_k^{-\frac{1}{1-\epsilon_k}} \frac{p_k}{\mu_k} \alpha_k g_{hk} M_k^{-\frac{\epsilon_k}{\epsilon_k - 1}} y_k = \frac{\alpha_k g_{hk} p_k y_k}{\mu_k p_h}$$

where we used the fact that $p_k = M_k^{\frac{1}{1-\epsilon_k}} \mu(k,i)\lambda(k,i)$ and that $\mu_k = \mu(k,i)$ for $i \in k$. □

Proof of Proposition 1. For notational simplicity, let $\check{x}$ denote $\log x$ for any variable x. Taking logarithms of (A2) and using the pricing rule ($p_h = \mu_h \lambda_h$), we get:

$$\check{\lambda}(k,i) = \beta_k \check{w} + \alpha_k \sum_{h \in N_K^+} g_{hk} \left(\check{\mu}_h + \check{\tau}_h\right) + \alpha_k \sum_{h \in N_k^+} g_{hk} \check{\lambda}_h.$$

From Lemma 2, we can write the above equation in terms of sector level marginal cost as:

$$\check{\lambda}_k = \frac{1}{\epsilon_k - 1} \check{M}_k + \beta_k \check{w} + \alpha_k \sum_{h \in N_k^+} g_{hk} \left(\check{\mu}_h + \check{\tau}_h\right) + \alpha_k \sum_{h \in N_k^+} g_{hk} \check{\lambda}_h$$

Define column vector $\boldsymbol{\eta} = \{\frac{1}{\epsilon_k - 1} \check{M}_k\}_{k=1}^M$. Moreover normalize $w = 1$. Writing the above equation for all k's in vector notation we get:

$$\begin{aligned}
\check{\boldsymbol{\lambda}} =& \boldsymbol{\eta} + \mathbf{AG}'\check{\boldsymbol{\mu}} + \mathbf{AG}'\boldsymbol{\tau} + \mathbf{AG}'\boldsymbol{\lambda} \Rightarrow \\
\check{\boldsymbol{\lambda}} =& \left[\mathbf{I} - \mathbf{AG}'\right]^{-1} \left(\boldsymbol{\eta} + \mathbf{AG}'(\check{\boldsymbol{\mu}} + \check{\boldsymbol{\tau}})\right).
\end{aligned} \tag{A5}$$

Finally, by differentiating we get:

$$\begin{aligned}
\frac{\partial \check{\lambda}(k,i)}{\partial \check{\tau}_h} =& \frac{\partial \check{\lambda}_k}{\partial \check{\tau}_h} = \boldsymbol{e}_k' \left[\mathbf{I} - \mathbf{AG}'\right]^{-1} \mathbf{AG}' \boldsymbol{e}_h = \boldsymbol{e}_k' \left[\left[\mathbf{I} - \mathbf{AG}'\right]^{-1} - \mathbf{I}\right] \boldsymbol{e}_h \\
=& \boldsymbol{e}_h' \left[\left[\mathbf{I} - \mathbf{GA}\right]^{-1} - \mathbf{I}\right] \boldsymbol{e}_k.
\end{aligned}$$

Whenever $k \neq h$ the above equation has a form:

$$\frac{\partial \check{\lambda}(k,i)}{\partial \check{\tau}_h} = \boldsymbol{e}_h' \left[\mathbf{I} - \mathbf{GA}\right]^{-1} \boldsymbol{e}_k,$$

which completes the proof. □

Proof of Proposition 2. From (A5) we have:

$$\check{\lambda}_k = \sum_{\ell=1}^{M} \upsilon_{\ell k}\eta_\ell + \sum_{\ell=1}^{M} \upsilon_{\ell k}\check{\mu}_\ell + \sum_{\ell=1}^{M} \upsilon_{\ell k}\check{\tau}_\ell - \check{\mu}_k - \check{\tau}_k,$$

where $\upsilon_{\ell k}$ is the Input Rank of input ℓ for firms in k. Using Lemma 2 we get:

$$\check{\lambda}(k,i) = \sum_{\ell=1}^{M} \upsilon_{\ell k}\eta_\ell + \sum_{\ell=1}^{M} \upsilon_{\ell k}\check{\mu}_\ell + \sum_{\ell=1}^{M} \upsilon_{\ell k}\check{\tau}_\ell - \check{\mu}_k - \check{\tau}_k - \eta_k,$$

which gives:

$$\lambda(k,i) = \exp\left(\sum_{\ell=1}^{M} \upsilon_{\ell k}\eta_\ell + \sum_{\ell=1}^{M} \upsilon_{\ell k}\check{\mu}_\ell + \sum_{\ell=1}^{M} \upsilon_{\ell k}\check{\tau}_\ell - \check{\mu}_k - \check{\tau}_k - \eta_k\right)$$

where exp denotes the exponential function. Define:

$$\delta_k \equiv \sum_{\ell=1}^{M} \upsilon_{\ell k}\eta_\ell + \sum_{\ell=1}^{M} \upsilon_{\ell k}\check{\mu}_\ell + \sum_{\ell=1}^{M} \upsilon_{\ell k}\check{\tau}_\ell - \check{\mu}_k - \check{\tau}_k - \eta_k.$$

By simple differentiation of $\lambda(k,i)$ we get:

$$\frac{\partial \lambda(k,i)}{\partial \tau_h} = \frac{\upsilon_{hk}}{\tau_h} e^{\delta_k},$$

from where if follows directly that $\frac{\partial^2 \lambda(k,i)}{\partial \tau_h \partial \upsilon_{hk}} > 0$.

To calculate $\frac{\partial^2 \lambda(k,i)}{\partial \tau_h \partial \epsilon_h}$ we recall that $\mu_h = \frac{\epsilon_h}{\epsilon_h - 1}$ and $\eta_k = \frac{1}{\epsilon_k - 1}\check{M}_k$, an thus:

$$\frac{\partial^2 \lambda(k,i)}{\partial \tau_h \partial \epsilon_h} = \frac{\upsilon_{hk}^2}{\tau_h} e^{\delta_k} \left(\frac{\partial \eta_h}{\partial \epsilon_h} + \frac{\partial \check{\mu}_h}{\partial \epsilon_h}\right) = \frac{\upsilon_{hk}^2}{\tau_h} e^{\delta_k} \frac{1 - \epsilon_h(1 + \check{M}_h)}{(\epsilon_h - 1)^2 \epsilon_h} < 0,$$

where the last inequality comes from the fact that $M_h \geq 1$ and $\epsilon_h > 1$, which completes the proof.

$\square$

Appendix B: Imperfect information

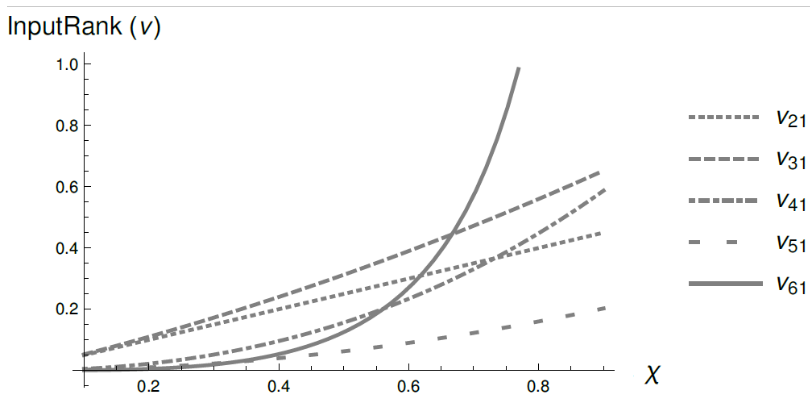
In this Appendix, we consider the case of a producer that does not fully observe her supply network. Intuitively, a firm cannot outreach the entire web of specialized suppliers when production networks are more sophisticated. To fix ideas, it is worth looking back at the fictional supply network reported in Figure 1. Any time the manager of a firm 1 tries to collect information about upstream transactions, say about transactions between firm 3 and its suppliers, she has a limited ability to know the quality and quantity of deliveries. She can call the direct supplier to ask or, alternatively, she can gather information on the market when, for example, prices and qualities of upstream inputs are relatively standard terms to include in a contract. At any upstream step, e.g. from firm 4 up to firm 6, the same problem starts all over again.

We capture this phenomena in a stylized way by assuming that firm i in a sector k does not observe the full network $\mathbf{G}$ but rather its subnetwork $\chi_k\mathbf{G}$, where one can think of χ_k as a parameter capturing the share of transactions in the production network observed by firms form sector k. We assume that the extent to which the network is observed (χ_k) is output-specific, i.e., it varies according to the peculiar contractual environment of the downstream market. In this, we follow the seminal intuition by Rauch (1999), who sketched the idea of a network search on international trade when products are differentiated or homogeneous. In line with that intuition, we can think of χ_k as a search barrier in supply networks. Even more realistically, one may consider an extension to the case when firms in sector k observe suppliers of firms in sector h with independent probabilities χ_{hk}. Then, we would replace a scalar χ_k with a diagonal matrix $\mathbf{H}_k$ that has diagonal elements equal to χ_{hk}. Taking the output as a reference, we align with Nunn (2007), who looks at an average measure of input contract intensity to infer the thickness and the relation-specificity of the markets.

Hence, when assessing the importance of suppliers from sector h, instead of relying on Input Rank $\nu_{hk}(\mathbf{GA}) = \mathbf{e}_h'(\mathbf{I}-\mathbf{GA})^{-\mathbf{1}}\mathbf{e}_k$, firms from sector k consider Input Rank

$\nu_{hk}(\chi_k \mathbf{GA}) = \mathbf{e}'_h(\mathbf{I} - \chi_\mathbf{k}\mathbf{GA})^{-1}\mathbf{e}_k$. Since $\chi_k \leq 1$, firm $i \in k$ underestimates the importance of indirect suppliers. When χ_k is smaller, suppliers that are relatively closer to the final producer will have a relatively higher Input Rank than more distant suppliers. To illustrate this property, we calculate the Input Rank using the network from Figure 1 at changing values of the dampening rates χ (assumed to be equal for each node k) and plot it in Figure B1.

Figure B1: Input Rank as a function the searching parameter χ

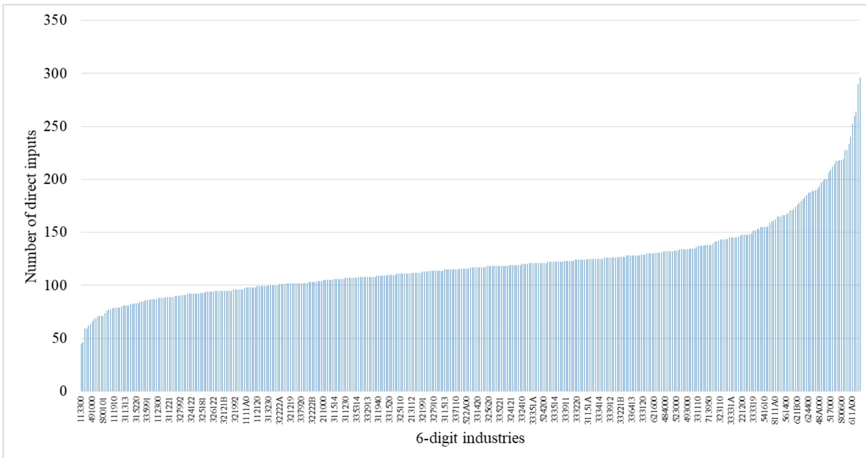


Note: Input Ranks for the output of firm 1 calculated for a fictional supply network depicted in Figure 1, assuming symmetric input deliveries and labor intensities, as functions of χ.

When plotting Figure B1, we assume that for any fixed output node k and any two of its suppliers h_1 and h_2, we have $g_{h_1k} = g_{h_2k}$. Interestingly, although there are more (upstream) paths connecting firm 1 to firm 6 than firm 1 to firm 2, firm 2 will have a disproportionately higher Input Rank when the dampening rate becomes smaller and smaller. In a nutshell, a limited ability to outreach on the input markets implies that downstream buyers underestimate the role of longer paths in production network. In other words, search barriers can prevent the downstream buyers to explore production processes that are too distant in the supply structure.

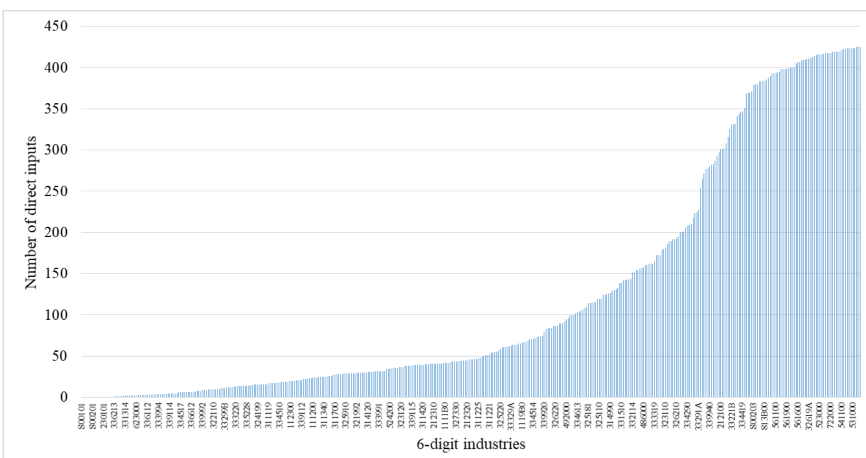
Appendix C: Tables and Graphs

Figure C1: In-degree distribution of Input-Output Network from U.S. BEA 2002 I-O tables



Note: Number of input industries by output ordered on the x-axis. Average: 122. Minimum at the Logging industry (code 113300) is 45. Maximum at the Retail Trade (code 4A0000) is 296.

Figure C2: Out-degree distribution of Input-Output Network from U.S. BEA 2002 I-O tables



Note: Number of buying industries by output ordered on the x-axis. Average: 122. Minimum at the Museums, Historical Sites, Zoos, and Parks (code 712000) is 0. Maximum at the Wholesale Trade (code 420000) is 425.

Table C1: Top 10 inputs of the Automobile Manufacturing (I-O code: 336111)

Abs. Rank	I-O code	Industry name	Input Rank
1	336111	Automobile Manufacturing	.10263
2	336300	Motor Vehicle Parts Manufacturing	.05608
3	420000	Wholesale Trade	.01183
4	550000	Management of Companies and Enterprises	.01103
5	531000	Real Estate	.00965
6	331110	Iron and Steel Mills and Ferroalloy Manufacturing	.00834
7	211000	Oil and Gas Extraction	.00721
8	533000	Lessors of Nonfinancial Intangible Assets	.00521
9	221100	Electric Power Generation, Transmission, and Distribution	.00492
10	324110	Petroleum Refineries	.00492

Note: The full Input Rank vector of the Automobile Manufacturing (code 336111) has been computed starting from all 421 industries classified at the 6-digit in the U.S. BEA 2002 tables.

Table C2: Top 10 inputs of the Electronic Computer Manufacturing (I-O code: 334111)

Abs. Rank	I-O code	Industry name	Input Rank
1	334111	Electronic Computer Manufacturing	.09650
2	334112	Computer Storage Device Manufacturing	.01882
3	334418	Printed Circuit Assembly Manufacturing	.01472
4	420000	Wholesale Trade	.01354
5	334413	Semiconductor and Related Device Manufacturing	.01318
6	550000	Management of Companies and Enterprises	.01220
7	531000	Real Estate	.00981
8	511200	Software Publishers	.00675
9	33411A	Computer Terminals and Other Computer Peripheral Equipment	.00626
10	541800	Advertising and Related Services	.00459

Note: The full Input Rank vector of the Electronic Computer Manufacturing (code 334111) has been computed starting from all 421 industries classified at the 6-digit in the U.S. BEA 2002 tables.

Table C3: Descriptive statistics of industry-level variables

Variable	Mean	St. Dev.	Min.	Max	N. obs.
Input Rank	.0021	.0183	.0001	.5530	1,131,406
Upstreamness	3.2752	.8865	1.0029	8.7470	1,131,406
Direct requirement coeff.	.0009	.0073	0	.4194	1,131,406
Elasticity of substitution	8.7239	11.8644	1.3	108.5019	1,131,406

Note: Input Rank, Upstreamness, and direct requirement coefficients are based on U.S. I-O 2002 Tables sourced from BEA (2002). Elasticities of substitution for inputs are sourced from Broda and Weinstein (2006). Number of sample observations refer to the last columnn of baseline estimates in Table 9.

Table C4: Input Rank and vertical integration - Different functional forms

Dependent variable: input is integrated = 1	LPM	LPM	Logit	Logit	Probit	Probit
$Input\ Rank_{hk}$	.026***	.023***	1.315***	1.259***	.141***	.132***
	(.009)	(.009)	(.064)	(.070)	(.036)	(.027)
$Complements\ [\epsilon > med\,]_h$		-.002		.324		-.185
		(.005)		(.420)		(.172)
$Input\ Rank_{hk} \times\ Complements\ [\epsilon > med\,]_h$		.019		1.145		.065
		(.017)		(.099)		(.058)
$Upstreamness_{hk}$	-.001**	-.001**	.547***	.511***	-.229***	-217***
	(.001)	(.001)	(.051)	(.061)	(.029)	(.035)
$Leontief\ Inverse_{hk}$	.031***	.051***	.069***	.065***	.059***	.045***
	(.011)	(.015)	(.017)	(.016)	(.008)	(.008)
$Direct\ Requirement_{hk}$	.001	-.001	1.002	1.016	.009	.008
	(.001)	(.001)	(.157)	(.165)	(.008)	(.008)
$Constant$	.006***	.006***	.003***	.004***	-2.738***	-2.669***
	(.001)	(.001)	(.001)	(.001)	(.038)	(.050)
N. obs.	7,805,667	1,257,911	7,805,667	1,257,911	7,805,667	1,257,911
N. Parent companies	20,467	3,510	20,467	3,510	20,467	3,510
Parent fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Clustered errors	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R-squared	.081	.094	-	-	-	-
Pseudo R-squared (McFadden's)	-	-	.145	.179	.158	.189
Log pseudo-likelihood	-	-	-40,544.433	-29,743.55	-37,549.143	-29,589.110

Note: Odds ratios are reported when we run a logit model, in the form $\frac{Pr(y_{hr(k)}=1)}{Pr(y_{hr(k)}=0)}$. Marginal effects and beta coefficients are reported for probit specifications and linear probability models, respectively. Errors are clustered by I-O output industries. Variables are standardized at their mean and variance. *, **, *** stand for $p < .10$, $p < .05$, and $p < .01$, respectively. Please note, the variable $Input\ Elasticity\ [\epsilon > med\,]_h$ is available only for traded industries.