On the Geography of Global Value Chains*

Pol Antràs
Harvard University and NBER
pantras@fas.harvard.edu

Alonso de Gortari
Harvard University
adegortaribriseno@fas.harvard.edu

May 24, 2017

Abstract

This paper develops a multi-stage general-equilibrium model of global value chains (GVCs) and studies the specialization of countries within GVCs in a world with barriers to international trade. With costly trade, the optimal location of production of a given stage in a GVC is not only a function of the marginal cost at which that stage can be produced in a given country, but is also shaped by the proximity of that location to the precedent and the subsequent desired locations of production. We show that, other things equal, it is optimal to locate relatively downstream stages of production in relatively central locations. We also develop and estimate a tractable, quantifiable version of our model that illustrates how changes in trade costs affect the extent to which various countries participate in domestic, regional or global value chains, and traces the real income consequences of these changes.

*We thank Arnaud Costinot and Iván Werning for useful conversations during the most upstream stages of this project. Roc Armenter, Rob Johnson, Myrto Kalouptsidi, Sam Kortum, Marc Melitz, Eduardo Morales, Felix Tintelnot, and Kei-Mu Yi provided very useful feedback on preliminary versions of the paper. We are also grateful to seminar audiences at Princeton, the North American Econometric Society Summer Meeting in Philadelphia, the SED in Toulouse, the Federal Reserve Bank of Dallas, Geneva, Chicago Booth, MIT, Stanford, Harvard, Brown, Clark, UCLA, Wisconsin, CREI-UPF, Autònoma in Barcelona, Toulouse, CEMFI and Toronto for useful feedback. Le Kang, Daniel Ramos, BooKang Seol, Maria Voronina, and Diana Zhu provided valuable research assistance at various stages of this paper's production chain. Antràs acknowledges support from the NSF (proposal #1628852). All errors are our own.

1 Introduction

In recent decades, technological progress and falling trade barriers have allowed firms to slice up their value chains, retaining within their domestic economies only a subset of the stages in these value chains. The rise of global value chains (GVCs) has dramatically changed the landscape of the international organization of production, placing the specialization of countries *within* GVCs at the center stage. Where in GVCs are different countries specializing? Should countries use specific policies to place themselves in particularly appealing segments of GVCs? These are questions being posed in the policy arena for which the academic literature has yet to provide satisfactory answers.

This paper studies the specialization of countries within GVCs in a world with barriers to international trade. Although we are motivated by normative questions, the focus of this paper is on outlining the implications of the existence of *exogenously* given trade costs for the equilibrium shape of GVCs. The role of trade barriers on the geography of GVCs is interesting in its own right and has been relatively underexplored in the literature, perhaps due to the technical difficulties that such an analysis entails. More specifically, characterizing the allocation of production stages to countries is not straightforward because the optimal location of production of a given stage in a GVC is not only a function of the marginal cost at which that stage can be produced in a given country, but is also shaped by the proximity of that location to the precedent and the subsequent desired locations of production.

We start off our analysis in section 2 by illustrating these interdependencies in a simple partial equilibrium environment. We consider the problem of a *lead firm* choosing the location of its various production stages in an environment with costly trade. A key insight from our partial-equilibrium framework is that the relevance of geography (or trade costs) in shaping the location of the various stages of a GVC is more and more pronounced as one moves towards more and more downstream stages of a value chain. Intuitively, whenever trade costs are largely proportional to the gross value of the good being transported, these costs compound along the value chain, thus implying that trade costs erode more value added in downstream relative to upstream stages. In a parameterized example of our framework, this differential effect of trade costs takes the simple form of a stage-specific 'trade elasticity' that is increasing in the position of a stage in the value chain. The fact that trade costs are proportional to gross value follows from our iceberg formulation of these costs, a formulation that is not only theoretically appealing, but is also a reasonable approximation to reality.[1]

Having characterized the key properties of the solution to the lead-firm problem, we next show how it can be 'decentralized'. More specifically, we consider an environment in which there is no lead firm coordinating the chain, and instead stand-alone producers of the various stages in a GVC make cost-minimizing sourcing decisions by purchasing the good completed up to the prior

[1]The fact that import duties and insurance costs are approximately proportional to the value of the goods being shipped should be largely uncontroversial. For shipping costs, weight and volume are naturally also relevant, but as shown by Brancaccio, Kalouptsidi and Papageorgiou (2017), search frictions in the shipping industry allow shipping companies to extract rents from exporters by charging shipping fees that are increasing in the value of the goods in transit.

stage from their least-cost source. The partial equilibrium of this decentralized economy coincides with the solution to the lead-firm problem – and in fact can be recast as a dynamic programming formulation of the lead-firm problem – but it is dramatically simpler to compute. For a chain entailing N stages with each of these stages potentially being performed in one of J countries, characterizing the J optimal GVCs that service consumers in each country requires only $J \times N \times J$ computations, instead of the lead firm having to optimize over J^N potential paths for each of the J locations of consumption (for a total of $J \times J^N$ computations).

Although the results of our partial equilibrium model suggest that more central countries should have comparative advantage in relatively downstream stages within GVCs, formally demonstrating such a result requires developing a general-equilibrium model of GVCs in which production costs are endogenously determined and also shaped by trade barriers. With that goal in mind and also to explore the real income implications of changes in trade costs, in section 3 we develop a Ricardian model of trade in which the combination of labor productivity and trade costs differences across countries shapes the equilibrium position of countries in GVCs. More specifically, we adapt Eaton and Kortum's (2002) Ricardian model to a multi-stage production environment and derive sharp predictions for the *average* participation of countries in different segments of GVCs.

Previous attempts to extend the Ricardian model of trade to a multi-stage, multi-country environment (e.g., Yi, 2003, 2010, Johnson and Moxnes, 2016, Fally and Hillberry, 2016) have focused on the quantification of relatively low-dimensional models with two stages or two countries. Indeed, as we describe in section 3, it is not obvious how to exploit the extreme-value distribution results invoked by Eaton and Kortum (2002) in a multi-stage environment in which cost-minimizing location decisions are a function of the various cost 'draws' obtained by producers worldwide at various stages in the value chain. The reason for this is that neither the sum nor the product of Fréchet random variables are themselves distributed Fréchet, and thus previous approaches have been forced to resort to numerical analyses and simulated method of moments estimation.

We propose two alternative approaches to restore the tractability of Eaton and Kortum (2002) in a Ricardian model with multi-stage production. The first approach consists in simply treating the *overall* unit cost of production of a GVC flowing through a sequence of countries as a draw from a Fréchet random variable with a location parameter that is a function of the states of technology and wage levels of *all* countries involved in that GVC, as well as of the trade costs incurred in that chain. The second approach maintains the standard assumption that labor productivity is stage-specific and drawn from a Fréchet distribution, but instead considers a decentralized equilibrium in which, producers of a particular stage in a GVC have incomplete information about the productivity of certain suppliers upstream from them. More specifically, we assume that firms know their productivity and that of the suppliers immediately upstream from them (i.e., their tier-one suppliers) when they commit to sourcing from a particular supplier, but they do not know the precise productivity of their suppliers' suppliers (i.e., tier-two suppliers, tier-three suppliers, and so on). Interestingly, we find that these two alternative approaches are isomorphic, in the sense that they yield the exact same equilibrium equations.[2]

Under these two alternative assumptions, we show in section 4 that our model generates a closed-form expression for the probability of any potential path of production constituting the cost-minimizing path to service consumers in any country. These probabilities are analogous to the trade shares in Eaton and Kortum (2002), and indeed our model nests their framework in the absence of multi-stage production. Exploiting properties of the resulting distribution of final-good and input prices, we show that our model also delivers closed-form expressions for final-good and input trade flows across countries, which can easily be mapped to the various entries of a world Input-Output table, or WIOT for short. Various versions of these type of world Input-Output tables have become available in recent years, including the World Input Output Database, the OECD's TiVA statistics, and the Eora MRIO database. Our Ricardian multi-stage framework also delivers a simple formula relating real income to the relative prevalence of purely domestic value chains, a formula that generalizes the 'gains from trade' formula in Arkolakis et al. (2012). Although our set of general-equilibrium equations is a bit more cumbersome than in Eaton and Kortum (2002), we show how the proof of existence and uniqueness in Alvarez and Lucas (2007) can be easily (though tediously) adapted to our setting. Finally, we formally establish the existence of a centrality-downstreamness nexus, by which the average downstreamness of a country in GVCs should be increasing in this country's centrality (holding other determinants of comparative advantage constant). After introducing our main data sources, in section 5, we provide suggestive empirical evidence for this centrality-downstreamness nexus and for a key mechanism of the model – namely, the fact that the trade elasticity is larger for downstream stages than for upstream stages.

In section 6, we leverage the tractability of our framework to back out the model's fundamental parameters from data on the various entries of a WIOT. Our empirical approach constitutes a blend of calibration and estimation. First, we show that when abstracting from variation in domestic costs across countries, our equilibrium conditions unveil a simple way to back out the matrix of bilateral trade costs across countries from data on final-good trade flows within and across countries. Our approach is akin to that in Head and Ries (2001), but it requires the use of only final-good trade flows. We also fix a key parameter that governs the shape of the Fréchet distributions of productivity to (roughly) match the aggregate trade elasticity implied by our model. Conditional on a set of countries J and a number of stages N, we then estimate the remaining parameters of the model via a generalized method of moments (GMM), in which we target the diagonal entries of a WIOT.

We perform this exercise for two distinct and complementary samples. First, we use 2014 data from the World Input-Output Database, a source which is deemed to provide high-quality and reliable data on intermediate input and final-good bilateral trade flows across countries for a sample of 43 countries and the rest of the world. The main downside of this database is that the bulk of the countries in the database are high- and medium-income countries in Europe, Asia and North America. In order to study the geography of GVCs worldwide, we also present results using

[2] The approach of building some form of incomplete information (or ex-ante uncertainty) into the Eaton and Kortum (2002) framework is similar in spirit to the one pursued by Tintelnot (2017) and Antràs, Fort and Tintelnot (forthcoming).

the broader sample of 190 countries in the Eora MRIO database. This data source is admittedly less reliable, but it allows us to estimate the model for 101 countries (or consolidated countries) in which all continents and income-levels are more properly represented. In both cases, we find that the model is able to match the targeted moments remarkably well, and it also provides a very good fit for the cells of the WIOT that are not directly targeted in the estimation.

Armed with estimates of the fundamental parameters of the model, we conclude the paper in section 7 by performing counterfactual exercises that illustrate how changes in trade barriers affect the extent to which various countries participate in domestic, regional or global value chains, and traces the real income consequences of these changes. We find that the gains from trade (i.e., the income losses from reverting to autarky) emanating from our model are modestly larger than those obtained from a version of our model without multi-stage production. This variant of our model is a generalization of the Eaton and Kortum (2002) model calibrated to match *exactly* the WIOT. When studying trade costs reductions relative to their calibrated levels, we find much higher income gains, both in absolute terms, but also relative to the version of our model without multi-stage production. These larger gains partly reflect the increased participation of low-income countries in GVCs.

Our paper most closely relates to the burgeoning literature on GVCs. On the theoretical front, in recent years a few theoretical frameworks have been developed highlighting the role of the sequentiality of production for the global sourcing decisions of firms. Among others, this literature includes the work of Harms, Lorz, and Urban (2012), Baldwin and Venables (2013), Costinot *et al.* (2013), Antràs and Chor (2013), Kikuchi *et al.* (2014), Fally and Hillberry (2014), and Tyazhelnikov (2016).[3] A key limitation of this body of theoretical work is that it either completely abstracts from modeling trade costs or it introduces such barriers in highly stylized ways (i.e., assuming common trade costs across all country-pairs). On the empirical front, a growing body of work, starting with the seminal work of Johnson and Noguera (2012), has been concerned with tracing the value-added content of trade flows and using those flows to better document the rise of GVCs and the participation of various countries in this phenomenon (see Koopman et al., 2014, Johnson, 2014, Timmer et al., 2014, de Gortari, 2017).[4] A parallel empirical literature has developed indices of the relative positioning of industries and countries in GVCs (see Fally, 2012, Antràs et al., 2012, Alfaro et al., 2015). On the quantitative side, and as mentioned above, our work builds on and expands on previous work by Yi (2003, 2010), Johnson and Moxnes (2016) and Fally and Hillberry (2016). Other authors, and most notably Caliendo and Parro (2015), have developed quantitative frameworks with Input-Output linkages across countries, but in models with a roundabout production structure without an explicit sequentiality of production. The connection between our framework and these previous contributions is further explored in de Gortari (2017), who blends several strands of this literature by generalizing the formulas on value-added content and

[3] This literature is in turn inspired by earlier contributions to modeling multi-stage production, such as Dixit and Grossman (1982), Sanyal and Jones (1982), Kremer (1993), Yi (2003) and Kohler (2004).

[4] An important precursor to this literature is Hummels et al. (2001), who combined international trade and Input-Output data to construct indices of vertical specialization.

downstreamness within the context of a multi-sector Ricardian model with sequential production. Finally, some implications of the rise of offshoring and GVCs for trade policy have been studied by Antràs and Staiger (2012) and Bown et al. (2016), but in much more stylized frameworks than studied in this paper.

The rest of the paper is structure as follows. Section 2 develops our partial equilibrium model and highlights some of its key features. Section 3 describes the assumptions of the general equilibrium model, and section 4 characterizes its equilibrium. Section 5 introduces our data sources and provides suggestive empirical evidence for some of the key features of our model. Section 6 covers the estimation of our model and section 7 explores several counterfactuals. All proofs and several details on data sources and the estimation are relegated to the Appendix and Online Appendix.

2 Partial Equilibrium: Interdependencies and Compounding

In this section, we develop a simple model of firm behavior that formalizes the problem faced by a firm choosing the location of its various production stages in an environment with costly trade. For the time being, we consider the problem of a firm (or, more precisely, of a competitive fringe of firms) producing a particular good. We defer a discussion of the general equilibrium aspects of the model to section 3.

2.1 Environment

There are J countries in which consumers derive utility from consuming a final good. The good is produced combining N stages that need to be performed sequentially. The last stage of production can be interpreted as assembly and is indexed by N. We will often denote the set of countries $\{1, ..., J\}$ by $\mathcal{J}$ and the set of production stages $\{1, ..., N\}$ by $\mathcal{N}$. At each stage $n > 1$, production combines a local composite factor (which encompasses primitive factors of production and a bundle of materials), with the good finished up to the previous stage $n-1$. Production in the initial stage $n = 1$ only uses the composite factor. The cost of the composite factor varies across countries and is denoted by c_i in country i.[5] Countries also differ in their geography, as captured by a $J \times J$ matrix of iceberg trade coefficients $\tau_{ij} \geq 1$, where τ_{ij} denotes the units of the finished or unfinished good that need to be shipped from i for one unit to reach j. Firms are perfectly competitive and the optimal location $\ell(n) \in \mathcal{J}$ of the different stages $n \in \mathcal{N}$ of the value chain is dictated by cost minimization. Because of marginal-cost pricing, we will somewhat abuse notation and denote by $p^n_{\ell(n)}$ the unit cost of production of a good completed up to stage n in country $\ell(n)$. That good is available in country $\ell(n+1)$ at a cost $p^n_{\ell(n)}\tau_{\ell(n)\ell(n+1)}$.

We summarize technology via the following sequential cost function associated with a path of

[5] For now we take this cost as given, but in the general equilibrium analysis in section 3, we will break c_i into the cost of labor and of a bundle of intermediate inputs we call *materials*. This will allow our model to encompass previous Ricardian models – and most notably Eaton and Kortum (2002) – featuring roundabout production.

production $\boldsymbol{\ell} = \{\ell(1), \ell(2), ..., \ell(N)\}$:

$$p^n_{\ell(n)}(\boldsymbol{\ell}) = g^n_{\ell(n)}\left(c_{\ell(n)}, p^{n-1}_{\ell(n-1)}(\boldsymbol{\ell})\,\tau_{\ell(n-1)\ell(n)}\right), \text{ for all } n \in \mathcal{N}. \tag{1}$$

The stage- and country-specific cost functions $g^n_{\ell(n)}$ in equation (1) are assumed to feature constant-returns-to-scale and diminishing marginal products. As mentioned before, we let the cost of the first stage depend only on the local composite factor, so constant returns to scale implies $p^1_{\ell(1)}(\boldsymbol{\ell}) = g^1_{\ell(1)}(c_{\ell(1)})$ for all paths $\boldsymbol{\ell}$, with the function $g^1_{\ell(1)}$ necessarily being linear in $c_{\ell(1)}$.

Note that equation (1) also applies to the assembly stage N, and a good assembled in $\ell(N)$ after following the path $\boldsymbol{\ell}$ is available in any country j at a cost $p^F_j(\boldsymbol{\ell}) = p^N_{\ell(N)}(\boldsymbol{\ell})\,\tau_{\ell(N)j}$ (we use the superscript F to denote finished goods). For each country $j \in \mathcal{J}$, the goal is then to choose the optimal path of production $\boldsymbol{\ell}^j = \left\{\ell^j(1), \ell^j(2), ..., \ell^j(N)\right\} \in \mathcal{J}^N$ that minimizes the cost $p^F_j(\boldsymbol{\ell})$ of providing the good to consumers in that country j.

At various points in the paper, we will find it useful to focus on the case in which cross-country differences in technology are associated with Ricardian differences in the efficiency with which the local composite factor is used in different stages, and in which the function $g^n_{\ell(n)}$ is a Cobb-Douglas aggregator of the composite factor and the product finished up to the previous stage. More specifically, we write

$$p^n_{\ell(n)}(\boldsymbol{\ell}) = \left(a^n_{\ell(n)} c_{\ell(n)}\right)^{\alpha_n} \left(p^{n-1}_{\ell(n-1)}(\boldsymbol{\ell})\,\tau_{\ell(n-1)\ell(n)}\right)^{1-\alpha_n}, \text{ for all } n \in \mathcal{N}, \tag{2}$$

where α_n denotes the cost share of the composite factor at stage n and $a^n_{\ell(n)}$ is the unit factor requirement at stage n in country $\ell(n)$. Because the initial stage of production uses solely the local composite factor, we have $\alpha_1 = 1$.

2.2 *Lead-Firm* Problem

We consider first the problem of a *lead firm* choosing the location of production of all stages $n \in \mathcal{N}$, in order to minimize the overall cost of serving consumers in a given country j. Using $p^F_j(\boldsymbol{\ell}) = p^N_{\ell(N)}(\boldsymbol{\ell})\,\tau_{\ell(N)j}$ and iterating (2), this problem reduces to:

$$\boldsymbol{\ell}^j = \arg\min_{\boldsymbol{\ell}\in\mathcal{J}^N} p^F_j(\boldsymbol{\ell}) = \arg\min_{\boldsymbol{\ell}\in\mathcal{J}^N} \left\{\prod_{n=1}^{N}\left(a^n_{\ell(n)} c_{\ell(n)}\right)^{\alpha_n\beta_n} \times \prod_{n=1}^{N-1}\left(\tau_{\ell(n)\ell(n+1)}\right)^{\beta_n} \times \tau_{\ell(N)j}\right\} \tag{3}$$

where

$$\beta_n \equiv \prod_{m=n+1}^{N}(1-\alpha_m), \tag{4}$$

and where we use the convention $\prod_{m=N+1}^{N}(1-\alpha_m) = 1$. Note that $\sum_{n=1}^{N}\alpha_n\beta_n = 1$.

We next highlight two important features of program (3). First, notice that when trade costs are identical for all country-pairs (i.e., $\tau_{ij} = \tau$ for all i and j), the last two terms reduce to a constant that is independent of the path of production. In such a case, we can break the cost-minimization problem in (3) into a sequence of N independent cost-minimization problems in which the optimal location of stage n is simply given by $\ell^{j}(n) = \arg\min_{i}\{a_{i}^{n}c_{i}\}$, and is thus independent of the country of consumption j. Notice, however that this result requires no differences between internal and external trade costs (i.e., $\tau_{ij} = \tau$ also for $i = j$), and thus this case is isomorphic, up to a productivity shifter, to an environment with costless trade. With a general geography of trade costs, a lead firm can no longer perform cost minimization independently stage-by-stage, and instead it needs to optimize over the whole path of production. Intuitively, the location $\ell(n)$ minimizing production costs $a_{\ell(n)}^{n}c_{\ell(n)}$ might not be part of a firm's optimal path if the optimal locations for stages $n-1$ and $n+1$ are sufficiently far from $\ell(n)$. A direct implication of this result is that the presence of arbitrary trade costs turns a problem of dimensionality $N \times J$ into J much more complex problems of dimensionality J^{N} each. As we will see below, however, the dimensionality of program (3) can be dramatically reduced using dynamic programming.

A second noteworthy aspect of the minimand in equation (3) is that the trade-cost elasticity of the unit cost of serving consumers in country j increases along the value chain. More specifically, note from equation (4) that, as long as $\alpha_{n} > 0$ for all n, we have $\beta_{1} < \beta_{2} < ... < \beta_{N} = 1$. For the particular case in which overall value added is a symmetric Cobb-Douglas aggregator of the value added of all stages (i.e., $\alpha_{n}\beta_{n} = 1/N$, for all n), the program in (3) reduces to

$$\boldsymbol{\ell}^{j} = \arg\min_{\boldsymbol{\ell}\in\mathcal{J}^{N}} p_{j}^{F}(\boldsymbol{\ell}) = \arg\min_{\boldsymbol{\ell}\in\mathcal{J}^{N}}\left\{\prod_{n=1}^{N}\left(a_{\ell(n)}^{n}c_{\ell(n)}\right)^{1/N}\times\prod_{n=1}^{N-1}\left(\tau_{\ell(n)\ell(n+1)}\right)^{n/N}\times\tau_{\ell(N)j}\right\}, \quad (5)$$

and the trade-cost elasticity increases linearly with the downstreamness n of a stage.

The reason for this compounding effect of trade costs stems from the fact that the costs of transporting goods have been modeled (realistically, as we argued in the Introduction) to be proportional to the gross value of the good being transacted, rather than being assumed proportional to the value added at that stage. Thus, as the value of the good rises along the value chain, so does the amount of resources used to transport the goods across locations. An implication of this compounding effect is that, in choosing their optimal path of production, firms will be relatively more concerned about reducing trade costs in relatively downstream stages than in relatively upstream stages. As we will illustrate below and formally demonstrate when exploring the general equilibrium of our model, this feature of the cost function will generate a centrality-downstreamness nexus by which, *ceteris paribus*, relatively more central countries will tend to gain comparative advantage and specialize in relatively downstream stages.[6]

Although we have derived this compounding effect of trade costs for the case of Ricardian technological differences and Cobb-Douglas cost functions, we show in Appendix A.1 that the same result applies for arbitrary constant-returns-to-scale technologies of the type in equation (1). More

[6] Building on the results in Costinot (2012), we can briefly anticipate this result with the following example. Suppose that trade costs can be decomposed as $\tau_{ij} = \left(\rho_{i}\rho_{j}\right)^{-1}$, where we take ρ_{i} is an index of the *centrality* of country i. In such a case, it is straightforward to show that leaving aside other determinants of comparative advantage, the unit cost of servicing consumers in country j is log-supermodular in a country's centrality ρ_{i} and a stage's downstreamness n (see more on this in section 4).

specifically, denoting by β_n the elasticity of $p_j^F(\boldsymbol{\ell})$ with respect to $\tau_{\ell(n)\ell(n+1)}$, we show that β_n is again necessarily non-decreasing in n even when these elasticities are not pinned down by exogenous parameters. Thus, the result that firms will be particularly concerned about minimizing trade costs in downstream stages is quite general.[7]

2.3 Decentralization and Dynamic Programming

We have so far characterized the problem of a *lead firm* with full information on the productivity of the various potential producers of each stage n in each country j. This characterization relies on strong informational assumptions, so we now consider an alternative environment in which no individual firm coordinates the whole value chain. Instead, we assume that a value chain consists of a series of stage-specific producers that simply minimize their cost of production taking into account their composite factor cost, their productivity, and the cost at which they can obtain the good finished up to the immediately preceding stage. Similarly, consumers in country j simply purchase the final good from whichever assembler (i.e., stage N producer) worldwide can provide the finished good at the lowest price.

From equation (1), a producer of stage n in country $\ell(n)$ would choose to procure the good finished up to stage $n-1$ by simply solving $\min_{\ell(n-1)\in\mathcal{J}}\left\{p_{\ell(n-1)}^{n-1}\tau_{\ell(n-1)\ell(n)}\right\}$, where $p_{\ell(n-1)}^{n-1}$ is the optimal (free-on-board) price charged by producers of stage $n-1$ in country $\ell(n-1)$. Importantly, producers seek to simply minimize sourcing costs regardless of their own composite factor cost, their productivity and the future path of the good after flowing through $\ell(n)$ at stage n. Furthermore, the resulting price at which this producer can sell the good finished up to stage n to producers of stage $n+1$ is only a function of $a_{\ell(n)}^n c_{\ell(n)}$ and this minimum price $\min_{\ell(n-1)\in\mathcal{J}}\left\{p_{\ell(n-1)}^{n-1}\tau_{\ell(n-1)\ell(n)}\right\}$. Producers of the initial stage $n=1$ only use their local composite factor, and thus $p_{\ell(1)}^1 = a_{\ell(1)}^1 c_{\ell(1)}$.

With constant returns to scale, the identity of the specific firms making these decisions is of course immaterial, so this formulation is entirely consistent with our previous *lead firm* using dynamic programming to solve for the optimal path of production leading to consumption in each country $j\in\mathcal{J}$. More specifically, instead of solving program (3) in a brute force manner, the lead firm can break the problem into a series of stage- and country-specific optimal sourcing problems (as in the decentralized formulation above), and then solve the problem via forward induction (starting in the most upstream stages). Invoking the principle of optimality, we can then establish (see Appendix A.2) that the resulting optimal path of production $\boldsymbol{\ell}^j = \left\{\ell^j(1), \ell^j(2), ..., \ell^j(N)\right\}\in\mathcal{J}^N$ that minimizes the cost $p_j^F(\boldsymbol{\ell})$ in this decentralized formulation of the problem will coincide with the one we obtained solving the *lead-firm* problem in (3) by exhaustive search.

A key advantage of this dynamic programming approach is that it only requires $J\times N\times J$ computations to obtain the optimal production path for *all* destinations of final consumption,

[7] For example, for the case of a symmetric Leontief technology and production costs equal to 1 in all countries and stages (i.e., $a_{\ell(n)}^n c_{\ell(n)} = 1$ for all n and $\ell(n)$), we obtain

$$p_j^F(\boldsymbol{\ell}) = \tau_{\ell(N)j} + \tau_{\ell(N)j}\tau_{\ell(N-1)\ell(N)} + \tau_{\ell(N)j}\tau_{\ell(N-1)\ell(N)}\tau_{\ell(N-2)\ell(N-1)} + \tau_{\ell(N)j}\tau_{\ell(N-1)\ell(N)}\tau_{\ell(N-2)\ell(N-1)}\tau_{\ell(N-3)\ell(N-2)} + \cdots,$$

which again illustrates the larger relative importance of downstream trade costs.

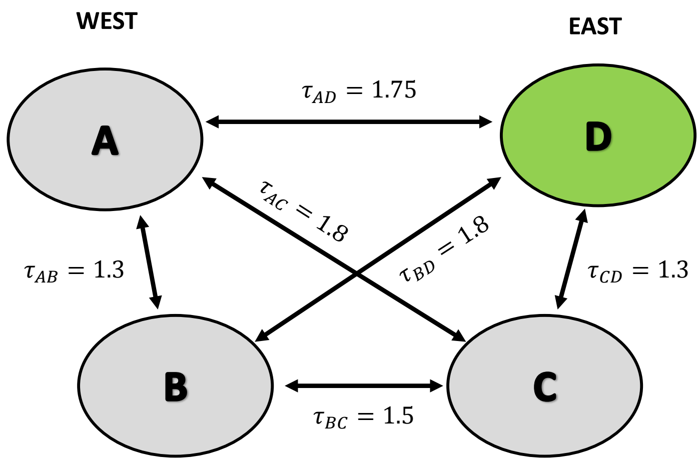

Figure 1: An Example with Four Countries

instead of having to optimize over J^N potential paths for each country j.[8] For example, with 200 countries and 5 stages, this amounts to only 200,000 computations rather than 64 trillion computations.[9] Although it might be clear from our discussion above, it is worth stressing that the isomorphism between the lead-firm problem and the decentralized problem holds true for any constant-returns-to-scale technology, and not only for the Cobb-Douglas one in (2).

2.4 An Example

We close this section by illustrating some of the salient and distinctive features of this partial model of sequential production via a simple example. We consider a world with four countries ($J = 4$) and four stages ($N = 4$). Technology is given by the symmetric Cobb-Douglas specification in (5), with $\alpha_n \beta_n = 1/4$ for all n. The four countries are divided into two regions, the West (comprising countries A and B) and the East (comprising countries C and D). The 'geography' of this example is illustrated in Figure 1. Note that we impose a great deal of symmetry: intra-regional trade costs are common in both regions, and inter-regional costs between A and C are identical to those between B and D. On the other hand, trade costs between B and C are lower than between A and D. For simplicity, all domestic trade costs are set to 0, so $\tau_{ii} = 1$ for $i = A, B, C, D$. We are interested in solving for the optimal path of a four-stage production process leading to consumption in country D (in green in the figure). Note that shipping to D directly is least costly when shipping from D itself, followed by C (the other country in the East), then by A and finally by B, which is the most remote country relative to D.

We compute the optimal path leading to D for different levels of trade costs starting with a benchmark with $\tau_{AB} = \tau_{CD} = 1.3$, $\tau_{BC} = 1.5$, $\tau_{AD} = 1.75$, $\tau_{AC} = \tau_{BD} = 1.8$, and then

[8] See Appendix A.2 for more details. This same point has been made in contemporaneous work by Tyazhelnikov (2016).

[9] Though the dimensionality of the lead firm's problem is huge, for the particular case with Cobb-Douglas technologies, in Appendix A.2 we show that the problem can also be written as a zero-one integer programming problem, for which many extremely quick and efficient algorithms are available (see, for instance, http://www.gurobi.com).

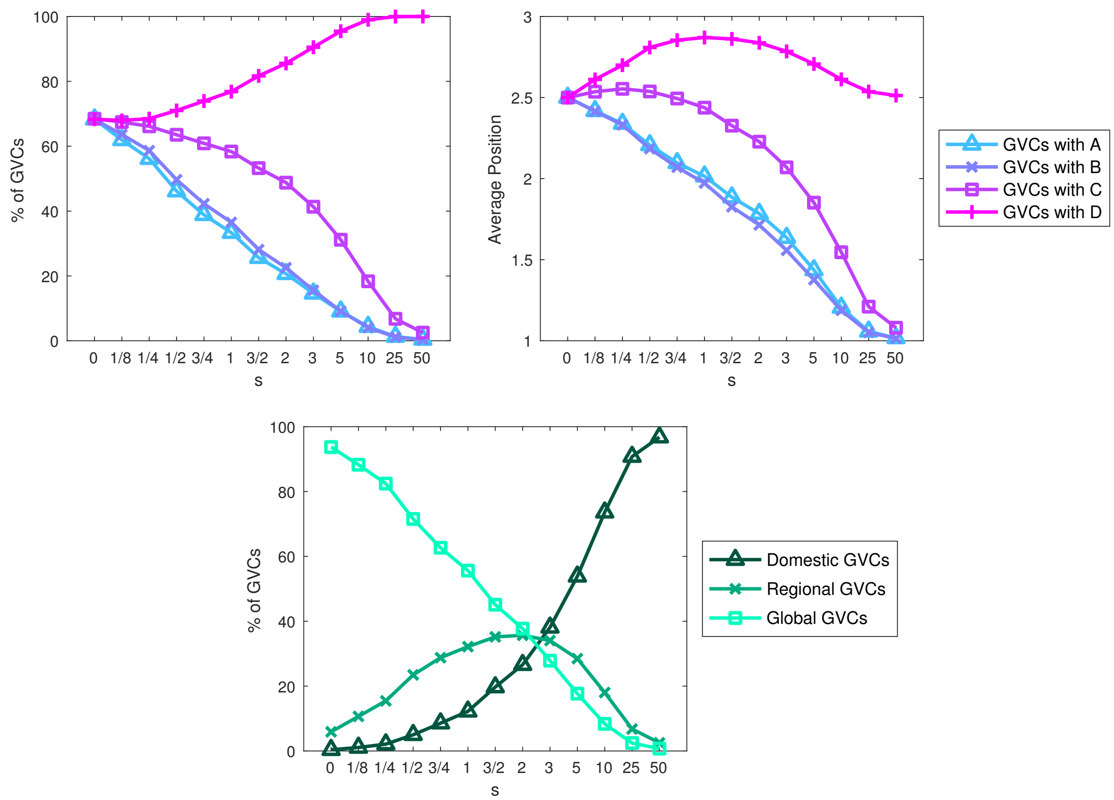

Figure 2: Some Features of Optimal Production Paths

scale these international trade costs up or down by a shifter s (so starting from τ_{ij}, we instead use $\tilde{\tau}_{ij}(s) = 1 + s \times (\tau_{ij} - 1)$).[10] For each matrix of trade costs, we run one million simulations with production costs $a_j^n c_j$ being drawn independently for each stage n and each country j from a lognormal distribution with mean 0 and variance 1. By choosing a common distribution across countries and stages, we seek to isolate the role of trade costs in shaping the optimal path of sequential value chains.

The results of these simulations are depicted in Figure 2 for various levels of s ranging from 0 (free trade) to 50 (which results in close to prohibitive trade costs). The upper left panel shows the average propensity of each country to appear in GVCs leading to consumption in D. The upper right panel depicts the average position (or downstreamness) of countries in these GVCs. Finally, the lower panel decomposes GVCs into purely domestic ones (with all production stages in D), purely regional ones (with some stages in C and D, but not in A or B) and global ones (involving at least one stage in A or B).

Several aspects of Figure 2 are worth highlighting. First, focusing on the upper left panel, notice that country B, which is farthest away from country D, appears slightly more often in value chains leading to D than its Western neighbor A does. The reason for this surprising fact is tightly related to the sequential nature of production. Even though, A is closer to D than B is, B is

[10]These parameters are chosen such that for all values of s considered, the triangle inequality holds for any three given countries.

relatively close to D's Eastern neighbor C, and this makes this 'remote' country B a particularly appealing location from which to set off value chains that will flow to D through C.[11] A second noteworthy aspect, apparent from the upper right panel of Figure 2, is that remoteness appears to shape the average position of a country in GVCs, a fact we anticipated above. More specifically, country B, which is farthest away from D, is on average the most upstream of all countries, followed by its Western neighbor A, and then by C, with D being naturally the country positioned most downstream in value chains leading to consumption in D. Finally, the lower panel of Figure 2 illustrates how the progressive reduction of international trade costs first gives rise to GVCs that are largely regional in nature, and then later to truly global value chains involving inter-regional trade. It is also worth highlighting that even for fairly low trade costs, purely domestic GVCs remain quite prevalent, much more so than would be predicted by an analogous model without sequentiality (see the Online Appendix B.1). The reason for this is the compounding effect of trade costs, which other things equal makes it costly to offshore intermediate production stages in chains in which D has comparative advantage in the most upstream *and* downstream stages.

3 General Equilibrium Model

We next embed the model of firm behavior developed in the last section into a full-fledged general equilibrium model.

3.1 Environment

We continue to assume a world with J countries (indexed by i or j) where consumers now derive utility from consuming a continuum of final-good varieties (indexed by z). Preferences are CES and given by

$$u\left(\left\{y_i^N(z)\right\}_{z=0}^1\right) = \left(\int_0^1 \left(y_i^N(z)\right)^{(\sigma-1)/\sigma} dz\right)^{\sigma/(\sigma-1)}, \quad \sigma > 1. \tag{6}$$

Production of each of the final-good varieties is as described in the previous section: production processes entail N sequential stages (indexed by n) and are characterized by the Ricardian, Cobb-Douglas specification in (2).

We let countries differ in three key aspects: (i) their technological efficiency, as determined by the unit composite factor requirements $a_i^n(z)$, (ii) their geography, as captured by a $J \times J$ matrix of iceberg trade cost $\tau_{ij} \geq 1$, and (iii) their size, as reflected by the measure L_i of 'equipped' labor available for production in each country i (labor is inelastically supplied and commands a wage w_i).

The local composite factor used at each stage comprises labor and an aggregator of final-good varieties that corresponds exactly to the CES aggregator in (6). In other words, part of final-good

[11] As we show in Online Appendix B.1, in an analogous world without sequentiality, the above pattern would not hold and the relative prevalence of countries would be strictly monotonic in the level trade costs incurred when shipping to the assembly location.

production is not absorbed by consumers, but rather by firms that use those goods as a bundle of materials. This roundabout structure of production is standard in recent Ricardian models (see Eaton and Kortum, 2002, Alvarez and Lucas, 2007, or Caliendo and Parro, 2014), so we adopt it for comparability (see, in particular, section 7). We should stress, however, that our model features intermediate input flows across countries even in the absence of these production 'loops'. We let the cost c_i of the composite factor in country i be captured by a Cobb-Douglas aggregator $c_i = (w_i)^{\gamma_i} (P_i)^{1-\gamma_i}$, where P_i is the ideal price index associated with the CES aggregator in (6). Although allowing for variation in value added shares γ_i across countries is not important for our theoretical results, it will prove useful in allowing our model to better match world Input-Output tables.

This completes the discussion of the structure of our general-equilibrium model. In principle, given values for the unit composite factor requirements $a_i^n(z)$ and all other primitive parameters, the equilibrium of the model could be computed by (i) solving for the cost-minimizing path of production for each good z and each destination of consumption j given a vector of wages, and (ii) invoking labor-market clearing to reduce equilibrium wages to the solution of a fixed point problem. Such an approach, however, would not be particularly useful in order to formally characterize certain features of the equilibrium or to estimate the model in a computationally feasible and transparent manner. With that in mind, we next explore a particularly convenient parametrization of the unit factor requirements $a_i^n(z)$.

3.2 Technology

Building on the seminal work of Eaton and Kortum (2002), we propose a probabilistic specification of the unit factor requirements $a_i^n(z)$ that delivers a remarkably tractable multi-stage, multi-country Ricardian model. We are certainly not the first ones to explore such a multi-stage extension of the Eaton and Kortum (2002) framework. Yi (2010) and Johnson and Moxnes (2016), for instance, consider a 'natural' extension in which each productivity parameter $1/a_i^n(z)$ is assumed stochastic and drawn independently (across goods and stages) from a type II (or Fréchet) extreme-value probability distribution, as in Eaton and Kortum (2002). A key limitation of their approach is that the minimum cost associated with a given GVC path is not characterized by a particularly tractable distribution. The reason for this is that, although the minimum of a series of Fréchet draws is itself distributed Fréchet, the product of Fréchet random variables is *not* distributed Fréchet.[12] As a result, these papers need to resort to numerical methods to approximate the solution of their models, even when restricting the analysis to two-stage chains. We instead develop two alternative approaches that will permit a sharp and exact characterization of some of the features of the equilibrium for an arbitrary number of stages, and that will be readily amenable to structural (generalized method of moments) estimation using world Input-Output tables.

[12] Assuming a linear cost function (i.e., perfect complementarity) does not provide tractability either because the sum of Fréchet random variables is not distributed Fréchet either.

A. *Lead-Firm* Approach

We begin by revisiting the problem of a lead firm choosing the location of the various stages of production with full knowledge of the realized unit requirements $a_i^n(z)$ for each stage in each country. The key innovation we propose, relative to Eaton and Kortum (2002), is to introduce randomness to the overall cost of production of a given value chain, rather than to the productivity of each stage independently. Intuitively, a given production path $\boldsymbol{\ell} = \{\ell(1), \ell(2), ..., \ell(N)\} \in \mathcal{J}^N$ will be associated with an *average* cost that is naturally a function of trade costs, composite factor costs and the state of technology of the various countries involved in the chain. Yet, compatibility problems, production delays, or simple mistakes can generate idiosyncratic noise around that average. More formally, and building on the cost function in (3), we assume that the overall 'productivity' of a given chain $\boldsymbol{\ell}$ is characterized by

$$\Pr\left(\prod_{n=1}^{N}\left(a_{\ell(n)}^{n}(z)\right)^{\alpha_n\beta_n} \geq a\right) = \exp\left\{-a^{\theta}\prod_{n=1}^{N}\left(T_{\ell(n)}\right)^{\alpha_n\beta_n}\right\}, \tag{7}$$

which amounts to assuming that $\prod_{n=1}^{N}\left(a_{\ell(n)}^{n}(z)\right)^{-\alpha_n\beta_n}$ is distributed Fréchet with a shape parameter given by θ, and a location parameter that is a function of the states of technology in all countries in the chain, as captured by $\prod_{n=1}^{N}\left(T_{\ell(n)}\right)^{\alpha_n\beta_n}$. A direct implication of this assumption is that the unit cost associated with serving consumers in a given country j via a given chain $\boldsymbol{\ell}$ is also distributed Fréchet. More precisely, denoting by $p_j^F(\boldsymbol{\ell}, z)$ the price paid by consumers in j for a good z produced following the path $\boldsymbol{\ell}$, we have

$$\Pr\left(p_j^F(\boldsymbol{\ell}, z) \geq p\right) = \exp\left\{-p^{\theta} \times \prod_{n=1}^{N}\left(\left(c_{\ell(n)}\right)^{-\theta} T_{\ell(n)}\right)^{\alpha_n\beta_n} \times \prod_{n=1}^{N-1}\left(\tau_{\ell(n)\ell(n+1)}\right)^{-\theta\beta_n} \times \left(\tau_{\ell(N)j}\right)^{-\theta}\right\}, \tag{8}$$

independently of the final good z under consideration. This result will be key for neatly characterizing the equilibrium, as we will show in the next section.

B. Decentralized Approach

We also develop an alternative approach closer in spirit to the stage-specific productivity randomness in Yi (2010) and Johnson and Moxnes (2016), which also achieves tractability and, in fact, delivers an identical set of equilibrium conditions to those we will derive under the specification in (7). On the technology side, we now assume that $1/a_i^n(z)$ is drawn independently (across goods and stages) from a Fréchet distribution satisfying

$$\Pr\left(a_i^n(z)^{\alpha_n\beta_n} \geq a\right) = \exp\left\{-a^{\theta}\left(T_i\right)^{\alpha_n\beta_n}\right\}. \tag{9}$$

Note that this formulation imposes a common variance (as captured by the shape parameter θ) of the contribution of each stage n's productivity – i.e., $a_i^n(z)^{-\alpha_n\beta_n}$ – to the overall productivity of a value chain. This assumption ensures that we do not mechanically introduce heterogeneity across stages in the trade-cost elasticity related to the importance of these stages in production.[13]

In order to make this alternative approach tractable, we relax the assumption that firms choose the optimal path of production with full knowledge of the productivity levels with which all stages of production in their chain could be produced in different countries. More specifically, we explore an environment akin to the decentralized equilibrium developed in section 2.3, in which stage-specific producers simply attempt to minimize the cost of production of their stage. Unlike in section 2.3, we assume, however, that these stage-specific producers do not observe realized upstream prices before making sourcing decisions, and can only forecast these prices based on information on the productivity levels of their potential direct (or tier-one) suppliers in various countries. These tier-one supplier productivity levels are not sufficient statistics for sourcing prices because upstream marginal costs also depend on the productivity of suppliers further upstream (i.e., tier-two suppliers, tier-three suppliers and so on). The idea behind this formulation is that firms need to pre-commit to purchase from particular suppliers based on information they gather from inspecting (e.g., through factory visits) all their potential immediate suppliers. Ex-post, a supplier's marginal cost might be higher or lower than expected because this supplier may face unexpectedly high or low sourcing costs itself. Although the pre-commitment to buy from a particular source naturally affects the nature of ex-post competition, we assume that buyers have all the bargaining power and continue to be able to source upstream inputs at marginal cost.

Because this decentralized approach with incomplete information is a bit more cumbersome than the formulation in (7), we only illustrate how it works for the simple case with only two stages, input production (stage 1) and assembly (stage 2). In Appendix A.3, we show how the approach naturally generalizes to the case $N > 2$.

Input producers of a given good z in a given country $\ell(1) \in \mathcal{J}$ observe their productivity $1/a^1_{\ell(1)}(z)$, and simply hire labor and buy materials to minimize unit production costs, which results in $p^1_{\ell(1)}(z) = a^1_{\ell(1)}(z)\, c_{\ell(1)}$. Assemblers of good z in any country $\ell(2) \in \mathcal{J}$ observe their own productivity $1/a^2_{\ell(2)}(z)$, as well as that of all potential input producers worldwide, and solve

$$p^2_{\ell(2)}(z) = \min_{\ell(1)\in\mathcal{J}} \left\{ \left(a^2_{\ell(2)}(z)\, c_{\ell(2)}\right)^{\alpha_2} \left(a^1_{\ell(1)}(z)\, c_{\ell(1)} \tau_{\ell(1)\ell(2)}\right)^{1-\alpha_2} \right\}.$$

Independently of the values of $a^2_{\ell(2)}(z)$, $c_{\ell(2)}$, and α_2, the solution of this problem simply entails procuring the input from the location $\ell^*(1)$ satisfying $\ell^*(1) = \arg\min \left\{ \left(a^1_{\ell(1)}(z)\, c_{\ell(1)} \tau_{\ell(1)\ell(2)}\right)^{1-\alpha_2} \right\}$. As is well-known, the Fréchet assumption in (9) will make characterizing this problem fairly straightforward. Consider finally the problem of retailers in each country j seeking to procure a final good z to local consumers at a minimum cost. These retailers observe the productivity $1/a^2_{\ell(2)}(z)$ of all

[13] There is no economic reason to think that that the trade-cost elasticity should vary with the contribution of a stage to value added, and such variation would obviously obfuscate our result showing that this elasticity rises along the value chain.

assemblers worldwide, but *not* the productivity of input producers, and thus seek to solve

$$p_j^F(z) = \min_{\ell(2)\in\mathcal{J}} \left\{ \left(a_{\ell(2)}^2(z)\, c_{\ell(2)} \right)^{\alpha_2} \mathbb{E}\left[a_{\ell^*(1)}^1(z)\, c_{\ell^*(1)} \tau_{\ell^*(1)\ell(2)} \right]^{1-\alpha_2} \tau_{\ell(2)j} \right\}. \tag{10}$$

If retailers could observe the particular realizations of input producers, the expectation in (10) would be replaced by the realization of $a_{\ell(1)}^1(z)\, c_{\ell(1)} \tau_{\ell(1)\ell(2)}$ in all $\ell(1) \in \mathcal{J}$, and characterizing the optimal choice would be complicated because it would depend on the joint distribution of $a_{\ell(2)}^2(z)$ and $a_{\ell(1)}^1(z)$,which is not Fréchet under (9). As we will demonstrate in section 4, with our incomplete information assumption, the expectation in (10) does not depend on the particular realizations of upstream productivity draws, and this will allow us to apply the well-know properties of the univariate Fréchet distribution in (9) to characterize the problem of retailers.

4 Characterization of the Equilibrium

In this section, we characterize the general equilibrium of our model. We proceed in five steps. First, we leverage our extreme-value representation of GVC productivity to obtain closed-form expressions for the relative prevalence (in value terms) of different GVCs in the world equilibrium. Second, we show how to manipulate these relative market shares of different GVCs to obtain expressions for bilateral intermediate input and final-good flows across countries, which can be mapped to observable data from world Input-Output tables. Third, we study the existence and uniqueness of the general equilibrium. Fourth, we obtain expressions for the gains from trade in our model and compare them to those in Eaton and Kortum (2002). Fifth, we formalize the link between downstreamness and centrality that we hinted at in section 2.

4.1 Relative Prevalence of Different GVCs and Equilibrium Prices

Let us begin with the lead-firm version of our model, in which the price paid by consumers in j for a good produced following the path $\boldsymbol{\ell} \in \mathcal{J}^N$ is given by the Fréchet distribution in (8). In such a case, we can readily invoke a few of the results in Eaton and Kortum (2002) to characterize the equilibrium prices and the relative prevalence of different GVCs. First, it is straightforward to verify that the probability of a given GVC $\boldsymbol{\ell}$ being the cost-minimizing production path for serving consumers in j is given by

$$\pi_{\boldsymbol{\ell}j} = \frac{\prod_{n=1}^{N-1} \left(\left(T_{\ell(n)}\right)^{\alpha_n} \left(\left(c_{\ell(n)}\right)^{\alpha_n} \tau_{\ell(n)\ell(n+1)} \right)^{-\theta} \right)^{\beta_n} \times \left(T_{\ell(N)}\right)^{\alpha_N} \left(\left(c_{\ell(N)}\right)^{\alpha_N} \tau_{\ell(N)j} \right)^{-\theta}}{\Theta_j}, \tag{11}$$

where

$$\Theta_j = \sum_{\boldsymbol{\ell}\in\mathcal{J}^N} \prod_{n=1}^{N-1} \left(\left(T_{\ell(n)}\right)^{\alpha_n} \left(\left(c_{\ell(n)}\right)^{\alpha_n} \tau_{\ell(n)\ell(n+1)} \right)^{-\theta} \right)^{\beta_n} \times \left(T_{\ell(N)}\right)^{\alpha_N} \left(\left(c_{\ell(N)}\right)^{\alpha_N} \tau_{\ell(N)j} \right)^{-\theta}, \tag{12}$$

and where remember that $c_i = (w_i)^{\gamma_i} (P_i)^{1-\gamma_i}$. With a unit measure of final goods, $\pi_{\boldsymbol{\ell} j}$ also corresponds to the share of GVCs ending in j for which $\boldsymbol{\ell}$ is the cost-minimizing production path.[14]

Second, and as in Eaton and Kortum (2002), the distribution of final-good prices $p_j^F(\boldsymbol{\ell}, z)$ paid by consumers in j satisfies

$$\Pr\left(p_j^F(\boldsymbol{\ell}, z) \leq p\right) = 1 - \exp\left\{-\Theta_j p^\theta\right\}. \tag{13}$$

Because the distribution of final-good prices in j is independent of the path of production $\boldsymbol{\ell}$, it follows that the probabilities in $\pi_{\boldsymbol{\ell} j}$ also constitute the shares of country j's income spent on final goods produced under all possible paths $\boldsymbol{\ell} \in \mathcal{J}^N$.

As is clear from equation (11), GVCs that involve countries with higher states of technology T_i or lower composite factor costs c_i will tend to feature disproportionately in production paths leading to consumption in j. Furthermore, and consistently with our discussion in section 2, high trade costs penalize the participation of countries in GVCs, but such an effect is disproportionately large for downstream stages relative to upstream stages. This is captured by the fact that the 'trade elasticity' associated with stage n is given by $\theta\beta_n$, and β_n is increasing in n with $\beta_N = 1$.

Following the same steps as in Eaton and Kortum (2002), we can further solve for the exact ideal price index P_j in country j associated with (6)

$$P_j = \kappa (\Theta_j)^{-1/\theta}, \tag{14}$$

where $\kappa = \left[\Gamma\left(\frac{\theta+1-\sigma}{\theta}\right)\right]^{1/(1-\sigma)}$ and Γ is the gamma function. For the price index to be well defined, we impose $\sigma - 1 < \theta$.

So far we have focused on the 'randomness-in-the-chain' formulation in (8). Consider now our alternative approach with stage-specific randomness captured by (9) and incomplete information. As in section 3, we will focus here on the case with two stages and leave the more general case to Appendix A.3. Take two countries $\ell(1)$ and $\ell(2)$ and consider the probability $\pi_{\boldsymbol{\ell} j}$ of a GVC flowing through $\ell(1)$ and $\ell(2)$ before reaching consumers in j. This probability is simply the product of (i) the probability of $\ell(1)$ being the cost-minimizing location of input production conditional on assembly happening in $\ell(2)$, and (ii) the probability of $\ell(2)$ being the cost-minimizing location of assembly for GVC serving consumers in j. Denoting $\mathcal{E}_{\ell(2)} = \mathbb{E}\left[\tau_{\ell^*(1)\ell(2)} a^1_{\ell^*(1)}(z)\, c_{\ell^*(1)}\right]^{1-\alpha_2}$, and using the properties of the Fréchet distribution, it is easy to verify that we can write $\pi_{\boldsymbol{\ell} j}$ as

$$\pi_{\boldsymbol{\ell} j} = \underbrace{\frac{\left(T_{\ell(1)}\right)^{1-\alpha_2}\left(c_{\ell(1)}\tau_{\ell(1)\ell(2)}\right)^{-\theta(1-\alpha_2)}}{\sum\limits_{k\in\mathcal{J}}\left(T_k\right)^{1-\alpha_2}\left(c_k\tau_{k\ell(2)}\right)^{-\theta(1-\alpha_2)}}}_{\Pr(\ell(1)|\ell(2))} \times \underbrace{\frac{\left(T_{\ell(2)}\right)^{\alpha_2}\left(\left(c_{\ell(2)}\right)^{\alpha_2}\tau_{\ell(2)j}\right)^{-\theta}\left(\mathcal{E}_{\ell(2)}\right)^{-\theta}}{\sum\limits_{i\in\mathcal{J}}\left(T_i\right)^{\alpha_2}\left(\left(c_i\right)^{\alpha_2}\left(\tau_{ij}\right)\right)^{-\theta}\left(\mathcal{E}_i\right)^{-\theta}}}_{\Pr(\ell(2))}. \tag{15}$$

A bit less trivially, but also exploiting well-known properties of the Fréchet distribution, it can

[14]Note that when $N = 1$, we necessarily have $\alpha_N = 1$, and the formulas (11) and (12) collapse to the well-know trade share formulas in Eaton and Kortum (2002).

be shown that

$$\mathcal{E}_{\ell(2)} = \mathbb{E}\left[\tau_{\ell^*(1)\ell(2)} a^1_{\ell^*(1)}(z)\, c_{\ell^*(1)}\right]^{1-\alpha_2} = \varsigma \left(\sum_{k\in\mathcal{J}} (T_k)^{1-\alpha_2} \left(c_k \tau_{k\ell(2)}\right)^{-\theta(1-\alpha_2)}\right)^{-1/\theta},$$

for some constant $\varsigma > 0$. This allows us to reduce (15) to

$$\pi_{\boldsymbol{\ell}j} = \frac{\left(T_{\ell(1)}\right)^{1-\alpha_2}\left(c_{\ell(1)}\tau_{\ell(1)\ell(2)}\right)^{-\theta(1-\alpha_2)}\left(T_{\ell(2)}\right)^{\alpha_2}\left(\left(c_{\ell(2)}\right)^{\alpha_2}\tau_{\ell(2)j}\right)^{-\theta}}{\sum_{k\in\mathcal{J}}\sum_{i\in\mathcal{J}} (T_k)^{1-\alpha_2}\left(c_k\tau_{ki}\right)^{-\theta(1-\alpha_2)}(T_i)^{\alpha_2}\left((c_i)^{\alpha_2}(\tau_{ij})\right)^{-\theta}}. \tag{16}$$

It should be clear that this expression is identical to (11) – plugging in (12) – for the special case $N = 2$. It is also straightforward to verify that the distribution of final-good prices $p_j^F(\boldsymbol{\ell}, z)$ paid by consumers in j is independent of the actual path of production $\boldsymbol{\ell}$ and is again characterized, as in equation (13), by $\Pr\left(p_j^F(\boldsymbol{\ell}, z) \le p\right) = 1 - \exp\left\{-\tilde{\Theta}_j p^\theta\right\}$, where $\tilde{\Theta}_j$ is the denominator in (16), and is the analog of Θ_j in (12) when $N = 2$.

In sum, this alternative specification of the stochastic nature of technology delivers the exact same distribution of GVCs and of consumer prices as the one in which the overall GVC unit cost is distributed Fréchet. As mentioned above and as demonstrated in Appendix A.3, this isomorphism carries over to the case $N > 2$.

4.2 Mapping the Model to Observables

So far, we have just described how to adapt the Eaton and Kortum (2002) probabilistic approach to apply to trade shares in terms of specific production paths (or GVCs) rather than in terms of trade volumes. Unfortunately, these 'GVC trade shares' are not observable in the data, so we next describe how to map the model to the type of information available in world Input-Output tables. These sources of data provide information on (i) the value of final-good consumption in country j originating in assembly plants (producing stage N) in all other countries i, and (ii) the value of intermediate input purchases used by firms in j originating from producers in all other countries i.

Consider first the implications of our model for final-good consumption. Notice that for final goods to flow from a given source country i to a given destination country j, it must be the case that country i is in position N in a chain serving consumers in country j. Define the set of GVCs flowing through i at position n by $\Lambda_i^n \in \mathcal{J}^{N-1}$, or formally, $\Lambda_i^n = \left\{\boldsymbol{\ell} \in \mathcal{J}^N \mid \ell(n) = i\right\}$. Note then that Λ_i^N corresponds to the set of chains in which assembly is carried out in i. With this notation, the overall relative prevalence of all GVCs serving consumers in j in which country i is in assembly (position N) can be expressed as

$$\pi_{ij}^F = \frac{\sum_{\ell\in\Lambda_i^N}\prod_{n=1}^{N-1}\left(\left(T_{\ell(n)}\right)^{\alpha_n}\left(\left(c_{\ell(n)}\right)^{\alpha_n}\tau_{\ell(n)\ell(n+1)}\right)^{-\theta}\right)^{\beta_n} \times (T_i)^{\alpha_N}\left((c_i)^{\alpha_N}\tau_{ij}\right)^{-\theta}}{\Theta_j}. \tag{17}$$

Because these flows occur at the same expected price for all goods regardless of the actual source country i, it follows that the shares π_{ij}^{F} also correspond to the final consumption shares reported in world Input-Output tables. Our model thus provides explicit formulas for these world Input-Output entries as a function of the parameters of our model and the endogenous composite factor cost c_i, which we can solve for in general equilibrium. Note also that final-good trade flows between any two countries i and j are then simply given by $\pi_{ij}^{F} \times w_j L_j$, since spending on final goods in country j must equal aggregate income, and labor is the only factor of production (when we estimate the model, we will incorporate trade imbalances).

Computing intermediate input flows between any two countries i and j is a bit more tedious, but equally straightforward. To begin, we need to distinguish between two types of intermediate input flows. First, at any stage of production, firms in country j purchase a bundle of materials at cost P_j from firms worldwide, and part of that spending originates in country i. Because the bundle of intermediate corresponds exactly to the consumption CES aggregator, the share of j's input purchases originating in i is again given by π_{ij}^{F} in (17).[15] Furthermore, note that any time the bundle of intermediates is used in production, spending on it in country j equals a multiple $\left(1-\gamma_j\right)/\gamma_j$ of spending on labor. As a result, aggregate flows between i and j of this type of intermediates are given by $\pi_{ij}^{F} \times \left(1-\gamma_j\right)/\gamma_j \times w_j L_j$.

In our multi-stage model, there is a second type of intermediate input flows across countries. In particular, firms in j also import a semi-finished product from i in sequential GVCs in which i *immediately* precedes j. To compute these flows, let us begin by denoting by $\Lambda_{k\to i}^{n} \in \mathcal{J}^{N-2}$ the set of GVCs that flow through k at position $n \leq N-1$ and through i at position $n+1$, or more formally, $\Lambda_{k\to i}^{n} = \left\{ \boldsymbol{\ell} \in \mathcal{J}^{N} \mid \ell(n) = k \text{ and } \ell(n+1) = i \right\}$. The probability that this subset of GVCs emerges in equilibrium in GVCs serving consumers in j is given by

$$\Pr\left(\Lambda_{k\to i}^{n}, j\right) = \frac{\sum_{\boldsymbol{\ell}\in\Lambda_{k\to i}^{n}} \prod_{n=1}^{N-1} \left(\left(T_{\ell(n)}\right)^{\alpha_n} \left(\left(c_{\ell(n)}\right)^{\alpha_n} \tau_{\ell(n)\ell(n+1)} \right)^{-\theta} \right)^{\beta_n} \times \left(T_{\ell(N)}\right)^{\alpha_N} \left(\left(c_{\ell(N)}\right)^{\alpha_N} \tau_{\ell(N)j} \right)^{-\theta}}{\Theta_j}.$$

Note further that all final goods sold in j, command the same expected price regardless of the actual chain, and thus $\Pr\left(\Lambda_{k\to i}^{n}, j\right)$ corresponds to the share of total spending in country j associated with chains that flow through k at position $n \leq N-1$ and through i at position $n+1$ before reaching country j after assembly. Moreover, the value of the trade flow between countries k and i at positions n and $n+1$ is a share β_n of the total spending on that chain in country j.[16] The latter spending comprises final-good consumption $(w_j L_j)$ and spending in the intermediate input bundle $\left(\left(1-\gamma_j\right)/\gamma_j \times w_j L_j\right)$. To find the overall spending of intermediate input purchases by firms in i importing from firms in k immediately upstream from them, we thus just need to aggregate across

[15] In the Eaton and Kortum (2002) model, these are the only type of intermediate input flows and thus there is a unique 'trade share' π_{ij} regardless of the nature of the goods flowing between country i and country j.

[16] This can be verified by iterating (2) and referring to the definition of β_n in (4).

destinations markets j and neighboring stages n and $n+1$ to obtain

$$\mathcal{X}_{ki} = \sum_{j\in\mathcal{J}} \sum_{n=1}^{N-1} \beta_n \Pr\left(\Lambda^n_{k\to i}, j\right) \frac{1}{\gamma_j} w_j L_j.$$

Together with the input flows associated with the more standard roundabout structure of production, we finally obtain that the share of input purchases by firms in i originating in country k is given by:

$$\pi^X_{ki} = \frac{\pi^F_{ki} \frac{1-\gamma_i}{\gamma_i} w_i L_i + \sum_{j\in\mathcal{J}} \sum_{n=1}^{N-1} \beta_n \Pr\left(\Lambda^n_{k\to i}, j\right) \frac{1}{\gamma_j} w_j L_j}{\sum_{k'\in\mathcal{J}} \pi^F_{k'i} \frac{1-\gamma_i}{\gamma_i} w_i L_i + \sum_{k'\in\mathcal{J}} \sum_{j\in\mathcal{J}} \sum_{n=1}^{N-1} \beta_n \Pr\left(\Lambda^n_{k'\to i}, j\right) \frac{1}{\gamma_j} w_j L_j}. \tag{18}$$

Although computing these intermediate input shares is somewhat cumbersome, notice that our model provides an explicit expression for these shares, which have an empirical counterpart in world Input-Output tables (see more on this in section 5).[17]

4.3 General Equilibrium

So far, we have characterized trade flows as a function of the vectors of equilibrium wages $\boldsymbol{w} = (w_1, ..., w_J)$ and of input bundle costs $\boldsymbol{P} = (P_1, ..., P_J)$. We next describe how these vectors are pinned down in general equilibrium.

Notice first that invoking (14) and $c_i = (w_i)^{\gamma_i} (P_i)^{1-\gamma_i}$, we can solve for the vector $\boldsymbol{P}$ as a function of the vector $\boldsymbol{w}$ from the system of equations:

$$P_j = \kappa \left(\sum_{\boldsymbol{\ell}\in\mathcal{J}^N} \prod_{n=1}^{N-1} \left(\left(T_{\ell(n)}\right)^{\alpha_n} \left(\left(c_{\ell(n)}\right)^{\alpha_n} \tau_{\ell(n)\ell(n+1)} \right)^{-\theta} \right)^{\beta_n} \times \left(T_{\ell(N)}\right)^{\alpha_N} \left(\left(c_{\ell(N)}\right)^{\alpha_N} \tau_{\ell(N)j} \right)^{-\theta} \right)^{-1/\theta}, \tag{19}$$

for all $j \in \mathcal{J}$.

To solve for equilibrium wages, notice that for all GVCs, stage n value added (or labor income) accounts for a share $\gamma_{\ell(n)}\alpha_n\beta_n$ of the value of the finished good emanating from that GVC. Furthermore, total spending in any country j is given by the sum of final-good spending ($w_j L_j$) and spending in the intermediate input bundle ($\left(1-\gamma_j\right)/\gamma_j \times w_j L_j$). The share of that spending going to GVCs in which country i is in position n is given by $\Pr\left(\Lambda^n_i, j\right) = \sum_{\boldsymbol{\ell}\in\Lambda^n_i} \pi_{\boldsymbol{\ell}j}$, where remember that we have defined $\Lambda^n_i = \left\{\boldsymbol{\ell}\in\mathcal{J}^N \mid \ell(n) = i\right\}$ and $\pi_{\boldsymbol{\ell}j}$ is given in equation (11). It thus follows that the equilibrium wage vector is determined by the solution of the following system of equations

$$\frac{1}{\gamma_i} w_i L_i = \sum_{j\in\mathcal{J}} \sum_{n\in\mathcal{N}} \alpha_n\beta_n \times \Pr\left(\Lambda^n_i, j\right) \times \frac{1}{\gamma_j} w_j L_j. \tag{20}$$

The system of equations is nonlinear because $\Pr\left(\Lambda^n_i, j\right)$ is a nonlinear function of wages themselves,

[17] We have demonstrated how trade flows within GVCs aggregate into bilateral intermediate input flows. De Gortari (2017) develops a more general framework to study the complementary problem of disentangling the shape of GVCs from aggregate data on bilateral intermediate input flows.

and of the vector $\boldsymbol{P}$, which is in turn a function of the vector of wages $\boldsymbol{w}$.

When $N = 1$, we have that $\alpha_N \beta_N = 1$ and $\Pr\left(\Lambda_i^n, j\right) = \pi_{ij} = \left(\tau_{ij} c_i\right)^{-\theta} T_i / \sum_k \left(\tau_{kj} c_k k\right)^{-\theta} T_k$. The equilibrium then boils down to a simple generalization of the general equilibrium in Eaton and Kortum (2002) and Alvarez and Lucas (2007), with cross-country variation in how the composite factor aggregates value added and the bundle of intermediate inputs.

In Online Appendix B.2, we build on Alvarez and Lucas (2007) to show that, given a vector of wages $\mathbf{w}$, the system of equations in (19) delivers a unique vector of input bundle costs $\boldsymbol{P}$. In that Appendix, we also demonstrate the existence of a solution $\mathbf{w}^* \in \mathbb{R}_{++}^J$ to the system of equations in (20) – with (19) plugged in – and we derive a set of sufficient conditions that ensure that this solution is unique.

4.4 Gains from Trade

We next study the implications of our framework for how changes in trade barriers affect real income in all countries. Consider a 'purely-domestic' value chain that performs all stages in a given country j to serve consumers in the same country j. Let us denote this chain $\boldsymbol{\ell} = (j, j, ..., j)$ by j^N. From equation (11), such a value chain would capture a share of country j's spending equal to

$$\pi_{j^N} = \Pr\left(j^N\right) = \frac{\left(\tau_{jj}\right)^{-\theta\left(1+\sum_{n=1}^{N-1}\beta_n\right)} \times \left(c_j\right)^{-\theta} T_j}{\Theta_j},$$

where we have used the fact that $\sum_{n=1}^{N} \alpha_n \beta_n = 1$. Combining this equation with (14) and $c_j = \left(w_j\right)^{\gamma_j} \left(P_j\right)^{1-\gamma_j}$, we can establish that real income in country j can be expressed as

$$\frac{w_j}{P_j} = \left(\kappa \left(\tau_{jj}\right)^{1+\sum_{n=1}^{N-1}\beta_n}\right)^{-1/\gamma_j} \left(\frac{T_j}{\pi_{j^N}}\right)^{1/\left(\theta\gamma_j\right)}. \tag{21}$$

Consider now a prohibitive increase in trade costs that brings about autarky, but leaves all other technological parameters (α_n, T_j, γ_j, θ, κ) unchanged. Because under autarky $\pi_{j^N} = 1$, we can conclude that the (percentage) real income gains from trade, relative to autarky, are given by $\left(\pi_{j^N}\right)^{-1/\left(\theta\gamma_j\right)} - 1$. This formula is analogous to the one that applies in the Eaton and Kortum (2002) framework (and the wider class of models studied by Arkolakis et al., 2012). An important difference, however, is that π_{j^N} is not the share of spending on domestic finished goods (π_{jj}^F in equation (17)), but rather the share of spending on goods that *only* embody domestic value added. The latter share π_{j^N} is necessarily lower than π_{jj}^F (and increasingly so, the larger number of stages), and thus the gains from trade emanating from our model are expected to be larger on this account. This result is similar to the one derived by Melitz and Redding (2014) in an Armington framework with sequential production, and also bears some resemblance to Ossa's (2015) argument that the gains from trade can be significantly larger in a multi-sector models, with stages in our model playing the role of sectors in his framework. One can also show that our Cobb-Douglas assumption in technology is not essential for this result: the gains from trade would still be given $\left(\pi_{j^N}\right)^{-1/\left(\theta\gamma_j\right)} - 1$

for any CES multi-stage production technology with an elasticity of substitution lower than one between the value added at different stages.[18]

Another key distinctive feature of the formula in (21) is that, unlike π_{jj}^F, π_{jN} cannot be directly observed in the data, and thus the sufficient statistic approach advocated by Arkolakis et al. (2012) is not feasible in our setting. Instead, one needs a model to structurally back out π_{jN} from available data. For a similar reason, the hat algebra approach to counterfactual analysis proposed by Dekkle et al. (2008) is not feasible in our setting.

Although we have argued above that $\pi_{jN} < \pi_{jj}^F$ implies larger gains from trade in our model than in models without sequential production, it should be noted that the values of γ_j and θ that are appropriate for our model might be different from those appropriate for a model without multi-stage production. First, remember that our model features an additional type of intermediate input flows relative to a model with roundabout production. In order to match the empirical ratio of value added to gross output in each country, our model will thus require setting relatively higher values of γ_j, which other things equal, will lead to lower gains from trade. As for the parameter θ governing the elasticity of trade flows to iceberg trade costs, we can no longer rely on simple gravity equation specifications to back out that parameter. Furthermore, our model suggests that the trade elasticity should on average be lower for intermediate inputs than for final goods, a prediction we will find some suggestive empirical support for below. This suggests that the proper way to calibrate our model entails setting a value of θ higher than the one that would be suitable to calibrate a Ricardian model without multi-stage production. As in the case of γ_j, the use of a larger value of θ would again generate a downward correction to the gains from trade. Overall, whether our model generates larger or smaller gains from trade than models without multi-stage production is an empirical question, and one which we will explore in section 7.

4.5 The Centrality-Downstreamness Nexus

We finally exploit the tractability of our framework to formally explore the role of a country's geography (and, in particular, its centrality) in shaping its average position in GVCs. In order to isolate the role of geography in shaping GVC positioning, we further focus on the 'symmetric' case $\alpha_n \beta_n = 1/N$ for all stages n, which amounts to assuming $\alpha_n = 1/n$ and $\beta_n = n/N$ for all $n \in \mathcal{N}$. Without this assumption, technology would not be symmetric in the value added originated at different stages, and thus the state of technology T_i of a country would affect different stages differentially, thereby generating *technological* comparative advantage.[19]

In order to formalize a centrality-downstreamness nexus, let us define the average upstreamness

[18] We thank Arnaud Costinot for this observation. It should be clear, however, that the π_{iN} one would back out from available data would depend on the multi-stage production function one specifies.

[19] For instance, if downstream stages contributed more to overall value added than upstream stages, we would obtain a prediction analogous to that in Costinot et al. (2013), namely that countries with better technologies T_i have comparative advantage in downstream stages.

of production of a given country i in value chains that seek to serve consumers in country j, by

$$U(i;j) = \sum_{n=1}^{N} (N - n + 1) \times \frac{\Pr(\Lambda_i^n, j)}{\sum_{n'=1}^{N} \Pr\left(\Lambda_i^{n'}, j\right)}, \tag{22}$$

where remember that $\Pr(\Lambda_i^n, j)$ is the probability that country i features in position n in value chains leading to consumption in country j, and corresponds to $\Pr(\Lambda_i^n, j) = \sum_{\boldsymbol{\ell} \in \Lambda_i^n} \pi_{\boldsymbol{\ell} j}$, with $\pi_{\boldsymbol{\ell} j}$ given in equation (11). The index $U(i;j)$ in (22) is thus a weighted average distance of country i from final consumers in value chains that service consumers in country j, and it is closely related to measure proposed by Antràs et al. (2012).[20] Although $U(i;j)$ in equation (22) uses probabilities rather than expenditure shares as weights, it can be verified that given our symmetry assumption, these probabilities correspond to the share of country i's value added at each stage in chains ending in consumption in country j.

We seek to establish a connection between the measure of upstreamness $U(i;j)$ and the centrality of country i. As in section 2, the structure of equation (11) already hints at a negative association between the two, since high values of trade costs (high τ_{ij}) in relatively downstream stages (high n) have a disproportionately negative effect on the likelihood of a given permutation of countries forming an equilibrium value chain. In order to develop a more precise formulation of this result, we assume that the easiness of trade between any two countries i and j can be decomposed as $(\tau_{ij})^{-\theta} = \rho_i \rho_j$, where we take ρ_i to be an index of the *centrality* of country i. Notice that if country i is more central than country j, then it is cheaper to ship from i to any other country in the world than it is to ship from country j. This is a rather strong notion of centrality but it has the virtue of providing the following stark result (which we prove in Appendix A.4):

Proposition 1 The more central a country i is (i.e., the higher is ρ_i), the lower is the average upstreamness $U(i;j)$ of this country in global value chains leading to consumers in any country $j \in J$.

In the next section, we will provide suggestive evidence consistent with this prediction.

5 Data Sources and Suggestive Evidence

Before turning to the structural estimation of our model, in this section we describe the data used in the paper and provide suggestive evidence for some of the key mechanisms in our model.

Our structural estimation relies exclusively on the data contained in world Input-Output tables (WIOTs). To understand the nature of the data, it is useful to refer to the schematic representation of a WIOT in Figure 3. This matrix is split into two blocks with the block on the left containing data on bilateral intermediate input trade flows across countries (denoted by X_{ij}) and the block on

[20] To be precise, the index $U(i;j)$ in (22) reflects the average distance of country i from final consumers ignoring country i's contribution to GVCs via the composite of materials used by firms worldwide in all production stages. See de Gortari (2017) for additional details.

		Input use & value added			Final use			Total use
		Country 1	$\cdots$	Country J	Country 1	$\cdots$	Country J	
Intermediate inputs supplied	Country 1	X_{11}	$\cdots$	X_{1J}	F_{11}	$\cdots$	F_{1J}	Y_1
	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
	Country J	X_{J1}	$\cdots$	X_{JJ}	F_{J1}	$\cdots$	F_{JJ}	Y_J
Value added		$w_1 L_1$	$\cdots$	$w_J L_J$				
Gross output		Y_1	$\cdots$	Y_J				

Figure 3: A schematic world Input-Output table.

the right containing the data for final good trade flows (denoted by F_{ij}). Each row represents the sales of each country to every other country, while each column represents the purchases of each country from every other country. More specifically, each row breaks down the sales of a country to every other country into sales for production purposes (i.e., intermediate input sales) and sales for consumption purposes (i.e., final good sales). Hence, the sum across a row equals a country's gross output (denoted by Y_j). Meanwhile, columns in the left block contain intermediate input purchases by each country so that the sum across a column equals gross output minus value added (the latter denoted by $w_j L_j$). Finally, summing down a column on the right block delivers aggregate final good consumption.[21]

Given the cells of a WIOT, it is straightforward to construct empirical analogs to our model's key equilibrium variables, namely the final-good π_{ij}^F and intermediate input shares π_{ij}^X in (17) and (18), as well as gross output and value added in each country. More specifically, denoting by a 'hat' these empirical moments, we have

$$\hat{\pi}_{ij}^F = \frac{F_{ij}}{\sum_{i' \in \mathcal{J}} F_{i'j}}, \qquad \hat{\pi}_{ij}^X = \frac{X_{ij}}{\sum_{i' \in \mathcal{J}} X_{i'j}}, \qquad \hat{Y}_j = \sum_{i \in \mathcal{J}} X_{ji} + \sum_{i \in \mathcal{J}} F_{ji}, \qquad \hat{w}_j L_j = \hat{Y}_j - \sum_{i \in \mathcal{J}} X_{ij}. \tag{23}$$

Building a WIOT of the type in Figure 3 is a formidable endeavor because it requires collecting trade and production data from many different sources, including national and supra-national statistical offices, but also because it necessarily requires assumptions and data analysis in order to make the data comparable. In this paper we work, for the most part, with the World Input Output Database (or WIOD for short), the outcome of a project that was carried out by a consortium of 12 research institutes headed by the University of Groningen in the Netherlands (see Timmer et al., 2015). We choose this dataset for our benchmark estimation because we believe that the assumptions put into its construction are less heroic than those contained in other sources. The main limitation of the WIOD is that it only covers 43 relatively developed countries, and includes no African country and only two countries in Latin America (Brazil and Mexico).[22] With that in

[21] Note that the difference between aggregate final consumption and value-added is the trade deficit/surplus. These deficits are nontrivial for certain countries and are taken into account in both our estimation and counterfactual exercises, as discussed below.

[22] Two releases of the WIOD are available. The 2016 release contains a WIOT covering 43 countries and the rest of the world for the period 2000-2014. A previous relase (2013) contained information for 40 countries and the rest of the world, for the period 1995-2011. See http://www.wiod.org.

mind, we will also estimate our model using the more comprehensive Eora MRIO database, which provides yearly WIOTs covering 190 countries and the rest of the world for the period 1990-2013. Although we find this dataset less reliable than the WIOD, its broader sample will enrich the set of counterfactuals studied in section 7.

Figure 4 depicts some salient characteristics of the data we employ in our structural estimation. The left panel plots the diagonal elements $\hat{\pi}_{ii}^F$ of the final-good matrix against the diagonal elements $\hat{\pi}_{ii}^X$ of the input matrix in the WIOD for the year 2014, with the size of each observation being proportional to each country's GDP. As is clear from the graph, most observations lie above the 45 degree line, indicating that the input $\hat{\pi}_{ij}^X$ and output $\hat{\pi}_{ij}^F$ matrices are asymmetric, and that countries tend to rely on foreign sources more prevalently for inputs than for final goods. This pattern, which is also observed in the Eora dataset (see Figure A.2 in the Appendix), is consistent with the notion that trade costs are more detrimental for downstream versus upstream stages.

The right panel of Figure 4 plots each country's gross output to final-good sales (GO/F) against its gross output to value added ratio (GO/VA). In a closed economy, these two ratios would naturally coincide and all observations would lie on the 45 degree line. In a globalized world, differences in these ratios provide a rough measure of the positioning of countries in GVCs. More specifically, for a given GO/VA, a high share of final output in overall sales (i.e., a low GO/F in the vertical axis) would indicate that a country is relatively downstream in GVCs. Similarly, for a given GO/F, a high ratio GO/VA would indicate that a country uses a relatively large amount of foreign inputs in production, which again would suggest a relatively downstream position of this country in GVCs. Hence, in a world in which countries are in markedly different segments of GVCs, the ratio GO/VA and GO/F might be expected to be negatively correlated. With this background, the figure indicates that although there are some deviations from the 45 degree line, cross-country variation in these ratios is much larger than within-country differences. These considerations motivate the introduction of cross-country variation in the value-added shares γ_j in our model. Without such variation, it would be impossible for our model to match the stark positive correlation of the ratios GO/VA and GO/F documented in the right panel of Figure 4.

Suggestive Evidence

Having described our main data sources, we complete this section by exploring the empirical relevance of a key mechanism of the model – namely, the fact that the trade elasticity is larger for downstream stages than for upstream stages – and of the centrality-downstreamness nexus result in Proposition 1. The empirical tests in this section are reduced-form in nature and not structurally related to our model, but we deem them to be informative nonetheless. We also note that the tests below use additional sources of data that are not employed in the structural estimation, so we will only discuss them succinctly as we introduce them.

We begin by studying empirically the compounding effect of trade costs, a key feature of our model. A crude way to assess the differential sensitivity of trade flows to trade costs at different stages of the value chain is to compare the elasticity of intermediate-input and final-good flows to

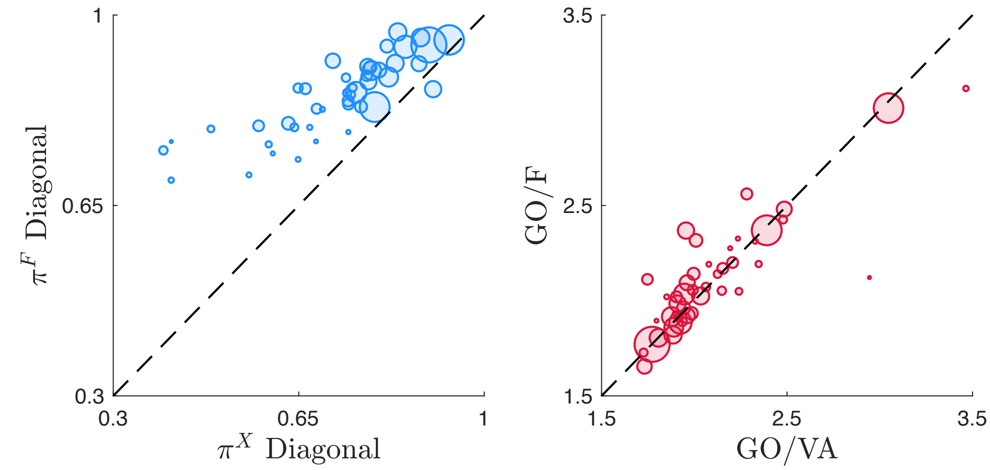

Figure 4: Some Key Features of the World Input Output Database

various proxies for trade costs τ_{ij}. In particular, and building on the gravity equation literature, consider projecting the bilateral trade cost parameters τ_{ij} on a vector of pair-specific variables including distance, contiguity and a common language indicator. More specifically, let

$$\ln \tau_{ij} = \ln \kappa + \delta_{dist} \ln Distance_{ij} + \delta_{con} Contiguity_{ij} + \delta_{lang} SameLanguage_{ij}.$$

As long as the coefficients δ_{dist}, δ_{con} and δ_{lang} are common for intermediate inputs and final goods, then any difference in the sensitivity of final-good versus intermediate-input trade flows to these bilateral gravity variables will be indication of a differential sensitivity of 'upstream' versus 'downstream' trade flows to trade costs. To assess the plausibility of this approach, consider the case of the distance elasticity δ_{dist}. Our key identification assumption in that case is that trade costs, as a percentage of the value of the good being shipped, are identical regardless of whether the good is an input or a final-good. If we observe final-good trade being more sensitive to distance than input trade, we will then conclude that final-good trade is more sensitive to trade costs than input trade is.

We implement this test in Table 1. In columns (1) and (2) of that table we report the results of a standard gravity specification in which the log of *aggregate* shipments from country i to country j are run on exporter and importer fixed effects, as well as the log of distance between i and j and dummy variables for whether i and j share a contiguous or share a common border. Our shipments data are from 2011 and correspond to the Eora MRIO database (described above). The data cover 190 countries and include information on domestic shipments (i.e., sales from i to i). The gravity variables are from the CEPII dataset for the year 2006 (the most recent one available), and the merge between these two data sources leaves us with information on 180 countries.

Our results in columns (1) and (2) are fairly standard. Distance reduces trade flows with an elasticity of around -1, while contiguity and common language have a sizeable positive effect on bilateral flows. Starting in column (3), we exploit a key advantage of the Eora MRIO database, namely the fact that it reports separately bilateral shipments of intermediate inputs (X_{ij}) and of finished goods (F_{ij}). In columns (3) and (4), we pool these observations and re-run the specifications in columns (1) and (2), while clustering at the country-pair level. It is clear that the results are almost identical to those in columns (1) and (2). More interestingly, in column (5) we document that the elasticity of trade flows to distance is significantly larger for final-good trade (-1.210) than for intermediate-input trade (-1.077). The difference is sizeable and highly statistically significant. In column (6), we document a similar phenomenon: the positive effect of contiguity and common language on trade flows is significantly attenuated when focusing on the intermediate-input component of trade. Finally, in column (7) we introduce a dummy variable for intranational shipments as well as its interaction with input trade. As is well-known from the border-effect literature, the domestic trade dummy is very large, but we again observe that it is significantly lower for input trade, a result consistent with the scatter plot in the left panel of Figure 4 discussed above.

Table 1. Trade Cost Elasticities for Final Goods and Intermediate Inputs

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Distance	-1.111***	-0.823***	-1.144***	-0.851***	-1.210***	-0.903***	-0.794***
	(0.019)	(0.014)	(0.019)	(0.014)	(0.021)	(0.015)	(0.015)
Distance × Input					0.133***	0.106***	0.098***
					(0.006)	(0.006)	(0.006)
Contiguity		2.187***		2.198***		2.287***	1.184***
		(0.111)		(0.112)		(0.120)	(0.099)
Contiguity × Input						-0.177***	-0.054
						(0.037)	(0.040)
Language		0.480***		0.507***		0.596***	0.513***
		(0.026)		(0.027)		(0.029)	(0.027)
Language × Input						-0.179***	-0.169***
						(0.013)	(0.013)
Domestic							5.635***
							(0.187)
Domestic × Input							-0.599***
							(0.067)
Observations	32,400	32,400	64,800	64,800	64,800	64,800	64,800
R^2	0.98	0.982	0.972	0.974	0.972	0.974	0.976

Notes: Standard errors clustered at the country-pair level reported. ***, **, and * denote 1, 5 and 10 percent significance levels. All regressions include exporter and importer fixed effects. Regressions in columns (3)-(7) also include a dummy variable for inputs flows.

Taken together, the results in Table 1 are highly suggestive of trade barriers impeding trade more severely in downstream stages than in upstream stages. In Online Appendix B.4, we further show that our results are not materially affected when pooling data from all years (1995-2013) for which the Eora dataset is available (instead of just using 2011 data).[23] We also repeat our tests using data from the two releases of the WIOD database, which cover a smaller and more homogenous set of countries. The results with the 2013 release of the WIOD continue to indicate a significantly lower distance elasticity and lower 'home bias' in intermediate-input relative to final-good trade. Nevertheless, with the 2016 release of the same dataset, we only find support for the second differential effect (see Online Appendix B.4 for details). This last result makes us interpret our results with caution. Another important caveat with the evidence above is that it is based on gravity-style specifications that, technically, are inconsistent with our theoretical framework. As equations (17) and (18) indicate, bilateral trade flows of final goods and intermediate inputs will typically be affected by trade costs associated with third countries (see Morales et al., 2014, and Adao et al,. 2017, for recent evidence of these third-market effects).

We next turn to examining the empirical relevance of the downstreamness-centrality nexus formalized in Proposition 1. For that purpose, we build on Antràs et al. (2012) who propose a measure of the positioning of countries in GVCs and study how this measure correlates with various country-level variables. More specifically, Antràs et al. (2012) propose a measure of industry "upstreamness" (or average distance of an industry's output from final use) and then compute the average upstreamness of a country's export vector using trade flow data from the BACI dataset for the year 2002. Column (1) of Table 2 reproduces exactly their baseline specification, which includes 120 countries, and correlates a country's upstreamness with its GDP per capita, rule of law, financial development, capital-labor ratio and human capital (schooling).[24] Only financial development and schooling have a statistically significant partial correlation with upstreamness.

In order to assess the relationship between upstreamness and centrality, we simply add a measure of centrality to the core specification in column (1). In particular, for each country j we compute $Centrality_j^{GDP} = \sum_i \left(GDP_i/Distance_{ji}\right)$ and $Centrality_j^{pop} = \sum_i \left(Population_i/Distance_{ji}\right)$, which capture a country's proximity to other countries with either large GDP or large population (or both). We are able to compute these measures for only 118 of the original 120 countries in Antràs et al. (2012), so for completeness, column (2) reproduces the results of running the same specification as in column (1) with only those 118 countries. Clearly, the results are not materially affected. More interestingly, in columns (3) and (4) we document a highly statistically significant negative relationship between upstreamness and each of the two measures of centrality. This partial correlation is not driven by the presence of the other covariates: column (5) shows that it persists when only controlling for GDP per capita, and column (6) demonstrates that it holds even unconditionally. In Online Appendix B.4, we plot this relationship and show that it is not driven by any outliers. Though these correlations cannot be interpreted causally, they are again suggestive of the empirical relevance of the nexus between centrality and downstreamness highlighted in Proposition 1.

[23] In fact, we have obtained extremely stable results when running these same regressions year by year for this same period 1995-2013.

[24] The source for each of these variables is discussed in Antràs et al. (2012).

Table 2. Export Upstreamness and Centrality

	(1)	(2)	(3)	(4)	(5)	(6)
Centrality (GDP weighted)			-0.173^{***}		-0.233^{***}	-0.155^{***}
			(0.065)		(0.061)	(0.044)
Centrality (population weighted)				-0.228^{***}		
				(0.084)		
Log(Y/L)	0.083	0.082	0.102	0.046	0.083^{*}	
	(0.142)	(0.142)	(0.138)	(0.148)	(0.046)	
Rule of Law	-0.029	-0.026	0.010	0.010		
	(0.103)	(0.104)	(0.105)	(0.105)		
Credit/Y	-0.437^{***}	-0.440^{***}	-0.375^{***}	-0.407^{***}		
	(0.136)	(0.137)	(0.130)	(0.135)		
Log(K/L)	0.156	0.159	0.163	0.188		
	(0.131)	(0.132)	(0.129)	(0.132)		
Schooling	-0.085^{***}	-0.085^{***}	-0.083^{***}	-0.094^{***}		
	(0.031)	(0.031)	(0.030)	(0.029)		
Observations	120	118	118	118	118	118
R^2	0.154	0.153	0.194	0.199	0.083	0.056

Notes: Robust standard errors reported. ***, **, and * denote 1, 5 and 10 percent significance levels.

6 Estimation

We next turn to a more structural empirical exploration of our multi-stage model. In our baseline, we will employ data from the WIOD for the year 2014, but we will subsequently replicate our estimation using the broader Eora dataset for the year 2013.

It is useful to begin by outlining the parameters we need to estimate or calibrate for a given number J of countries and N of stages. Geography is pinned down by the $J \times J$ matrix of iceberg trade costs τ_{ij}. Production depends on the Cobb-Douglas input expenditure shares α_n, which are stage-specific but common across countries, while the labor value-added shares γ_j are country-specific but common across stages. There are thus $N-1$ α_n's to estimate and J γ_j's. Lastly, labor productivity depends on a vector $J \times 1$ of country-specific state-of-technology levels T_j, while comparative advantage is governed by a single parameter θ. Although countries are also allowed to vary in terms of their supply of equipped labor, the particular values of L_j only affect the estimates of T_j and of equilibrium wages, but have no bearing on the counterfactuals discussed below. In order to keep the estimates of T_j in economically meaningful levels, we simply normalize equipped labor as $L_j = (capital_j)^{\frac{1}{3}} (population_j)^{\frac{2}{3}}$, where both capital and population are drawn from the Penn World Tables for the year 2013 (in the Eora estimation) and for 2014 (in the WIOD estimation).

To pin down trade costs, we follow the method proposed by Head and Ries (2001) and make the simplifying assumption that domestic trade costs are common across countries and normalized to 0, i.e., $\tau_{jj} = 1$ for all $j \in \mathcal{J}$. International trade costs, up to a power $-\theta$, can then be immediately read off the data through the use of equation (17) and our empirical analogs in (23):

$$\tau_{ij}^{-\theta} = \sqrt{\frac{\hat{\pi}_{ij}^F}{\hat{\pi}_{ii}^F} \frac{\hat{\pi}_{ji}^F}{\hat{\pi}_{jj}^F}}. \tag{24}$$

Trade costs are symmetric by construction, i.e. $\tau_{ij} = \tau_{ji}$, and in practice the triangle inequality (i.e., $\tau_{ij} \leq \tau_{ik}\tau_{kj}$) holds across more than 99.9% of triples.

A consequence of our approach to backing out trade costs is that the calibrated values for $\tau_{ij}^{-\theta}$ we obtain are unaffected by the particular value of θ chosen. Although the value of θ affects the equilibrium of our model beyond its effect on $\tau_{ij}^{-\theta}$, it turns out that the moments we employ for our structural estimation (see below) do *not* identify θ. More precisely, for every possible value of θ, there exists a re-normalization of T_j that yields the same equilibrium (conditional on the same set of parameters γ_j and α_n). With that in mind, we simply set $\theta = 5$ in our estimation. This value is slightly higher than is typically assumed in the literature, but our model predicts that the trade elasticity for final goods (i.e., θ) should be larger than the elasticity one would estimate with overall trade flows (which is a weighted average of $\theta\beta_1$, $\theta\beta_2$,...,θ).[25] We will return to this point below, after having estimated the β_n's (i.e., the α_n's).

Another tricky parameter to calibrate is the number of stages N. We initially estimate the model for the case $N = 2$. This is the simplest multi-stage model one could estimate, and as we will discuss extensively below, the data actually appears to favor this value of N over larger ones.

Having pinned down the matrix of trade costs $\tau_{ij}^{-\theta}$, θ, and N, we estimate the remaining parameters of the model by targeting specific moments of a WIOT via the generalized method of moments (GMM). More specifically, we target four sets of moments, each constituting a $J \times 1$ vector: (i) the diagonal elements $\hat{\pi}_{jj}^F$ of the final-good matrix; (ii) the diagonal elements $\hat{\pi}_{jj}^X$ of the input matrix, (iii) the gross output to value added ratio GO_j/VA_j in each country, and (iv) the GDP share $\hat{w}_j L_j / \sum_i \hat{w}_i L_i$ of each country.[26]

The choice of these moments is motivated by the following considerations. First, note that the vector of state of technology parameters T_j naturally shapes wages and thus (for given L_j) variation in GDP shares across countries. These technology parameters also affect the extent to which countries rely on local versus foreign sources of inputs and final goods (and thus the shares $\hat{\pi}_{jj}^X$ and $\hat{\pi}_{jj}^F$). Meanwhile, the input expenditure shares α_n determine how fast the trade elasticity increases along GVCs, and are thus crucial in shaping the observed differences between the domestic

[25] Simonovska and Waugh (2014), in a widely cited study, find a range for the elasticity of trade between 2.47 and 5.51. Using U.S. import data, Antràs et al. (2017) estimate an elasticity of trade of 4.54. Finally, Caliendo and Parro (2014) find an aggregate elasticity of 4.49 using the previous release of the WIOD data.

[26] In order to guarantee that our model provides a proper quantitative evaluation of the general-equilibrium workings of the world economy, we place a higher weight on matching the empirical moments of larger economies. More precisely, our weighting matrix is a diagonal matrix with GDP shares in the diagonal.

input share $\hat{\pi}_{jj}^X$ and the domestic final expenditure share $\hat{\pi}_{jj}^F$. With $N = 2$, we need only estimate α_2 (since $\alpha_1 = 1$ by assumption). If α_2 were to be close to 1, the sequentiality of production would become immaterial, and all input trade would reflect a roundabout structure of production, as in Eaton and Kortum (2002). In such a case, we would not expect large asymmetries between $\hat{\pi}_{jj}^X$ and $\hat{\pi}_{jj}^F$. Our previously discussed evidence in the left panel of Figure 4 already indicated the empirical importance of those asymmetries, so we would not expect α_2 to be too close to 1.

Finally, the vector of gross output to value added ratios GO_j/VA_j is a natural target for the vector of country-specific value-added share parameters γ_j. To see this, note that with $N = 1$, the gross output to value-added ratio is given by

$$\frac{GO_j}{VA_j} = \frac{Y_j}{w_j L_j} = \frac{w_j L_j + \frac{1-\gamma_j}{\gamma_j} w_j L_j}{w_j L_j} = \frac{1}{\gamma_j}.$$

When $N > 1$, the expression for the gross output to value-added ratio is more complicated and the other parameters of the model – and most notably the input shares α_n – have an influence over GO_j/VA_j. To see this, consider our estimation with $N = 2$. For a given γ_j, the gross-output to value-added ratio will be close to $1/\gamma_j$ when the upstream stage of production is irrelevant for production (i.e., when $\alpha_2 \rightarrow 1$), since this corresponds to reducing N from 2 to 1. Conversely, when $\alpha_2 \rightarrow 0$, the downstream stage of production adds very little value, and the gross output to value added ratio is close to $2/\gamma_j$, since the same output is shipped twice but value is added essentially only once. In practice, for a general N, the gross output to value added ratio features variation (see the right panel of Figure 4 above) both because countries have different labor value-added shares but also because they find themselves at different degrees of upstreamness along the GVC; the interaction of both forces determine this statistic.

Before turning to a discussion of our estimation results, we briefly comment on our treatment of trade imbalances. As mentioned above (see footnote 21), these imbalances are empirically nontrivial and correspond to the difference between aggregate final consumption and value added. Following a common approach in the trade literature (see, in particular, Costinot and Rodríguez-Clare, 2015), we treat these deficits as exogenous parameters, and we adjust our general-equilibrium equations to account for the difference between income and spending (see Online Appendix B.3).

Estimation Results

We now turn to discussing our estimation results and overall fit of the model. We mostly focus our discussion on the results we obtain using the WIOD, but at the end of this section, we also briefly describe the results with the broader Eora database.

As anticipated above, the asymmetries between the input and final-output diagonal elements of the WIOD lead to an estimate of α_2 far removed from one. In particular, we obtain $\alpha_2 = 0.16$. The estimated values for the vectors of γ_j and T_j are reported in Appendix Table A.1. Figure 5 presents a comparison between the data and the targeted moments, with the size of each observation proportional to GDP. The values for the diagonal elements π_{jj}^X, the gross output to GDP ratios,

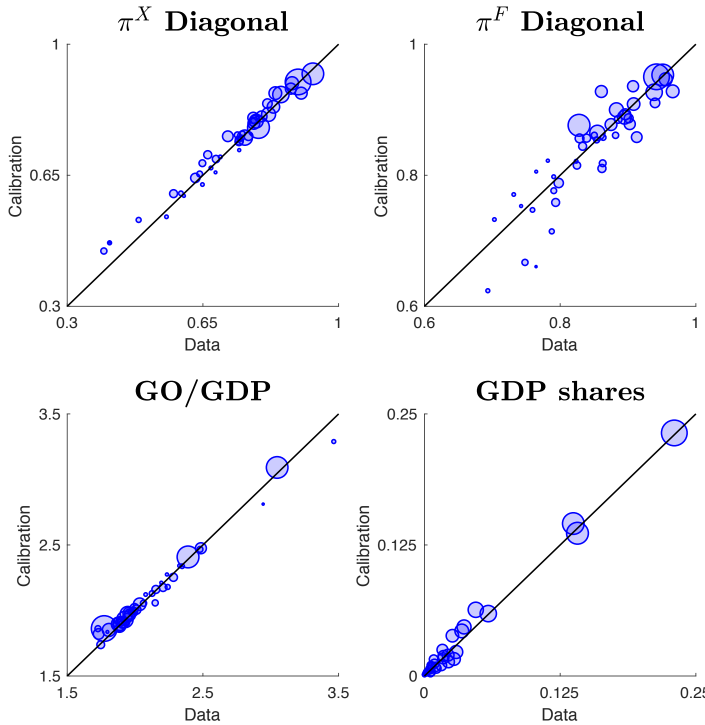

Figure 5: Targeted Moments

and GDP shares are all estimated very accurately, with correlations equal to 0.99, 0.97, and 0.99 with their empirical counterparts, respectively. The fit of the final-output diagonal elements π^F_{jj} is also very good (the correlation with data is 0.90), but it also presents some slight discrepancies, especially for some small countries (remember that our estimation algorithm weighs observations by country size).

Figure 6 performs a similar comparison between model and data but for moments that were not directly targeted in the estimation. The upper two charts present the non-diagonal elements of π^X and π^F, and those entries are also matched relatively accurately in both cases (with correlations equal to 0.83 and 0.91, respectively). The lower two charts explore how well our model matches the backward and forward GVC participation of various countries. Because we will later explore counterfactuals exercises that illustrate changes in the participation of countries in GVCs, it is desirable that our calibration matches these type of moments properly. These two measures of the positioning of countries in GVCs are proposed in Wang et al. (2017). The backward GVC participation index measures the share of a country's production of final goods and services that is accounted for by imported value added. More specifically, the numerator in the share includes foreign value added that is embodied in intermediate input imports used to produce final goods in a country, and it also includes domestic value added that has returned home embodied in those same imported inputs. The forward GVC participation index measures the share of a country's domestic value added that is exported worldwide embodied in intermediate goods that are consumed by both foreign and domestic firms downstream. Note that this second measure excludes domestic value added embodied in final goods that are exported directly to consumers. Our benchmark calibration fits both moments very well with correlations of 0.99 and 0.96, respectively.

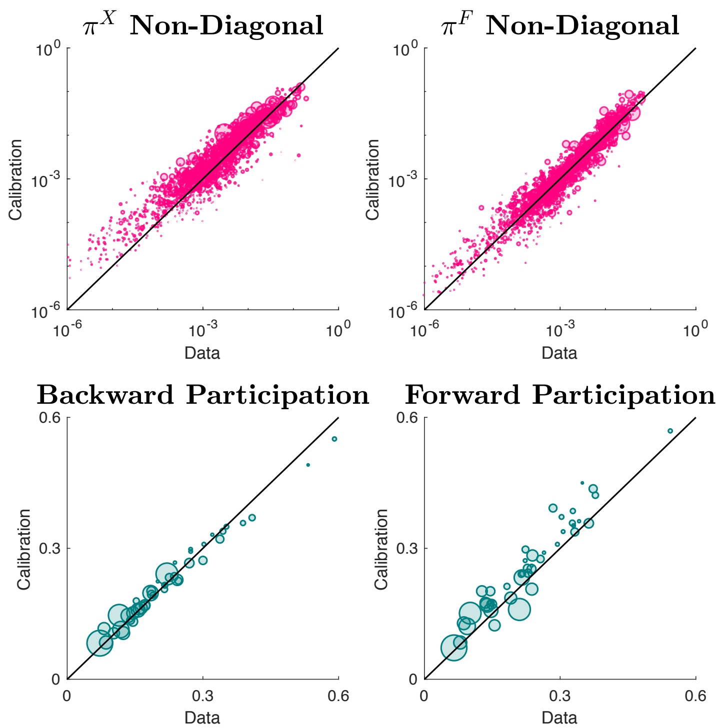

Figure 6: Untargeted Moments.

We next repeat our estimation with the use of the Eora data for 2013. Though the full Eora database contains 190 countries we consolidate it into a set of 101 country/regions in order to alleviate the burden of calibrating so many parameters.[27] Remarkably, we estimate an upstream input share of $\alpha_2 = 0.19$, which is very similar to the value of $\alpha_2 = 0.16$ found for the WIOD. In Appendix A.5, we provide estimates for the remaining parameters, and also illustrate the fit of the estimation via figures analogous to those in Figure 5 and 6. For both targeted and non-targeted moments, the fit continues to be extremely good.[28]

It is useful to compare our estimates of α_2 with those implied by our reduced-form results in Table 1, which also used the Eora dataset. Although the gravity-style specification in Table 1 cannot be mapped structurally to our model, the differential effect of distance on input and final-

[27] Specifically, we keep all countries with a population of 10 million or more and aggregate the rest into a set of 9 regions: Latin America and Caribbean, Central Europe, Eastern Europe, Western Europe, Scandinavia, Middle East and North Africa, Sub-Saharan Africa, Central Asia, and East Asia and Pacific

[28] The correlations between data and model for the four targeted moments are 0.96, 0.92, 0.89, and 0.99, respectively. The correlations for the off diagonal elements of π^X and π^F are 0.81 and 0.90, while the correlations for the backward and forward participation index stand at 0.84 and 0.69.

good trade is informative on the relative size of α_2. More specifically, the ratio of the elasticity of stage 1 output trade to stage 2 output trade is given by $1-\alpha_2$ in our model. Given the distance elasticities estimated in column (7) of Table 1, and assuming that all input trade is stage 1 output, we would then infer $\alpha_2 = 1 - 0.696/0.794 = 0.12$. Now, of course, in our model not all input trade is stage 1 output, since value added is combined with a bundle of materials at each stage, and the trade elasticity of that "roundabout" input trade is equal to that of final-good trade. Using the structural estimates of our model we find that around 18 percent of input trade takes this "roundabout" form. The actual elasticity of stage 1 input is thus lower than is implied by the results in Table 1 (0.675 rather than 0.696), and the implied α_2 is slightly larger ($\alpha_2 = 0.15$), and very close to the one we have estimated structurally.

Revisiting the Calibration of N

Up to now, we have fixed the number of stages to $N = 2$. Estimating our model for $N > 2$ is computationally more demanding but straightforward to carry out. In terms of the parameters to estimate, notice that this only amounts to estimating a longer vector of input shares α_n. Perhaps surprisingly, we have found that the structural estimation shuts down production stages that are more than one stage removed from final consumption, and delivers estimates for the other parameters that are identical to those in the benchmark with $N = 2$. To give a precise example, when we estimate the model with $N = 3$, our calibration delivers $\alpha_3 = 0.16$ and $\alpha_2 = 1$. The most upstream stage of production, $n = 1$, is thus effectively shut down (i.e., its output is negligible). The recovered parameters for γ_j are exactly the same as in our benchmark calibration while those for T_j are exactly those consistent with our benchmark calibration as well.[29]

Why does our model reject $N > 2$? A first important point to make is that we are calibrating an average N for the *whole* world economy, including sectors in which chains might be large (e.g., in some manufacturing sectors) but sectors in which they might be very short (e.g., certain types of services). Relatedly, the worldwide ratio of gross output to value added is 3.82 in manufacturing (in the 2014 WIOD), while it is 1.78 for non-manufacturing sectors. The fact that the aggregate value of N appears to be tightly related to the aggregate gross output to value added ratio (which is 2.13 in the WIOD) resonates with the theoretical results in the Input-Output model of Fally (2012). Yet, we should stress that Fally's result does not apply in our setting: by appropriate choices of α_n, a variant of our model with a large number of stages could be made consistent with gross output to value added ratios in the neighborhood of 2. Doing so, however, would demand setting relatively high values for the value-added intensity parameters α_2 and α_3, but those high values would in turn generate excessively high asymmetries between the diagonal elements of the input and final output matrices. Because, our GMM estimation penalizes those excessive deviations, we estimate a relatively low value of α_3 ($\alpha_3 = 0.16$), and a very large value for α_2 ($\alpha_2 = 1$), which effectively

[29] Note that the model with $N = 3$ involves an additional summatory for $n = 1$ in the country level index Θ_j even when this stage is shut down. Hence, the calibrated T_j's for $N > 2$ should equal $1/J^{N-2}$ times the T_j's for the benchmark calibration with $N = 2$.

shuts down the most upstream stage.

Some readers might still object that recovering the same estimated values for $N > 2$ is not synonymous with correctly identifying $N = 2$. For example, the moments that we target may be misspecified or not contain sufficient information for backing out the correct N. We next show through simulations that there is a precise sense in which recovering the same calibration for $N \geq 2$ implies that the true N is indeed equal to 2 and that $N > 2$ can be rejected.

Let us work with several values for the chain length ranging from $N = 1$, which is the model of Eaton and Kortum (2002), all the way to $N = 4$, and for a set of $J = 5$ imaginary countries. For each N, we simulate a set of primitives of the model and compute the general equilibrium. We then take the resulting simulated WIOT entries and apply our GMM estimation method with the exact same four sets of moments as above. Furthermore, for each true value of N we run our calibration for various possible values for the number of stages, i.e., $\hat{N} = 1, 2, 3, 4$. The spirit of the exercise is thus to examine whether our estimation method can successfully recover the true value of N. We perform this exercise 100 times for each N, so this amounts to 1600 calibrations in total (100 simulations for each N and four estimations per simulation, for each value $\hat{N} = 1, 2, 3, 4$).

Figure 7 plots our simulation results split into four panels, one for each true value of N. In each panel, the $x-$axis plots the value of the objective function that the calibration minimizes, i.e. the difference between the observed and estimated moments, while the $y-$axis plots the sum of squared percentage differences between the true values of the parameters underlying the simulated data and those estimated in the calibration (note the log-scale on both axis). In a nutshell, a lower value in the $x-$axis implies that our calibration is fitting the data more accurately while a lower value in the $y-$axis implies that our calibrated parameters are closer to the true parameters. Obviously, estimations that place us in the bottom left corner of each plot are particularly accurate.

Focus first on the top left scatterplot for $N = 1$, and note that the calibrations for all $\hat{N}$ do very well. This should not be surprising since the model with $\hat{N} > N$ is more flexible and thus nests a model with a lower $\hat{N}$. A crucial observation, however, is that all the points in the scatterplot lie in the bottom left of the graph, implying that a good fit of the moments occurs only if the true parameters are recovered. Turn next to the bottom right scatterplot for $N = 4$. In this case only the estimates for the empirical model with $\hat{N} = 4$ fit the data well, and notice that the true parameters are again recovered. In that same scatter plot, it is clear that the empirical model with $\hat{N} = 3$ does better than $\hat{N} = 2$, and both do better than $\hat{N} = 1$. This is also obvious since lower $\hat{N}$ implies less degrees of freedom. The key takeaway from Figure 7 is that in order to recover the correct N one need only have $\hat{N} \geq N$ and the calibration will recover the correct parameters regardless of the particular value of $\hat{N}$. This appears to be analogous to what occurs in our datasets with $N = 2$ and thus, to the extent that the data generating process behind the observable data is consistent with our model, we are able to reject $N > 2$.

It is important to stress, however, that our identification of N relies heavily on our assumption that the matrix of trade costs τ_{ij} and the vector of technology levels T_j is common for inputs and final goods. For example, one can show that an extension of the Eaton and Kortum (2002)

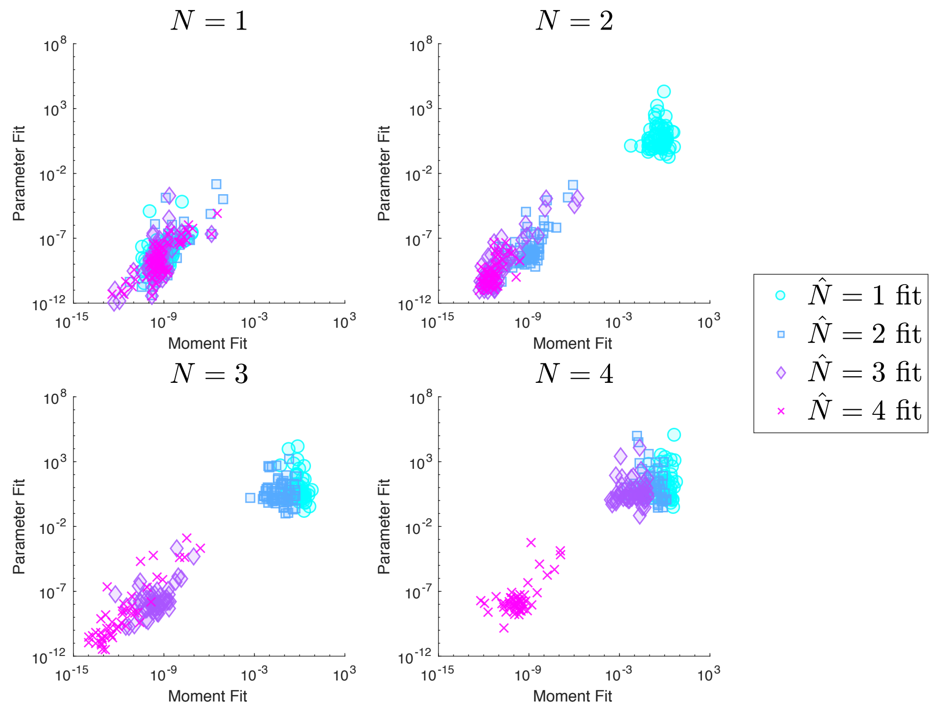

Figure 7: Calibration of N through simulations.

framework without multi-stage production (i.e., $N = 1$) could be calibrated to *exactly* match a WIOT, provided that one allows for (i) cross-country variation in value added shares γ_j, and (ii) arbitrary and potentially asymmetric trade costs for inputs and final goods. Intuitively, one could choose appropriate trade costs matrices τ_{ij}^X and τ_{ij}^F to reproduce the observed asymmetries in the input and output matrices π_{ij}^X and π_{ij}^F. The vector γ_j could then be set to ensure that the GO/VA ratios across countries are exactly nailed, while the technology parameters T_j could be chosen to match the observed cross-country variation in GDP levels. In sum, the data we use cannot reject $N = 1$ if one allows enough flexibility in the modeling of input and output trade costs.

7 Counterfactuals

Having estimated the fundamental parameters of the model, we next explore how counterfactual changes in trade costs, holding other parameters constant, alter the entries of world Input-Output tables, thereby affecting the real income and positioning of countries in GVCs.

Autarky and Zero Gravity

We begin by revisiting two focal counterfactual exercises in quantitative international trade, namely an increase in trade costs large enough to bring back autarky, and a complete elimination of trade barriers. Both of these counterfactuals are extreme in nature, but they are useful in understanding some distinctive features of our framework.

The real income gains of trade relative to autarky can be computed with the formula $\left(\pi_{jN}\right)^{-1/(\theta\gamma_j)} - 1$, as indicated by equation (21), although π_{jN} is not directly observable in the data and needs to be inferred from our model. For the sample of countries in the WIOD, the gains from trade range from a value of 3.3 percent for the United States to 75.9 percent for Luxembourg. The left-panel of Figure 8 plots these real income gains for the largest 25 economies in the WIOD sample.[30]

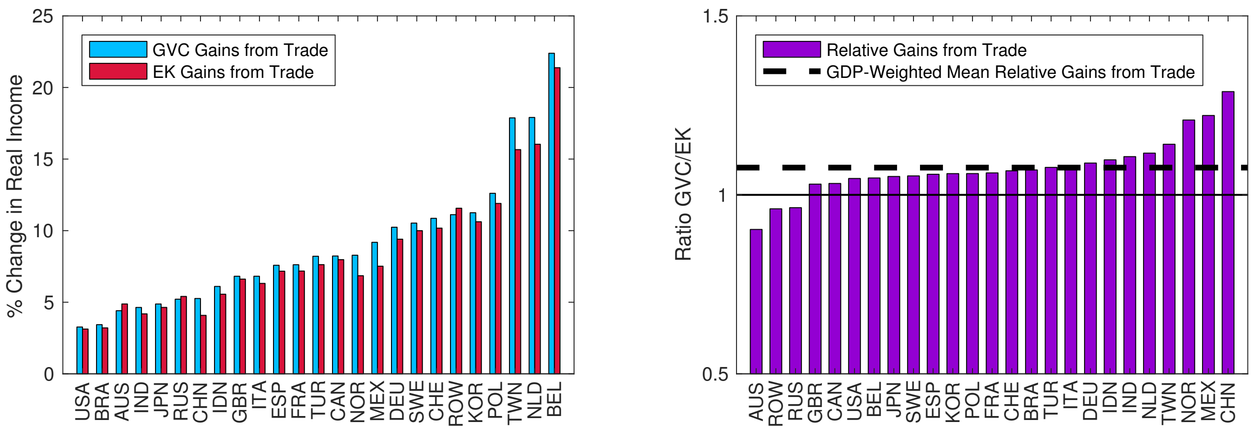

Figure 8: Gains from trade relative to autarky in GVC model ($N = 2$) versus EK model ($N = 1$), WIOD sample.

The figure also compares these gains (labeled 'GVC Gains from Trade') with those obtained in a comparison model without multi-stage production (labeled 'EK Gains from Trade') calibrated to match the WIOD for the year 2014. This comparison model is a modified Eaton and Kortum (2002) framework, with input trade reflecting roundabout production, but with cross-country variation in value added shares γ_j, and differential (and potentially asymmetric) trade costs for inputs and final goods. As mentioned at the end of section 6, by an appropriate choice of parameters, such a model can always exactly match a WIOT. Furthermore, similarly to Arkolakis et al. (2012), the real income losses from going to autarky can be computed using the formula $\left(\hat{\pi}^F_{jj}\left(\hat{\pi}^X_{jj}\right)^{1/\gamma^{EK}_j - 1}\right)^{-1/\theta_{EK}} - 1$, where $\gamma^{EK}_j = GO_j/VA_j$ is the value added to gross output ratio in country j, and where the variables with hats can be read off the data as in (23). As explained in section 4.4, the value of θ relevant for this Eaton and Kortum (2002) model (i.e., θ_{EK} in the formula) is naturally smaller than the one relevant for our framework (i.e., $\theta = 5$), since θ_{EK} here corresponds to the elasticity of overall trade, while θ in our GVC model corresponds to the trade elasticity for only final good trade. Using our estimate of α_2 and the relative prevalence of final-good trade, sequential input trade, and 'roundabout' input trade in our structural estimation leads us to calibrate $\theta_{EK} = 4.635$, which is very much consistent with available estimates of the overall trade elasticity (see footnote 25).

With this background in mind, the left panel of Figure 8 shows that our model with GVCs generates gains from trade that are generally higher than those emanating from a comparable

[30] This formula still measures the real income gains from trade in the presence of trade imbalances. The implications for real spending, however, may be quite different since autarky implies a closing of trade imbalances.

model without multi-stage production. The differences are, however, modest. Averaging across all 44 countries in the WIOD, the ratio of the (net) gains from trade in our GVC model versus those in a modified Eaton-Kortum model equals 1.075.[31] These modest differences arise despite the fact that the share π_{jN} of purely domestic GVCs is on average 29% lower than the final-good trade share π_{ij}^F (0.60 versus 0.85). As anticipated in section 4.4, the lower γ_j^{EK} and θ_{EK} (relative to γ_j and θ) are key factors attenuating the difference in the real income gains from trade. The right-panel of Figure 8 shows, however, that there is quite a lot of variation in the understatement of the gains from trade. China and Mexico, two of USA's largest trading partners, are the countries for which the Eaton-Kortum model underestimates these gains the most (by a factor 1.29 and 1.22, respectively). On the other hand, in a world with sequential production, the gains from trade are lower for certain countries, such as Australia and Russia.

So far, we have discussed our benchmark results with the WIOD. When performing counterfactuals with the broader sample of 101 countries and regions in the Eora database, we find similar results. The gains from trade in a world with GVCs are on average a factor 1.188 larger than in a comparable model without multi-stage production. Although this average ratio is larger than in the WIOD sample, we again find substantial heterogeneity in the relative gains across countries (see Figure A.5 in Appendix A.5), with the ratio being smaller for larger economies. As a result the GDP-weighted ratio of gains from trade in our model relative to a comparable Eaton-Kortum model is just 1.123, and quite in line with our WIOD results.[32]

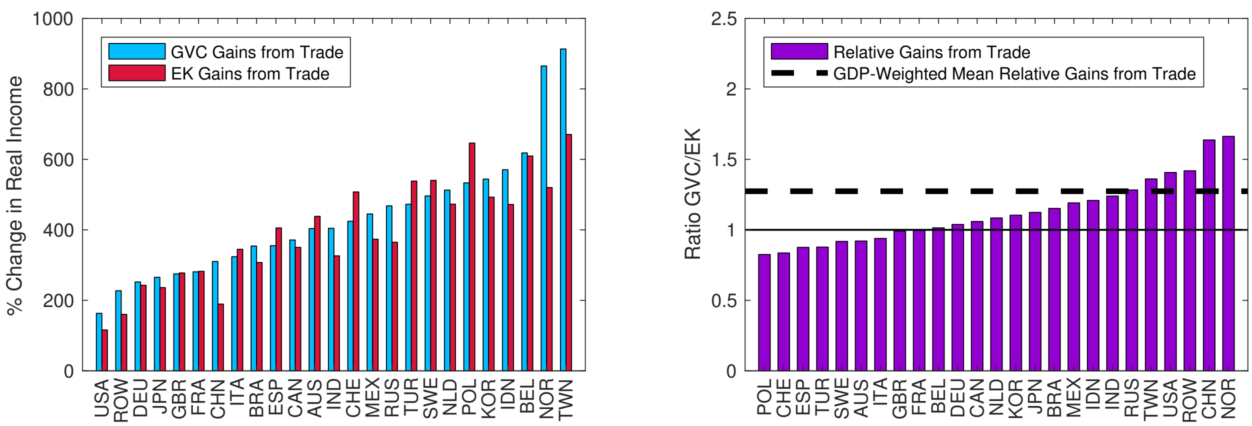

Figure 9: Gains from moving to zero gravity in GVC model ($N = 2$) versus EK model ($N = 1$), WIOD sample.

We next explore the implications of a (hypothetical) complete elimination of trade barriers. The real income consequences of a move to a world with zero gravity are much pronounced. Focusing on the 25 largest economies in the WIOD, Figure 9 shows that these gains range from 163% for the United States to a staggering 913% for Taiwan. Furthermore, these (net) percentage gains

[31] This corresponds to the unweighted average of these ratios. The GDP-weighted average is very similar (1.076) and is depicted as a dashed line in the right panel of Figure 8.

[32] Because of its minuscule own trade share, Ethiopia's gains from trade are extremely large (see the Online Appendix), so we remove this country when computing both the unweighted and weighted average welfare gains.

are on average a factor 1.099 higher than in a model without multi-stage production. For some countries, such as Norway or China, the modified EK model underestimates the real income gains by a very large factor (1.66 and 1.64, respectively). Furthermore, the differences are greater for richer countries. Overall, the GDP-weighted average of these ratio is 1.274, and appears as a dashed line in the right panel of Figure 9.

When repeating this exercise for the sample of countries in Eora, we find that the real income gains with GVCs are about one third larger than an Eaton-Kortum model without sequential production, with the average ratio equalling 1.303. Yet, this average masks substantial variation across countries and continents. Figure 10 breaks these real income gains by continent. As is clear, the Eaton-Kortum model underestimates the real income gains the most for Africa and the Middle East. In this case, the downward bias in the predicted income gains is uncorrelated with country income size, and the GDP-weighted average ratio is 1.345, which is similar to the one found with the WIOD. In Appendix A.5, we provide more details on the counterfactual exercises using the Eora dataset (see, in particular, Figure A.6).[33]

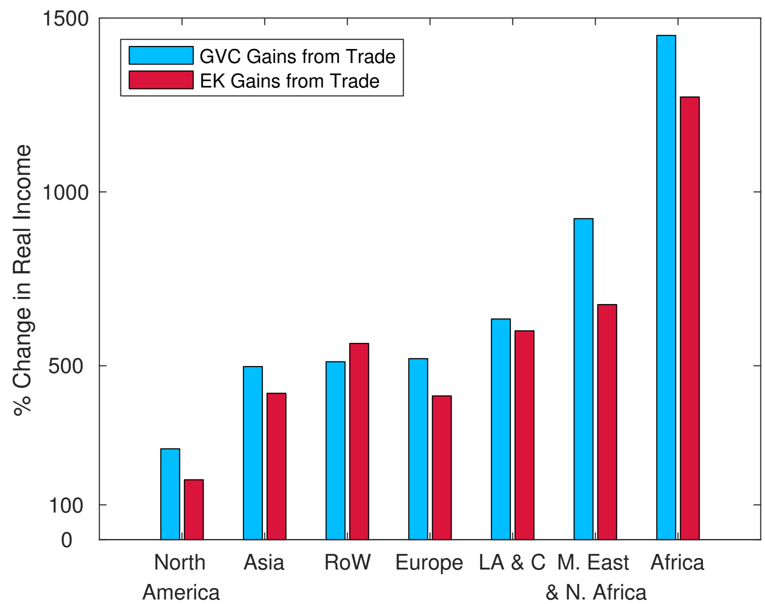

Figure 10: Gains from moving to zero gravity in GVC model ($N = 2$) versus EK model ($N = 1$), Eora sample

A Fifty Percent Reduction in Trade Barriers

We next consider a less extreme counterfactual associated with international trade costs falling by fifty percent, i.e., $\tau'_{ij} = 1+0.5(\tau_{ij}-1)$. We focus on studying the implications of this change for the equilibrium positioning of countries in GVCs in the WIOD sample. The real income implications of this change are reported in the Online Appendix.

[33] In the Online Appendix, we report the real income implications of these counterfactuals for all countries in the WIOD and Eora databases.

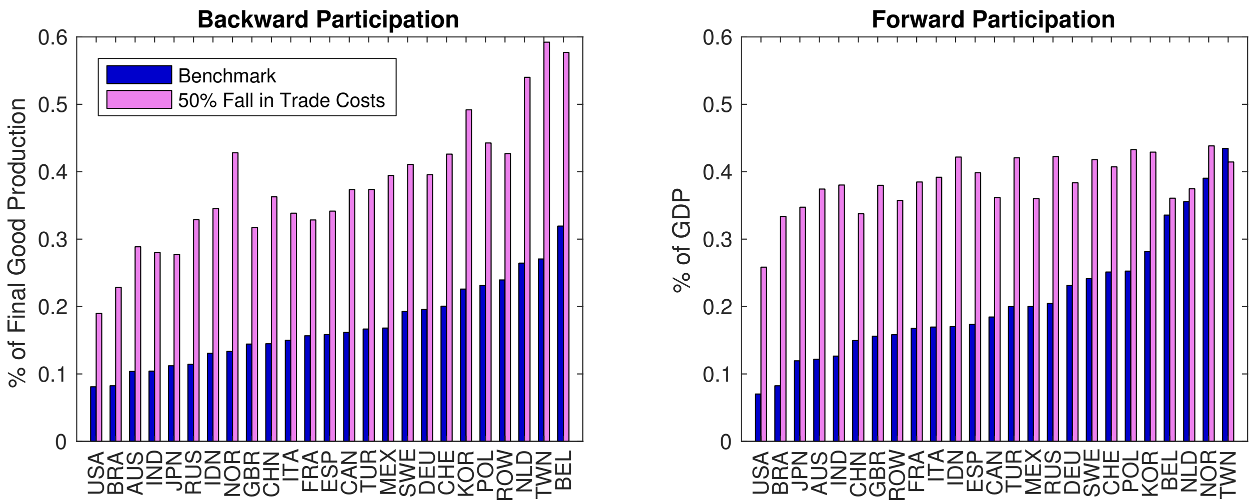

Figure 11: Change in GVC participation following a 50% trade cost reduction.

We begin in Figure 11 by plotting the resulting increase in backward and forward GVC participation in the largest 25 economies in the WIOD sample. As a reminder, the backward GVC participation index measures the extent to which a country's production of final goods uses imported inputs, while the forward GVC participation index measures the share of domestic value added that is exported embodied in intermediate goods. As Figure 11 indicates, both GVC participation indices increase markedly for all countries, but more so for countries that begin with small participation indices. For instance, the world's largest economy – the United States – is the least integrated according to both indices, both before and after this trade cost reduction, but its backward GVC participation doubles in size, while its forward GVC participation index more than triples in size.[34] The fifty percent reduction in trade barriers also shifts the relative positioning of countries in GVCs. For example, the Netherlands increases its backward participation substantially, with little impact on its forward participation rate and, as a result, the Netherlands becomes a more downstream producer in GVCs. In contrast, the USA moves upstream in GVCs, given the larger impact of the trade cost reduction on its forward GVC participation index. When repeating this exercise with the Eora dataset, we find very similar results (see Figure A.7 in the Appendix).

As discussed in the partial equilibrium example in section 2.4, the effects of trade cost reductions on the formation of regional versus global value chains is non-monotonic and depends on the degree of initial GVC integration. This result continues to hold in our general equilibrium model. To illustrate this, Figure 12 decomposes the change in the USA's GVC participation indices into seven bilateral indices related to USA's GVC participation indices vis-a-vis itself, its major trading partners and the main regions of the world. To give a precise example, the USA's backwards GVC participation with Canada equals the share of US production of final goods that is accounted for by Canadian value added embodied in imported intermediate inputs (used for US final-good

[34] The dispersion in GVC participation across countries falls dramatically in this counterfactual scenario since $\theta > 1$ implies that the variance in $\tau_{ij}^{-\theta}$ is higher when iceberg trade costs are at a higher level.

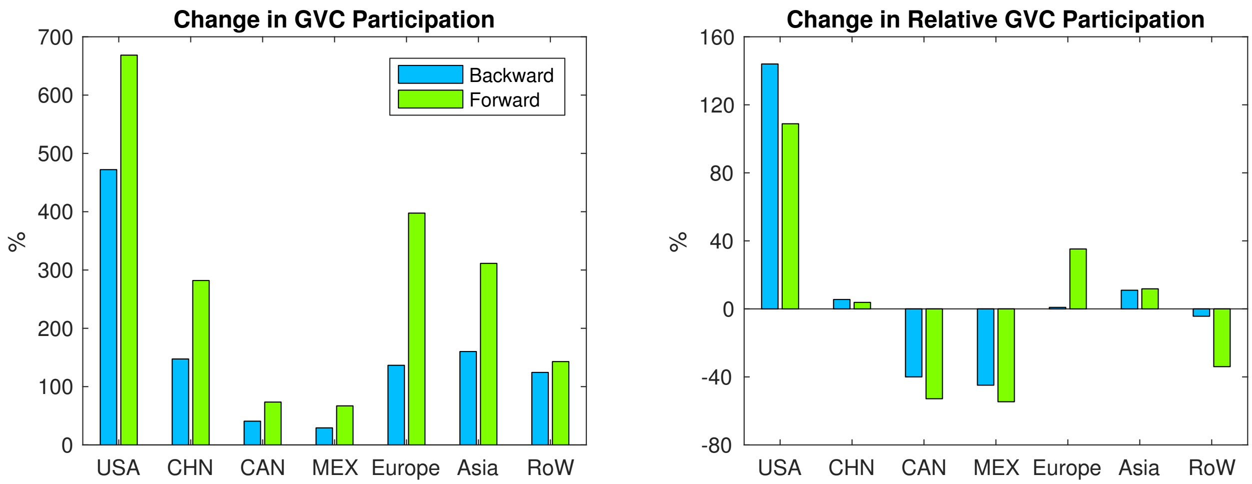

Figure 12: Change in USA bilateral GVC participation following a 50% trade cost reduction.

production), while the USA's forward GVC participation with Canada represents the share of US value added that is exported from the USA embodied in intermediate goods and is eventually consumed in Canada.

The left panel of Figure 12 then shows that a 50 percent reduction in trade barriers would naturally increase the USA's GVC participation with all regions of the world. Yet, the increase would be smallest for NAFTA countries (other than the USA). This indicates that most of the resulting GVC integration would be global rather than regional. Although the growth in the USA's GVC participation with itself is remarkable, we should stress that this does not reflect an increase in domestic value chains. Instead, this reflects an increase in the extent to which (i) US production of final goods uses domestic value added that was exported and re-imported upstream, and (ii) US value added that was exported but later re-imported and eventually consumed as final goods in the USA. The large growth rate in these USA-USA bilateral indices is largely explained by the fact that they start from a very low level (0.2% for the backward index and 0.4% for the forward index).

To further illustrate these differences, the right panel of Figure 12 plots the change in each bilateral GVC participation as a share of USA's total GVC participation. This graph further confirms that relative GVC participation actually falls for the NAFTA countries while increasing for China, Europe and other Asian countries (as well as the USA itself). When repeating this same exercise with the Eora sample of countries, our results are largely unchanged (see Figure A.8 in the Appendix).

Regional versus Global Value Chain Integration

The patterns unveiled in Figure 12 resonate with those in our partial equilibrium example in Figure 2, in which we emphasized that while the relative importance of global GVC integration monotonically increases when trade costs fall, the relative importance of regional GVC participation initially rises but eventually falls when trade costs are lowered sufficiently. Indeed, Figure 12 appears to indicate that current trade costs are at the level at which further reductions will boost global integration relative to regional integration.

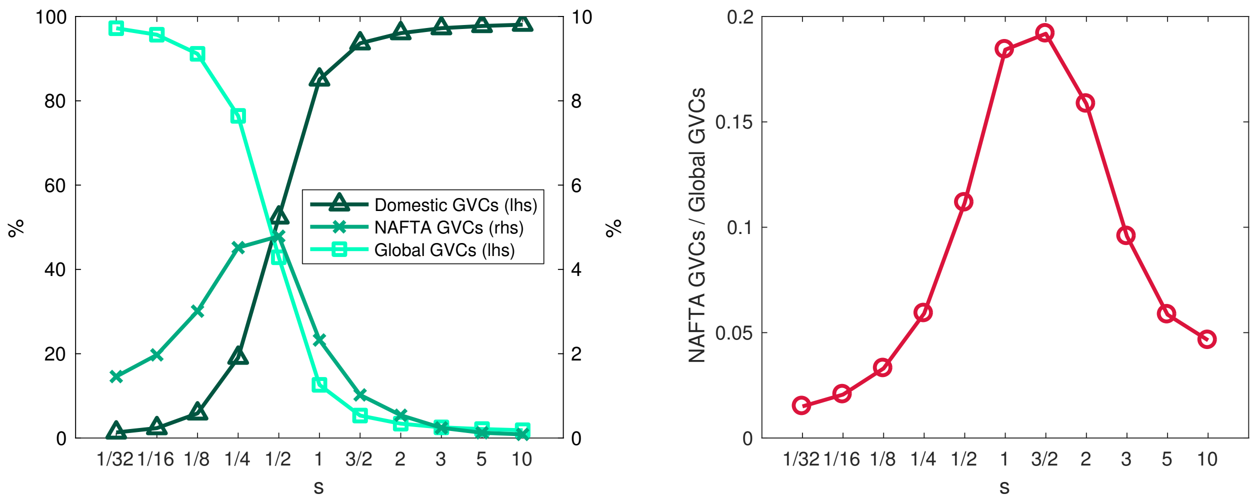

Figure 13: Regional vs Global Integration.

To further analyze the non-monotonic relation between regional and global integration, and in the spirit of section 2.4, we next explore the relative importance of domestic, regional and global value chains across several trade equilibria defined by a value of s such that $\tau'_{ij} = 1 + s(\tau_{ij} - 1)$, with τ_{ij} being our calibrated trade costs for the WIOD sample in 2014. Focusing on our estimated global economy with $N = 2$, we define a domestic GVC as $\boldsymbol{\ell}_d^{USA} = \{USA, USA\}$ and associate the prevalence of *domestic* value chains in overall US consumption with the share $\pi_{\boldsymbol{\ell}_d^{USA},USA}$ (see equation (16)). Similarly, we capture the relative prevalence of *regional* (or NAFTA) value chains in overall US consumption by $\sum_{\boldsymbol{\ell}_r^{USA}} \pi_{\boldsymbol{\ell}_r^{USA},USA}$, where $\boldsymbol{\ell}_r^{USA}$ are all chains that only include the USA, Canada or Mexico, with the exception of the chain $\boldsymbol{\ell}_d^{USA}$. Finally, we define the relative prevalence of *global* value chains in US consumption as $\sum_{\boldsymbol{\ell}_g^{USA}} \pi_{\boldsymbol{\ell}_g^{USA},USA}$, where $\boldsymbol{\ell}_g^{USA}$ are all the possible chains that involve at least one country outside of NAFTA. Naturally, the sum of these three relative measures is one. An important caveat is that, due to the use of a bundle of materials at each stage, what we label as domestic and regional value chains actually embody value added from countries outside NAFTA. In fact, for bounded trade costs, our model features no purely domestic value chains. Yet, the above taxonomy is useful for understanding the broad orientation of value chains serving US consumers for different levels of trade costs.

The left panel in Figure 13 plots these three measures for various values of s between $1/32 \simeq 0.031$ and 10. The resemblance of this chart with our partial-equilibrium Figure 2 is quite remarkable and provides evidence that the intuition presented in the partial equilibrium model carries through to this more general setting. Furthermore, the right panel of Figure 13 plots the ratio of the relative importance of regional (NAFTA) versus global value chains for the same values of s. Interestingly, our benchmark equilibrium, $s = 1$, is very close to the point at which the relative importance of regional value chains is maximized. Thus, as anticipated above in Figure 12, further reductions in trade costs will reduce the relative importance of regional global value chains in US consumption. In the Appendix, we demonstrate that the picture that emerges when repeating this exercise with the Eora dataset is again very similar (see Figure A.9).

8 Conclusion

In this paper, we have studied how trade barriers shape the location of production along GVCs. Relative to an environment with free trade, trade costs generate interdepencies in the sourcing decisions of firms. More specifically, when deciding on the location of production of a given stage, firms necessarily take into account where the good is coming from and where it will be shipped next. As a result, instead of solving N location decisions (where N is the number of stages), firms need to solve the much more computationally burdensome problem of finding an optimal *path* of production. Despite these complications, we have proposed tools to feasibly solve the model in high-dimensional environments.

After deriving these results in partial equilibrium, we have developed a multi-stage general-equilibrium model in which countries specialize in different segments of GVCs. We have demonstrated that, due to the compounding effect of trade costs along value chains, relatively central countries gain comparative advantage in relatively downstream stages of production. We have also borrowed from the seminal work of Eaton and Kortum (2002) to develop a tractable quantitative model of GVCs in a multi-country environment with costly trade. Relative to previous quantitative models of multi-stage production, our suggested approach maps more directly to world Input-Output tables, and allows for a straightforward structural estimation of the model. We finally illustrated some distinctive features of the model by performing counterfactual analyses.

Our framework is admittledly stylized and abstracts from many realistic features that we hope will be explored in future work. For instance, we have abstracted from explicitly modeling cross-industry variation in trade costs, the average level of technology of countries, and the length of production chains. In contemporaneous work, de Gortari (2017) develops a multi-industry version of our model – in a manner analogous to the Caliendo and Parro (2015) extension of the Eaton and Kortum (2002) model – and maps it to the industry-level information available in world Input-Output tables. Another potentially interesting avenue for future research would be to introduce scale economies (external or internal) into our analysis. In a previous version of the paper, we explored a variant of our model with external economies of scale featuring a proximity-concentration tradeoff. The interaction of trade costs and scale economies substantially enriches – but also complicates – the analysis. Although our dynamic programming approach is no longer feasible in that setting, the integer linear programming approach developed in the Appendix is still quite powerful in that environment. We believe that variants of that approach could also prove useful in extending our framework to include internal economies of scale and imperfect competition. It would also be interesting to incorporate contractual frictions into our framework and study the optimal governance of GVCs in a multi-stage, multi-country environment.

Beyond these extensions of our framework, we view our work as a stepping stone for a future analysis of the role and scope of man-made trade barriers in GVCs. Although we have focused on an analysis of the implications of exogenously given trade barriers, our theoretical framework should serve as a useful platform to launch a study of the role of trade policies, and of policies more broadly, in shaping the position of countries in value chains. Should countries actively pursue policies that foster their participation in GVCs? Should they implement policies aimed at moving them to particular stages of those chains? If so, what are the characteristics of these particularly appealing segments of GVCs? These are the type of questions we hope to tackle in future research.

References

Adao, Rodrigo, Arnaud Costinot and Dave Donaldson (2017) "Nonparametric Counterfactual Predictions in Neoclassical Models of International Trade," *American Economic Review*, 107(3): 633-89.

Alfaro, Laura, Pol Antràs, Paola Conconi, and Davin Chor (2015), "Internalizing Global Value Chains: A Firm-Level Analysis," NBER Working Paper 21582, September.

Anderson, James E., and Eric van Wincoop (2003), "Gravity with Gravitas: A Solution to the Border Puzzle," *American Economic Review* 93(1): 170-192.

Antràs, Pol, and Davin Chor (2013), "Organizing the Global Value Chain," *Econometrica* 81(6): 2127-2204.

Antràs, Pol, Davin Chor, Thibault Fally, and Russell Hillberry, (2012), "Measuring the Upstreamness of Production and Trade Flows," *American Economic Review Papers & Proceedings* 102(3): 412-416.

Antràs, Pol, Teresa Fort, and Felix Tintelnot (forthcoming), "The Margin of Global Sourcing: Theory and Evidence from U.S. Firms," *American Economic Review.*

Antràs, Pol, and Robert W. Staiger (2012), "Offshoring and the Role of Trade Agreements," *American Economic Review* 102, no. 7: 3140-3183.

Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare (2012), "New trade models, same old gains?" The American Economic Review 102.1: 94-130.

Baldwin, Richard, and Anthony Venables (2013), "Spiders and Snakes: Offshoring and Agglomeration in the Global Economy," *Journal of International Economics* 90(2): 245-254.

Blanchard, Emily, Chad Bown, and Robert Johnson (2016), "Global Supply Chains and Trade Policy", mimeo Dartmouth College.

Brancaccio, Giulia, Myrto Kalouptsidi, and Theodore Papageorgiou (2017), "Geography, Search Frictions and Trade Costs," mimeo Harvard University.

Caliendo, Lorenzo, and Fernando Parro (2015), "Estimates of the Trade and Welfare Effects of NAFTA," *Review of Economic Studies* 82 (1): 1-44.

Costinot, Arnaud (2009), "An Elementary Theory of Comparative Advantage," *Econometrica* 77:4, pp. 1165-1192.

Costinot, Arnaud, and Andrés Rodríguez-Clare (2015), "Trade Theory with Numbers: Quantifying the Consequences of Globalization" in *Handbook of International Economics*, vol. 4: pp. 197-261.

Costinot, Arnaud, Jonathan Vogel, and Su Wang, (2013), "An Elementary Theory of Global Supply Chains," *Review of Economic Studies* 80(1): 109-144.

de Gortari, Alonso (2017), “Disentangling Global Value Chains,” mimeo Harvard University.

Dekle, Robert, Jonathan Eaton, and Samuel Kortum (2008), “Global rebalancing with gravity: measuring the burden of adjustment.” IMF Economic Review 55.3: 511-540.

Dixit, Avinash, and Gene Grossman, (1982), “Trade and Protection with Multistage Production,” *Review of Economic Studies* 49(4): 583-594.

Ethier, Wilfred J. (1984), “Higher Dimensional Issues in Trade Theory,” Ch. 3 in *Handbook of International Economics*, Grossman, Gene M., and Kenneth Rogoff (Eds), Vol. 1.

Eaton, Jonathan and Samuel Kortum, (2002), “Technology, Geography, and Trade,” *Econometrica*, 70:5, 1741-1779.

Fally, Thibault (2012), “Production Staging: Measurement and Facts” mimeo University of Colorado Boulder.

Fally, Thibault, and Russell Hillbery (2015), “A Coasian Model of International Production Chains,” mimeo UC Berkeley.

Harms, Philipp, Oliver Lorz, and Dieter Urban, (2012), “Offshoring along the Production Chain,” *Canadian Journal of Economics* 45(1): 93-106.

Head, Keith, and Philippe Mayer (2014), “Gravity Equations: Workhorse, Toolkit, and Cookbook,” Ch. 3 in *Handbook of International Economics*, Gopinath, Gita, Elhanan Helpman and Kenneth Rogoff (Eds), Vol. 4.

Head, Keith, and John Ries (2001), “Increasing Returns Versus National Product Differentiation as an Explanation for the Pattern of U.S.-Canada Trade,” *American Economic Review*, 91(4): 858-876.

Hummels, David, Jun Ishii, and Kei-Mu Yi (2001), “The Nature and Growth of Vertical Specialization in World Trade,” *Journal of International Economics* 54 (1): 75–96.

Johnson, Robert, and Andreas Moxnes (2013), “Technology, Trade Costs, and the Pattern of Trade with Multi-Stage Production,” mimeo.

Johnson, Robert C. and Guillermo Noguera (2012), “Accounting for Intermediates: Production Sharing and Trade in Value Added,” *Journal of International Economics* 86(2): 224-236.

Johnson, Robert (2014), “Five Facts about Value-Added Exports and Implications for Macroeconomics and Trade Research,” *Journal of Economic Perspectives*, 28(2), pp. 119-142.

Lenzen, M., Kanemoto, K., Moran, D., Geschke, A. (2012), “Mapping the Structure of the World Economy,” *Environmental Science and Technology*, 46(15) pp 8374-8381.

Lenzen, M., Moran, D., Kanemoto, K., Geschke, A. (2013), “Building Eora: A Global Multi-regional Input-Output Database at High Country and Sector Resolution,” *Economic Systems Research*, 25:1, 20-49.

Kikuchi, Tomoo, Kazuo Nishimura, and John Stachurski (2014), "Transaction Costs, Span of Control and Competitive Equilibrium," mimeo.

Kohler, Wilhelm, (2004), "International Outsourcing and Factor Prices with Multistage Production," *Economic Journal* 114(494): C166-C185.

Koopman, Robert, William Powers, Zhi Wang, and Shang-Jin Wei (2010), "Give Credit Where Credit Is Due," NBER Working Paper 16426, September.

Koopman, Robert, Zhi Wang, and Shang-Jin Wei (2014), "Tracing Value-Added and Double Counting in Gross Exports", *American Economic Review* 104 (2): 459-94.

Kremer, Michael, (1993), "The O-Ring Theory of Economic Development," *Quarterly Journal of Economics* 108(3): 551-575.

Melitz, Marc J, and Stephen J Redding (2014), "Missing Gains from Trade?" *American Economic Review* 104 (5): pp. 317-21.

Morales, Eduardo, Gloria Sheu, and Andrés Zahler (2014), "Gravity and Extended Gravity: Using Moment Inequalities to Estimate a Model of Export Entry," NBER Working Paper No. 19916.

Ossa, Ralph (2015), "Why Trade Matters After All," *Journal of International Economics* 97(2): pp. 266-277.

Sanyal, Kalyan K., and Ronald W. Jones, (1982), "The Theory of Trade in Middle Products," *American Economic Review* 72(1): 16-31.

Timmer, Marcel P., Abdul Azeez Erumban, Bart Los, Robert Stehrer, and Gaaitzen J. de Vries (2014), "Slicing Up Global Value Chains." *Journal of Economic Perspective*s 28, no. 2: 99-118.

Timmer, Marcel P., Eric Dietzenbacher, Bart Los, Robert Stehrer, and Gaaitzen J. de Vries (2015), "An Illustrated User Guide to the World Input–Output Database: the Case of Global Automotive Production," *Review of International Economics*, 23: 575–605.

Tintelnot, Felix (2017), "Global Production with Export Platforms," *Quarterly Journal of Economics,* 132 (1): 157-209.

Tyazhelnikov, Vladimir (2016), "Production Clustering and Offshoring," mimeo UC Davis.

Wang, Zhi, Shang-Jin Wei, Xinding Yu, Kunfu Zhu (2017), "Measures of Participation in Global Value Chains and Global Business Cycles," NBER Working Paper 23222, March.

Yi, Kei-Mu, (2003), "Can Vertical Specialization Explain the Growth of World Trade?" *Journal of Political Economy* 111(1): 52-102.

A Appendix

A.1 Increasing Trade-Cost Elasticity

Define $\tilde{p}^{n-1}_{\ell(n)}(\boldsymbol{\ell}) = p^{n-1}_{\ell(n-1)}(\boldsymbol{\ell})\,\tau_{\ell(n-1)\ell(n)}$ as the price paid in $\ell(n)$ for the good finished up to stage $n-1$ in country $\ell(n-1)$, so that we can express the sequential unit cost function as

$$p^n_{\ell(n)}(\boldsymbol{\ell}) = g^n_{\ell(n)}\left(c_{\ell(n)}, \tilde{p}^{n-1}_{\ell(n)}(\boldsymbol{\ell})\right).$$

Define the elasticity of $p^F_j(\boldsymbol{\ell})$ with respect to the trade costs that stage n's production faces as

$$\beta^j_n = \frac{\partial \ln p^F_j(\boldsymbol{\ell})}{\partial \ln \tau_{\ell(n)\ell(n+1)}},$$

with the convention that $\ell(N+1) = j$ so that β^j_N is the elasticity of $p^F_j(\boldsymbol{\ell})$ with respect to the trade costs faced when shipping assembled goods to final consumers in j. Because $\tau_{\ell(n)\ell(n+1)}$ increases $\tilde{p}^n_{\ell(n+1)}(\boldsymbol{\ell})$ with a unit elasticity, the following recursion holds for all $n' > n$

$$\frac{\partial \ln p^{n'+1}_{\ell(n'+1)}(\boldsymbol{\ell})}{\partial \ln \tau_{\ell(n)\ell(n+1)}} = \frac{\partial \ln p^{n'+1}_{\ell(n'+1)}(\boldsymbol{\ell})}{\partial \ln \tilde{p}^{n'}_{\ell(n'+1)}(\boldsymbol{\ell})} \frac{\partial \ln p^{n'}_{\ell(n')}(\boldsymbol{\ell})}{\partial \ln \tau_{\ell(n)\ell(n+1)}}.$$

At the same time, the unit cost elasticity at stage $n+1$ satisfies

$$\frac{\partial \ln p^{n+1}_{\ell(n+1)}(\boldsymbol{\ell})}{\partial \ln \tau_{\ell(n)\ell(n+1)}} = \frac{\partial \ln p^{n+1}_{\ell(n+1)}(\boldsymbol{\ell})}{\partial \ln \tilde{p}^{n}_{\ell(n+1)}(\boldsymbol{\ell})}.$$

Hence, the elasticity of finished good prices can be decomposed as

$$\beta^j_n = \prod_{n'=n+1}^{N} \frac{\partial \ln p^{n'}_{\ell(n')}(\boldsymbol{\ell})}{\partial \ln \tilde{p}^{n'-1}_{\ell(n')}(\boldsymbol{\ell})}, \tag{A.1}$$

invoking the convention $\prod_{n'=N+1}^{N} f(n') = 1$ for any function $f(\cdot)$. Constant returns to scale in production implies that the function $g^n_{\ell(n)}$ is homogeneous of degree one. As a result, the elasticity of unit costs with respect to input prices is always less or equal than one, so for all $n > 1$ we have

$$\frac{\partial \ln p^{n}_{\ell(n)}(\boldsymbol{\ell})}{\partial \ln \tilde{p}^{n-1}_{\ell(n)}(\boldsymbol{\ell})} \leq 1,$$

with strict inequality whenever a stage adds value to the product. From equation (A.1), it is then clear that

$$\beta^1_j \leq \beta^2_j \leq \cdots \leq \beta^N_j = 1,$$

with strict inequality when value added is positive at all stages.

A.2 Fighting the Curse of Dimensionality: Dynamic and Linear Programming

When discussing the lead-firm problem in section 2.2, we mentioned that there are J^N sequences that deliver distinct finished good prices $p_j^F(\boldsymbol{\ell})$ in country j. Hence, solving for the optimal sequences $\boldsymbol{\ell}^j$ for all j by brute force requires J^{N+1} computations and is infeasible to do when J and N are sufficiently large. However, we show below that use of dynamic programming surmounts this problem by reducing the computation of all sequences to only $J \times N \times J$ computations. Furthermore, in the special case in which production is Cobb-Douglas, the minimization problem can be modeled with zero-one linear programming, for which very efficient algorithms exist.

Dynamic Programming

Define $\boldsymbol{\ell}_n^j \in \mathcal{J}^n$ as the optimal sequence for delivering the good completed up to stage n to producers in country j. This term can be found recursively for all $n = 1, \ldots, N$ by simply solving

$$\boldsymbol{\ell}_n^j = \arg\min_{k \in \mathcal{J}} p_k^n\left(\boldsymbol{\ell}_{n-1}^k\right)\tau_{kj}, \tag{A.2}$$

since the optimal source of the good completed up to stage n is independent of the local factor cost c_j at stage n, of the specifics of the cost function g_j^n, or of the future path of the good. For this same reason, we have written the pricing function p_k^n in terms of the $n-1$ stage sequence $\boldsymbol{\ell}_{n-1}^k$ since it does not depend on future stages of production (though it should be clear that p_k^n will also be a function of the production costs and technology available for producers at that chosen location k). The convention at $n = 1$ is that there is no input sequence so that $\boldsymbol{\ell}_0^k = \varnothing$ for all $k \in \mathcal{J}$ and the price depends only the composite factor cost: $p_k^1(\varnothing) = g_k^1(c_k)$.

The formulation in (A.2) makes it clear that the optimal path to deliver the assembled good to consumers in each country j, i.e., $\boldsymbol{\ell}^j = \boldsymbol{\ell}_N^j$, can be solved recursively by comparing J numbers for each location $j \in \mathcal{J}$ at each stage $n \in \mathcal{N}$, for a total of only $J \times N \times J$ computations.

To further understand this dynamic programming approach, Figure A.1 illustrates a case with 3 stages and 4 countries. Instead of computing $J^N = 64$ paths for each of the four locations of consumption, it suffices to determine the optimal source of (immediately) upstream inputs (which entails $J \times J = 16$ computations at stages $n = 2$ and $n = 3$, and for consumption). In the example, the optimal production path to serve consumers in A, B, and C is $A \to B \to B$, while the optimal path to serve consumers in D is $C \to D \to D$.

Linear Programming

In the special case in which production is Cobb-Douglas, the optimal sourcing sequence can be written as a log-linear minimization problem

$$\boldsymbol{\ell}^j = \arg\min_{\boldsymbol{\ell} \in \mathcal{J}^N} \sum_{n=1}^{N-1} \beta_n \ln \tau_{\ell(n)\ell(n+1)} + \ln \tau_{\ell(N)j} + \sum_{n=1}^{N} \alpha_n \beta_n \ln\left(a_{\ell(n)}^n c_{\ell(n)}\right).$$

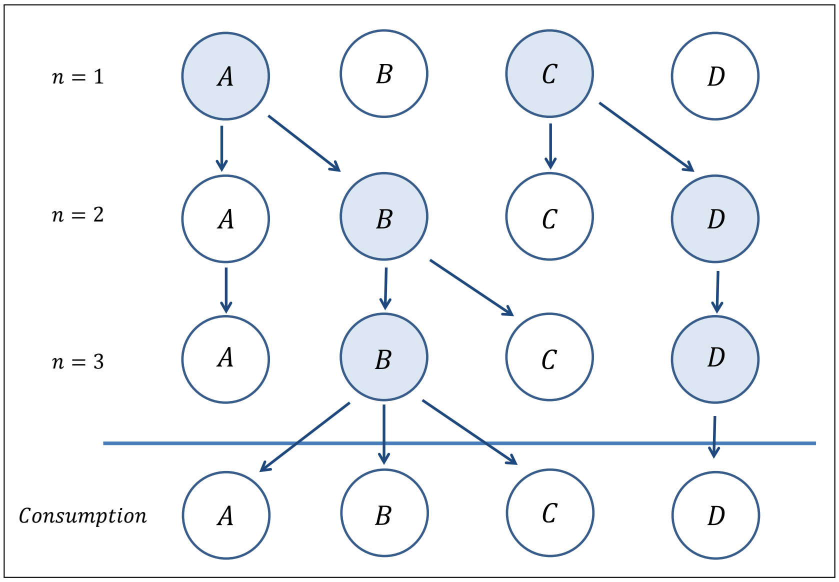

Figure A.1: Dynamic Programming – An Example with Four Countries and Three Stages

This can in turn be reformulated as the following zero-one integer linear programming problem

$$
\begin{aligned}
\boldsymbol{\ell}^j = \arg\min & \sum_{n=1}^{N-1} \beta_n \sum_{k\in\mathcal{J}} \sum_{k'\in\mathcal{J}} \zeta_{kk'}^n \left(\ln \tau_{kk'} + \alpha_n a_k^n c_k\right) + \sum_{k\in\mathcal{J}} \zeta_k^N \left(\ln \tau_{kj} + \alpha_N a_k^N c_k\right) \\
\text{s.t. } & \sum_{k'\in\mathcal{J}} \zeta_{k'k}^n = \sum_{k'\in\mathcal{J}} \zeta_{kk'}^{n+1}, \forall k \in \mathcal{J}, n = 1, \ldots, N-2 \\
& \sum_{k'\in\mathcal{J}} \zeta_{k'k}^{N-1} = \zeta_k^N, \forall k \in \mathcal{J} \\
& \sum_{k\in\mathcal{J}} \zeta_k^N = 1 \\
& \zeta_{kk'}^n, \zeta_k^N \in \{0,1\}.
\end{aligned}
$$

A.3 Decentralized Approach with $N \in \mathbb{N}^+$

This Appendix demonstrates how to generalize our approach with stage-specific randomness and incomplete information to an environment with more than two stages. It should be clear that the input sourcing decisions for the two most upstream stages work in the same way as outlined in section 3.2.B for a general number of stages $N > 2$. Let us quickly recap those decisions. Input producers of good z at the first stage set prices equal to the cost of labor and materials needed to produce a unit of the first-stage good: $p_{\ell(1)}^1(z) = a_{\ell(1)}^1(z)\, c_{\ell(1)}$. Meanwhile, a producer of z at stage $n = 2$ in country j observes the productivity draws of its tier-one input suppliers and thus sources inputs from $\ell_z^j(1) = \arg\min_{\ell(1)\in\mathcal{J}} \left\{ \left(a_{\ell(1)}(z)\, c_{\ell(1)} \tau_{(1)j)}\right)^{1-\alpha_2} \right\}$. However, producers at stage $n > 2$ only observe the productivity draws of their tier-one suppliers (i.e., those at $n-1$), and are forced to use their expectations over the productivity draws of upper tier input suppliers in order to form expectations over the prices at which they will ultimately buy from their tier-one suppliers (at $n-1$). This is because we have assumed that sourcing decisions are made before observing the prices at which tier-one suppliers will ultimately be able to sell at. In other words, when deciding on their optimal input sources, firms producing at stage $n+1$ can only form expectations over the input prices from stage $n-1$ that each of its own possible suppliers producing at stage n faces (or will face).

Let $\ell_z^j(n)$ be the tier-one sourcing decision of a firm producing good z at stage $n+1$ in j. Generalizing the approach in the main text, define the expectation

$$\mathcal{E}_j^n[s] = \mathbb{E}_n\left[\left(p^n_{\ell_z^j(n)}(z)\,\tau_{\ell_z^j(n)j}\right)^s\right],$$

for any $s>0$ and where we have written the expectation with an n subscript indicating that the expectation takes that unit costs (and prices) from stages $1,\ldots,n$ as unobserved. To be fully clear, a firm at $n+2$ observes the productivity draws from stage $n+1$ but does not know previous sourcing decisions. Hence it must form an expectation over the location from which its stage n suppliers source, $\ell_z^j(n)$, and use this to calculate the expected input prices $\mathcal{E}_j^n[s]$. As will become clear in the next paragraph, denoting the expectations for a general $s>0$ is useful since downstream firms between $n+2,\ldots,N$ and final consumers will all use the information on expected input prices at n but in different ways depending on the objective function they seek to minimize.

Substituting in the Cobb-Douglas production process in (2), we can write

$$\mathcal{E}_j^n[s] = \mathbb{E}_n\left[\left(a^n_{\ell_z^j(n)}(z)\,c_{\ell_z^j(n)}\right)^{\alpha_n s}\times\mathcal{E}^{n-1}_{\ell_z^j(n)}\left[(1-\alpha_n)\,s\right]\times\left(\tau_{\ell_z^j(n)j}\right)^s\right].$$

The crucial observation is that to determine expected input prices from stage n a firm must also incorporate expected input prices from stage $n-1$, and so on until input prices from all upstream stages have been incorporated. Note that productivity draws across stages of production are independent, but even more importantly, sourcing decisions across stages of production are also independent. Hence, one can use the law of iterated expectations to compute expected input prices from $n-1$, $\mathcal{E}^{n-1}_{\ell_z^j(n)}[\cdot]$, in the computation of expected prices at n in $\mathcal{E}_j^n[\cdot]$. The latter expectation is over $\ell_z^j(n)$ but once we condition on a specific value for $\ell_z^j(n)$, the expectation $\mathcal{E}^{n-1}_{\ell_z^j(n)}[\cdot]$ is a constant. Finally, note also that this recursion starts at $n=1$ with $\mathcal{E}_j^0[s]=1$ since only labor and materials are used in that initial stage.

Let us next illustrate why these definitions are useful. Consider the optimal sourcing strategies related to procuring the good finished up to stage $n<N$. Given the sequential cost function in (2), the problem faced by a stage $n+1$ producer in j can be written as

$$\ell_z^j(n) = \arg\min_{\ell(n)\in\mathcal{J}}\left\{\left(a^n_{\ell(n)}(z)\,c_{\ell(n)}\right)^{\alpha_n(1-\alpha_{n+1})}\times\mathcal{E}^{n-1}_{\ell(n)}\left[(1-\alpha_n)(1-\alpha_{n+1})\right]\times\left(\tau_{\ell(n)j}\right)^{1-\alpha_{n+1}}\right\}.$$

where the $1-\alpha_{n+1}$ superscript comes from the stage $n+1$ producer wanting to minimize its own expected input price and in which the stage n input price enters its own unit cost to this power.

Meanwhile, final consumers (or local retailers on their behalf) source their goods by solving

$$\ell_z^j(N) = \arg\min_{\ell(N)\in\mathcal{J}} \left\{ \left(a_{\ell(N)}^N (z)\, c_{\ell(N)} \right)^{\alpha_N} \times \mathcal{E}_{\ell(N)}^{N-1} [1-\alpha_N] \times \tau_{\ell(N)j} \right\}.$$

The probability of sourcing inputs from a specific location i at any stage n can be determined by invoking the properties of the Fréchet distribution, given that $1/a_i^n(z)$ is drawn independently (across goods and stages) from a Fréchet distribution satisfying

$$\Pr\left(a_j^n (z)^{\alpha_n \beta_n} \geq a \right) = \exp\left\{ -a^{\theta} (T_j)^{\alpha_n \beta_n} \right\}.$$

In particular, we obtain

$$\Pr\left(\ell_z^j(n) = i\right) = \frac{\left((T_i)^{\alpha_n} \left((c_i)^{\alpha_n} \tau_{ij} \right)^{-\theta} \right)^{\beta_n} \mathcal{E}_i^{n-1} \left[(1-\alpha_n)(1-\alpha_{n+1}) \right]^{-\beta_{n+1}\theta}}{\sum_{l\in J} \left((T_l)^{\alpha_n} \left((c_l)^{\alpha_n} \tau_{lj} \right)^{-\theta} \right)^{\beta_n} \mathcal{E}_l^{n-1} \left[(1-\alpha_n)(1-\alpha_{n+1}) \right]^{-\beta_{n+1}\theta}}.$$

These probabilities can now be leveraged in order to compute expected input prices. Define $\tilde{a}_{ij} = (c_i)^{\alpha_n s} \mathcal{E}_i^{n-1} [(1-\alpha_n) s] (\tau_{ij})^s$ so that $1/\left(a_i^{\alpha_n s} \tilde{a}_{ij}\right) \sim$ Fréchet $\left(T_i^{\alpha_n \beta_n} \tilde{a}_{ij}^{-\frac{\beta_n}{s}\theta}, \frac{\beta_n}{s}\theta \right)$ (note that the above distribution is the special case in which $s = 1-\alpha_{n+1}$). Then using the moment generating formula for the Fréchet distribution, it can be verified that

$$\mathcal{E}_j^n [s] = \left[\sum_{l\in J} T_l^{\alpha_n \beta_n} \tilde{a}_{lj}^{-\frac{\beta_n}{s}\theta} \right]^{-\frac{s}{\beta_n \theta}} \Gamma\left(1 + \frac{\beta_n}{s}\theta \right),$$

where Γ is the gamma function. From this equation it should also be clear why we are denoting $E_j^n[s]$ as a function of s, since as we move down the value chain we need to compute the upstream expectations at different 'moments'.

We are now ready to determine the equilibrium variables: (1) material prices P_j and (2) the distribution of GVCs. Material prices can be derived recursively using our expectations:

$$\begin{aligned} P_j = \left(\mathcal{E}_j^N [1-\sigma] \right)^{\frac{1}{1-\sigma}} &= \left[\sum_{l\in\mathcal{J}} (T_l)^{\alpha_N} \left((c_l)^{\alpha_N} \tau_{lj} \right)^{-\theta} \mathcal{E}_l^{N-1} \left[(1-\alpha_N)(1-\sigma) \right]^{-\frac{\theta}{1-\sigma}} \right]^{-\frac{1}{\theta}} \Gamma\left(1 + \frac{1-\sigma}{\theta} \right) \\ &= \left[\sum_{\ell\in\mathcal{J}} \prod_{n=1}^{N} \left(\left(T_{\ell(n)} \right)^{\alpha_n} \left(\left(c_{\ell(n)} \right)^{\alpha_n} \tau_{\ell(n)\ell(n+1)} \right)^{-\theta} \right)^{\beta_n} \right]^{-\frac{1}{\theta}} \prod_{n=1}^{N} \Gamma\left(1 + \frac{1-\sigma}{\beta_n \theta} \right)^{\frac{1}{1-\sigma}} \end{aligned}$$

Finally, since input decisions from n are independent from the decisions that firms at $n-1$ made then

$$
\begin{aligned}
\pi_{\boldsymbol{\ell}j} &= \Pr\left(\ell_z^j(N) = \ell(N) \middle| \ell_z^{\ell(N)}(N-1) = \ell(N-1)\right) \times \\
&\quad \times \prod_{n=2}^{N-1} \Pr\left(\ell_z^{\ell(n+1)}(n) = \ell(n) \middle| \ell_z^{\ell(n)}(n-1) = \ell(n-1)\right) \times \Pr\left(\ell_z^{\ell(2)}(1) = \ell(1)\right) \\
&= \Pr\left(\ell_z^j(N) = \ell(N)\right) \times \prod_{n=1}^{N} \Pr\left(\ell_z^{\ell(n+1)}(n) = \ell(n)\right) \\
&= \frac{\prod_{n=1}^{N-1}\left(\left(T_{\ell(n)}\right)^{\alpha_n}\left(\left(c_{\ell(n)}\right)^{\alpha_n}\tau_{\ell(n)\ell(n+1)}\right)^{-\theta}\right)^{\beta_n} \times \left(T_{\ell(N)}\right)^{\alpha_N}\left(\left(c_{\ell(N)}\right)^{\alpha_N}\tau_{\ell(N)j}\right)^{-\theta}}{\sum_{\boldsymbol{\ell}'\in\mathcal{J}}\prod_{n=1}^{N-1}\left(\left(T_{\ell'(n)}\right)^{\alpha_n}\left(\left(c_{\ell'(n)}\right)^{\alpha_n}\tau_{\ell'(n)\ell'(n+1)}\right)^{-\theta}\right)^{\beta_n} \times \left(T_{\ell'(N)}\right)^{\alpha_N}\left(\left(c_{\ell'(N)}\right)^{\alpha_N}\tau_{\ell'(N)j}\right)^{-\theta}},
\end{aligned}
$$

which is identical to equation (11) in the main text obtained in the 'randomness-in-the-chain' formulation of technology.

A.4 Proof of Centrality-Downstreamness Nexus

Let $(\tau_{ij})^{-\theta} = \rho_i\rho_j$. In such a case, the probability of country j sourcing through $\boldsymbol{\ell}$ reduces to

$$
\pi_{\boldsymbol{\ell}j} = \frac{\prod_{m=1}^{N}\left(T_{\ell(m)}\left(c_{\ell(m)}\right)^{-\theta}\right)^{\alpha_m\beta_m}\left(\rho_{\ell(m)}\right)^{\beta_{m-1}+\beta_m}}{\sum_{\boldsymbol{\ell}\in\mathcal{J}}\prod_{m=1}^{N}\left(T_{\ell(m)}\left(c_{\ell(m)}\right)^{-\theta}\right)^{\alpha_m\beta_m}\left(\rho_{\ell(m)}\right)^{\beta_{m-1}+\beta_m}}
$$

and is thus independent of the destination country j. The aggregate probability of observing country i in location n can thus be expressed as

$$
\Pr\left(\Lambda_i^n\right) = \sum_{\boldsymbol{\ell}\in\Lambda_i^n}\pi_{\boldsymbol{\ell}j} = \frac{\sum_{\boldsymbol{\ell}\in\Lambda_i^n}\prod_{m=1}^{N}\left(T_{\ell(m)}\left(c_{\ell(m)}\right)^{-\theta}\right)^{\alpha_m\beta_m}\left(\rho_{\ell(m)}\right)^{\beta_{m-1}+\beta_m}}{\sum_{k\in J}\sum_{\boldsymbol{\ell}\in\Lambda_k^n}\prod_{m=1}^{N}\left(T_{\ell(m)}\left(c_{\ell(m)}\right)^{-\theta}\right)^{\alpha_m\beta_m}\left(\rho_{\ell(m)}\right)^{\beta_{m-1}+\beta_m}}. \tag{A.3}
$$

But note that we can decompose this as

$$
\begin{aligned}
\Pr\left(\Lambda_i^n\right) &= \frac{\left(T_i\left(c_i\right)^{-\theta}\right)^{\alpha_n\beta_n}\left(\rho_i\right)^{\beta_{n-1}+\beta_n} \times \sum_{\boldsymbol{\ell}\in\Lambda_i^n}\prod_{m\neq n}\left(T_{\ell(m)}\left(c_{\ell(m)}\right)^{-\theta}\right)^{\alpha_m\beta_m}\left(\rho_{\ell(m)}\right)^{\beta_{m-1}+\beta_m}}{\sum_{k\in\mathcal{J}}\left(T_k\left(c_k\right)^{-\theta}\right)^{\alpha_n\beta_n}\left(\rho_k\right)^{\beta_{n-1}+\beta_n} \times \sum_{\boldsymbol{\ell}\in\Lambda_k^n}\prod_{m\neq n}\left(T_{\ell(m)}\left(c_{\ell(m)}\right)^{-\theta}\right)^{\alpha_m\beta_m}\left(\rho_{\ell(m)}\right)^{\beta_{m-1}+\beta_m}} \\
&= \frac{\left(T_i\left(c_i\right)^{-\theta}\right)^{\alpha_n\beta_n}\left(\rho_i\right)^{\beta_{n-1}+\beta_n}}{\sum_{k\in\mathcal{J}}\left(T_k\left(c_k\right)^{-\theta}\right)^{\alpha_n\beta_n}\left(\rho_k\right)^{\beta_{n-1}+\beta_n}}
\end{aligned}
$$

where the second line follows from the fact that, for GVCs in the sets Λ_i^n and Λ_k^n, the set of all possible paths excluding the location of stage n are necessarily identical (and independent of the country where n takes place), and thus the second terms in the numerator and denominator of the first line cancel out.

For the special symmetric case with $\alpha_n \beta_n = 1/N$ and $\alpha_n = 1/n$ we obtain that

$$\Pr\left(\Lambda_i^n\right) = \frac{\left(T_i \left(c_i\right)^{-\theta}\right)^{\frac{1}{N}} \left(\rho_i\right)^{\frac{2n-1}{N}}}{\sum\limits_{k \in \mathcal{J}} \left(T_k \left(c_k\right)^{-\theta}\right)^{\frac{1}{N}} \left(\rho_k\right)^{\frac{2n-1}{N}}}$$

Now consider our definition of upstreamness

$$U\left(i\right) = \sum_{n=1}^{N} \left(N - n + 1\right) \times \frac{\Pr\left(\Lambda_i^n\right)}{\sum\limits_{n'=1}^{N} \Pr\left(\Lambda_i^{n'}\right)}. \tag{A.4}$$

This is equivalent to the expect distance from final-good demand at which a country will contribute to global value chains. The expectation is defined over a country-specific probability distribution over stages, $f_i\left(n\right) = \Pr\left(\Lambda_i^n\right) / \sum_{n'=1}^{N} \Pr\left(\Lambda_i^{n'}\right)$.

Finally, note that for two countries with $\rho_{i'} > \rho_i$ and two inputs with $n' > n$ we necessarily have

$$\frac{f_{i'}\left(n'\right) / f_{i'}\left(n\right)}{f_i\left(n'\right) / f_i\left(n\right)} = \left(\frac{\rho_{i'}}{\rho_i}\right)^{2\left(n'-n\right)/N} > 1.$$

As a result, the probability functions $f_{i'}\left(n\right)$ and $f_i\left(n\right)$ satisfy the monotone likelihood ratio property in n. As is well known, this is a sufficient condition for $f_{i'}\left(n\right)$ to first-order stochastically dominate $f_i\left(n\right)$ when $\rho_{i'} > \rho_i$. But then it is immediate that $\mathbb{E}_{f_{i'}}\left[n\right] > \mathbb{E}_{f_i}\left[n\right]$, and thus the expected value in (A.4), which is simply $N + 1 - \mathbb{E}_{f_i}\left[n\right]$, will be lower for country i' than for country i when $\rho_{i'} > \rho_i$.

A.5 Further Estimation Results

WIOD for 2014

Table A.1 presents the values of γ_j and T_j for the sample of 44 countries in the WIOD found in our benchmark estimation with $N = 2$.

Table A.1: WIOD Calibration

	γ_j	T_j		γ_j	T_j
Australia	0.93	31.587	Ireland	0.87	0.396
Austria	0.91	6.134	Italy	0.89	10.419
Belgium	0.84	0.789	Japan	0.96	6.997
Bulgaria	0.78	0.006	South Korea	0.72	0.555
Brazil	1.00	0.011	Lithuania	0.95	0.056
Canada	0.96	5.395	Luxembourg	0.52	0.117
Switzerland	0.89	670.238	Latvia	0.79	0.028
China	0.59	0.116	Mexico	1.00	0.001
Cyprus	0.99	0.169	Malta	0.64	0.018
Czech Republic	0.73	0.129	Netherlands	0.88	1.108
Germany	0.92	47.132	Norway	0.93	0.444
Denmark	0.93	7.269	Poland	0.84	0.521
Spain	0.93	5.065	Portugal	0.95	0.163
Estonia	0.81	0.023	Romania	0.85	0.015
Finland	0.88	1.543	Russia	0.90	0.009
France	0.97	19.680	Slovakia	0.77	0.229
Great Britain	0.97	36.013	Slovenia	0.85	0.541
Greece	1.00	0.028	Sweden	0.94	25.031
Croatia	0.94	0.036	Turkey	0.91	0.045
Hungary	0.83	0.008	Taiwan	0.75	0.009
Indonesia	0.94	1.26E-05	USA	1.00	121.919
India	0.97	2.05E-06	Rest of World	0.79	0.009

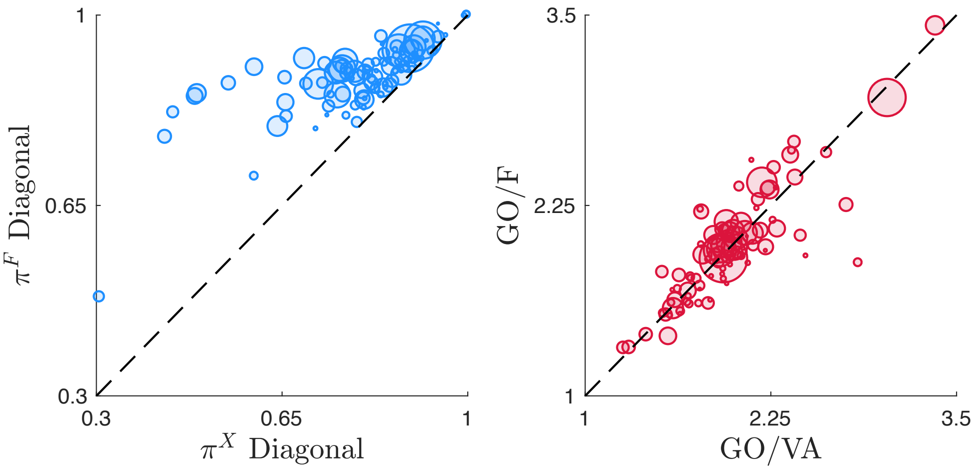

Figure A.2: Some Key Features of the Eora MRIO Dataset

Eora for 2013

Figure A.2 depicts some salient features of the Eora MRIO dataset for the year 2013. The figure is analogous to Figure 4 in the main text, and depicts the same qualitative patterns. The domestic shares are on average higher for final output than for inputs and there is wide dispersion in gross output to value added ratios and gross output to final output ratios, with both ratios being highly positively correlated.

We next turn to the estimation results for the Eora 2013 database when our model is calibrated to the same moments as the WIOD and with $N = 2$. As mentioned in the main text, we find $\alpha_2 = 0.19$. Table A.2 presents the values of γ_j and T_j for the sample of 101 country/regions. Figures A.3 and A.4 illustrate the goodness of fit of our model. As mentioned in the main text, the correlation between model and data is very high, even when considering untargeted moments. Figure A.5 presents the gains from trade with respect to autarky for the GVC and EK models for the largest 25 countries/regions. The gains are on average 19% higher across the full sample (12% when weighting by GDP size). Meanwhile, Figure A.6 presents the gains of a zero gravity world. Real income gains are on average 30% higher in the GVC world (34% when weighting by GDP size) than in a world without sequential production. Finally, Figures A.7 and A.8 present the changes in GVC participation following a 50% reduction in trade costs and Figure A.9 plots the dynamics of regional versus global integration when lowering trade costs; all three graphs look very similar to the ones with the WIOD data.

Table A.2: Eora Calibration

	γ_j	T_j		γ_j	T_j
Afghanistan	0.70	3.79E-05	Israel	0.96	383.437
Eastern Europe	0.79	4.11E-05	Italy	0.88	3.575
Algeria	0.98	3.08E-06	Japan	0.97	102.429
Western Europe	0.72	0.098	Kazakhstan	0.86	0.002
Angola	0.95	0.052	Kenya	0.93	2.80E-05
Latin America & Caribbean	0.85	0.605	Madagascar	0.87	0.001
Argentina	1.00	0.394	Malawi	0.61	1.72E-05
Australia	0.85	270.496	Malaysia	0.85	0.012
Central Europe	0.89	17.669	Mali	0.82	0.002
Central Asia	0.86	0.001	Mexico	1.00	0.001
Middle East & North Africa	0.96	1.079	Morocco	0.96	0.004
Bangladesh	0.94	5.38E-05	Mozambique	0.99	2.95E-06
Belgium	0.77	0.115	Myanmar	0.69	0.967
Benin	0.59	0.001	Nepal	0.90	2.84E-05
South Asia	0.40	0.020	Netherlands	0.80	0.057
Bolivia	0.88	0.162	Niger	0.97	3.33E-08
Sub-Saharan Africa	0.92	3.80E-04	Nigeria	1.00	2.61E-07
Brazil	0.86	0.300	Pakistan	0.90	0.002
East Asia & Pacific	0.79	6.598	Peru	0.89	0.047
Burkina Faso	0.94	1.37E-06	Philippines	0.99	3.69E-05
Burundi	0.87	6.52E-07	Poland	0.81	0.186
Cambodia	0.82	2.38E-04	Portugal	0.87	0.394
Cameroon	0.85	1.99E-04	South Korea	0.51	0.347
Canada	0.91	6.272	Romania	0.81	0.010
Chad	0.84	2.84E-04	Russia	0.93	0.058
Chile	0.91	0.070	Rwanda	0.94	0.001
China	0.63	0.082	Saudi Arabia	1.00	0.058
Colombia	1.00	0.001	Senegal	0.87	0.001
Cuba	0.95	0.004	Singapore	0.68	0.383
Czech Republic	0.69	0.125	Somalia	0.53	4.49E-05
Cote dIvoire	0.98	0.015	South Africa	0.88	0.012
North Korea	0.89	0.002	South Sudan	0.84	0.047
DR Congo	0.86	3.877	Spain	0.84	3.924
Scandinavia	0.89	10.130	Sri Lanka	0.98	0.001
Dominican Republic	0.93	0.004	Sudan	0.91	3.991
Ecuador	0.93	0.012	Syria	0.84	0.865
Egypt	1.00	0.001	Taiwan	0.82	6.672
Eritrea	0.81	1.03E-05	Thailand	0.73	0.383
Ethiopia	0.02	2.22E-05	Tunisia	0.93	0.010
France	0.94	21.699	Turkey	1.00	0.001
Germany	0.82	6.599	Uganda	0.83	1.30E-05
Ghana	0.94	1.81E-04	Ukraine	0.73	0.001
Greece	0.99	0.105	UK	0.90	11.837
Guatemala	0.97	0.316	Tanzania	0.43	8.08E-06
Guinea	0.76	1.83E-06	USA	0.94	133.312
Haiti	0.84	7.87E-05	Uzbekistan	0.95	0.001
Hong Kong	0.65	0.024	Venezuela	0.94	0.030
India	0.87	0.003	Viet Nam	0.69	9.67E-05
Indonesia	0.94	0.004	Yemen	0.91	2.45E-05
Iran	0.97	0.002	Zambia	0.93	1.06E-04
Iraq	1.00	1.65E-10			

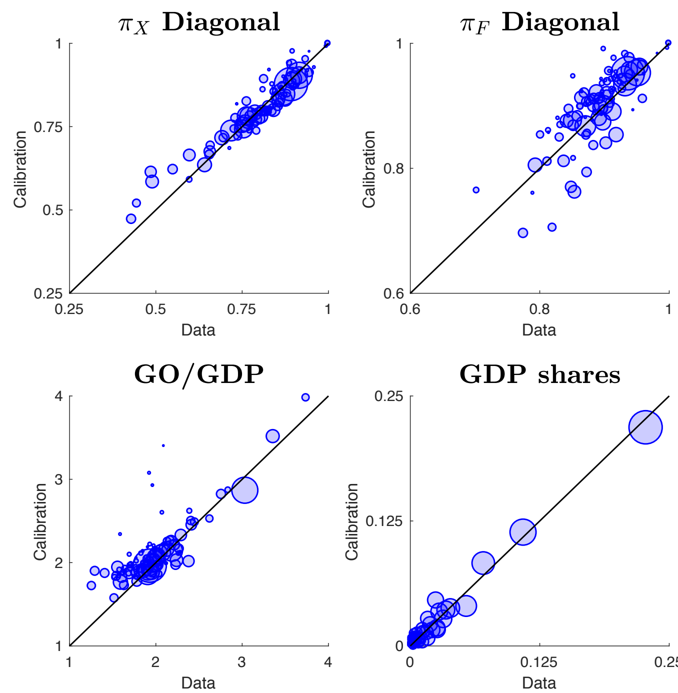

Figure A.3: Eora Targeted Moments.

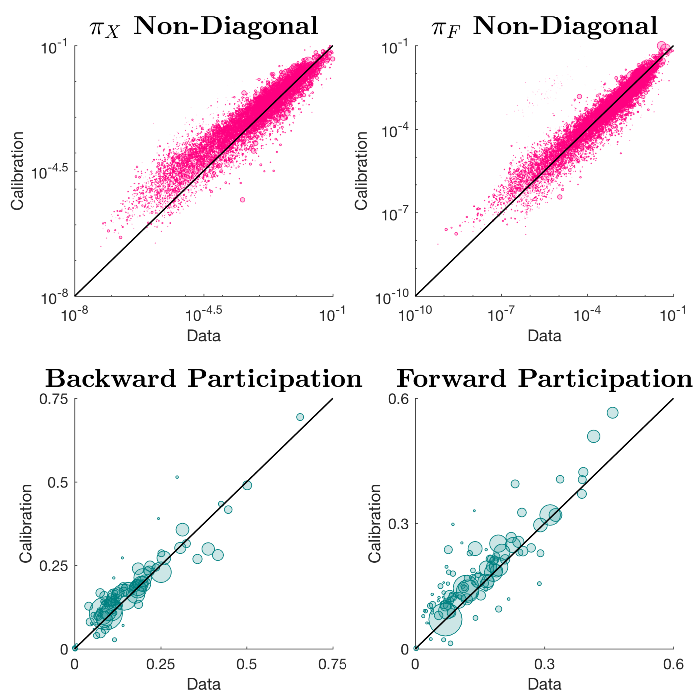

Figure A.4: Eora Untargeted Moments.

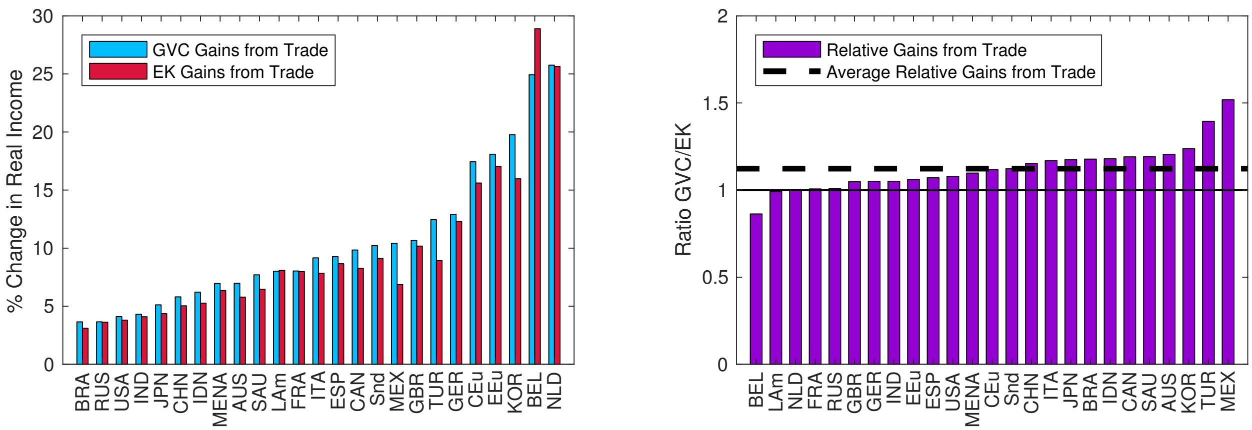

Figure A.5: Gains from trade relative to autarky in GVC model ($N = 2$) versus EK model ($N = 1$), Eora sample.

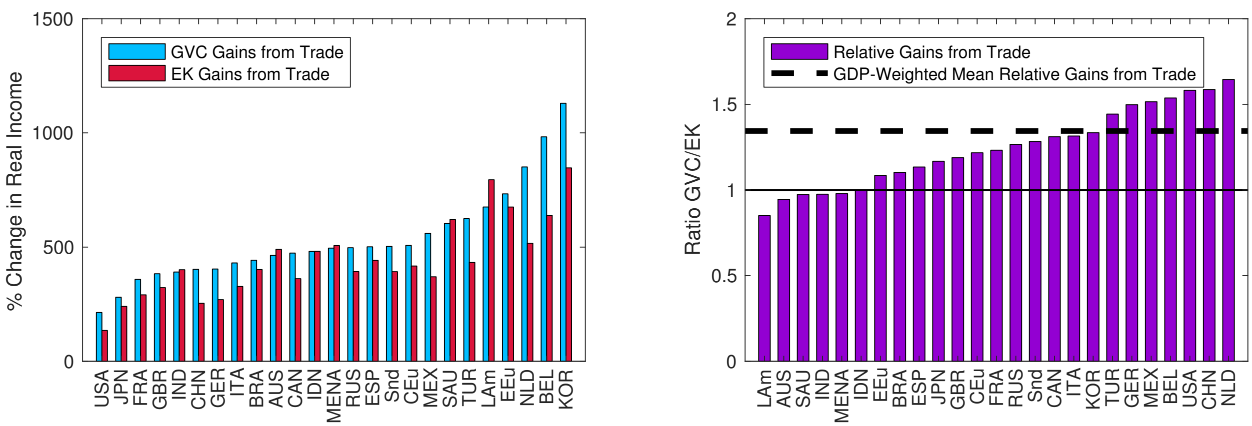

Figure A.6: Gains from moving to zero gravity in GVC model ($N = 2$) versus EK model ($N = 1$), Eora sample.

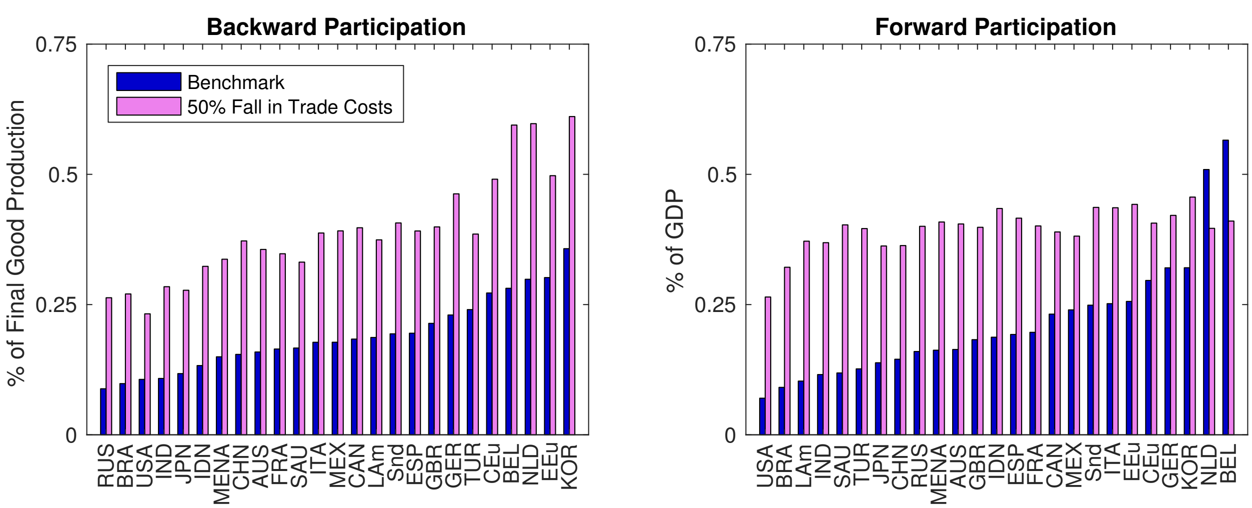

Figure A.7: Change in USA bilateral GVC participation following a 50% trade cost reduction, Eora sample.

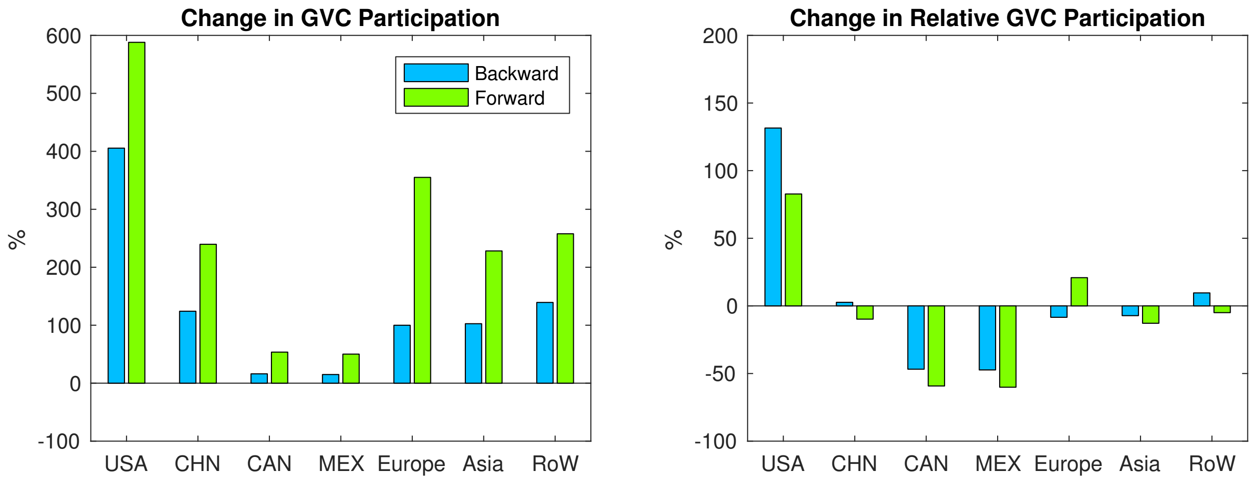

Figure A.8: Change in USA bilateral GVC participation following a 50% trade cost reduction, Eora sample.

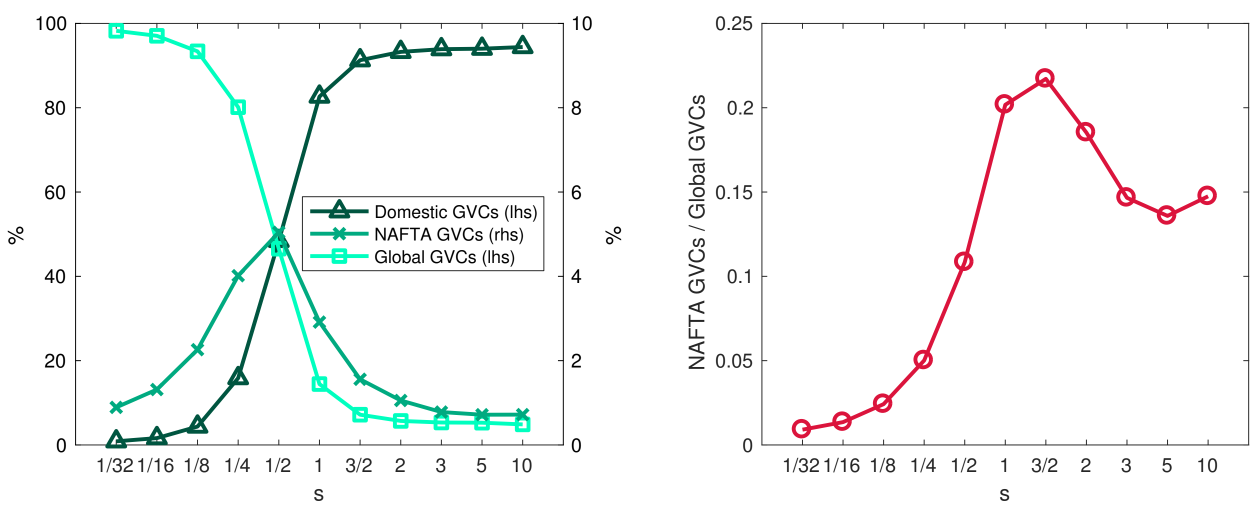

Figure A.9: Regional vs Global Integration, Eora sample.

On the Geography of Global Value Chains

Pol Antràs and Alonso de Gortari

B Online Appendix (Not for Publication)

B.1 The Partial Equilibrium Example without Sequentiality

In this Appendix, we revisit our partial equilibrium example with four countries and four stages in section 2.4, but we consider an alternative scenario without sequentiality. More specifically, we still consider a symmetric Cobb-Douglas technology with four 'stages' contributing to value added, but we assume that these four stages occur simultaneously and are combined into a non-tradeable final good. We continue to focus on serving consumers in country D, so this boils down to a "spider" sourcing model in which assemblers in D choose the optimal source for each of the required four inputs. The rest of the specifics of the exercise are as in section 2.4: for each level of trade costs considered, we run one million simulations with production costs $a_j^n c_j$ being drawn independently for each stage n and each country j from a lognormal distribution with mean 0 and variance 1.

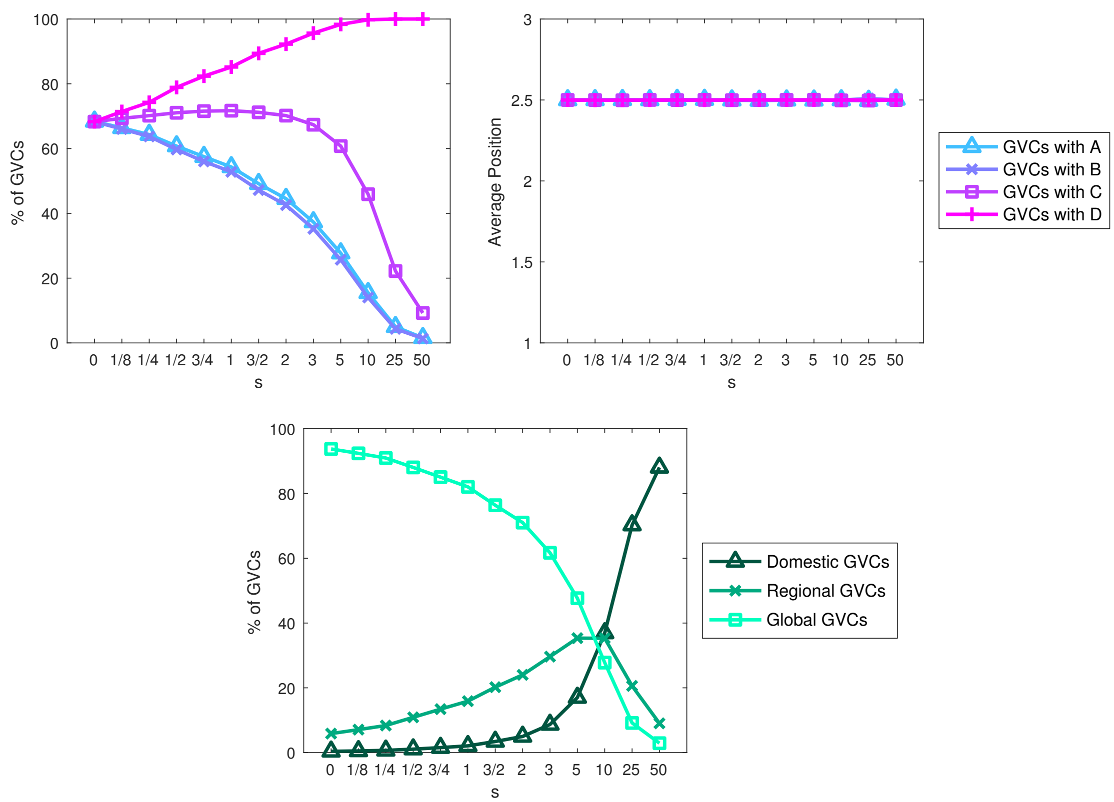

Figure B.1: Some Features of Optimal Sourcing Without Sequentiality

The results of this exercise are in Figure B.1 which is organized in a manner analogous to that in Figure 1. We continue to denote these sourcing strategies as GVCs, and also index stages from 1 to 4, although we should stress that all inputs are sourced simultaneously. For this reason, and unsurprisingly, the particular *position* or index of an input has no bearing for where it is sourced from. This is reflected in the upper right panel of Figure B.1, which shows that the average position of all countries is 2.5 for all trade costs. More interestingly, the upper left panel of Figure B.1 demonstrates that, in the absence of sequentiality, the relative prevalence of countries in GVCs serving D is strictly monotonic in the distance between these countries and D. In particular, the most remote country B is now less likely to be a source of inputs than country A, conversely to our findings in Figure 1. The lower panel of Figure B.1 unveils another interesting difference between sequential and non-sequential models of GVCs. Note, in particular, that relative to Figure 1, the relative prevalence of domestic GVCs (i.e., strategies in which all four inputs assembled in D are sourced in D itself) declines much faster with trade cost reductions. This share is close to 100% for prohibitively high trade costs, but for those in Figure 1 (i.e., $\tau_{AB} = \tau_{CD} = 1.3$, $\tau_{BC} = 1.5$, $\tau_{AD} = 1.75$, $\tau_{AC} = \tau_{BD} = 1.8$, and $s = 1$ in the Figures), 12.2% of GVCs are domestic with sequential production, but only 2.1% when inputs are all shipped simultaneously to D. When (net) trade costs are doubled (i.e., $s = 2$ in the figures), these shares are 26.6% and 5.0%, respectively.

B.2 Proof of Existence and Uniqueness

The aim of this Appendix is to study the existence and uniqueness of the general equilibrium of our model. Let us begin with some assumptions and definitions.

We shall assume throughout the following:

1. $\forall i \in J$: $\gamma_i \in (0,1]$.
2. $\sum_{n\in N} \alpha_n \beta_n = 1$.
3. There exist lower ($T_{\min}$, $\tau_{\min}$) and upper ($T_{\max}$, $\tau_{\max}$) bounds on τ_{ij} $\forall \{i,j\} \in \mathcal{J}^2$ and T_j $\forall j \in \mathcal{J}$.

Definition 2 ***(M-matrix)*** *An $n \times n$ matrix A is an M-matrix if the following* ***equivalent*** *statements hold:*

(i) A can be represented as $sI - B$, where I is $n \times n$ identity matrix, $s \in R_{++}$ is a constant and B is the matrix with positive elements and the moduli of B's eigenvalues are all $\leq s$.

(ii) A has a non-negative inverse.

Definition 3 ***(Excess demand)*** *The excess demand function $\mathbf{Z}(\mathbf{w})$ is defined as*

$$Z_i(\mathbf{w}) = \frac{1}{w_i}\left(\sum_{j\in\mathcal{J}}\sum_{n\in\mathcal{N}} \alpha_n\beta_n \times \Pr(\Lambda_i^n, j) \times \frac{1}{\gamma_j} w_j L_j\right) - \frac{1}{\gamma_i} L_i, \tag{B.1}$$

with $\Pr(\Lambda_i^n, j) = \sum_{\boldsymbol{\ell}\in\Lambda_i^n} \pi_{\boldsymbol{\ell}j}$, and where remember that $\Lambda_i^n = \left\{\boldsymbol{\ell} \in \mathcal{J}^N \mid \ell(n) = i\right\}$.

Definition 4 ***(Gross Substitutes)*** *The function* $\mathbf{F}(\mathbf{w}) : \mathbb{R}^J \to \mathbb{R}^J$ *has the gross substitutes property in* $\mathbf{w}$ *if*

$$\forall \{i,j\} \in \mathcal{J}^2, i \neq j : \qquad \frac{\partial F_i}{\partial w_j} > 0.$$

We next use these assumptions and definitions to develop proofs of existence and uniqueness that parallel those of Theorems 1-3 in Alvarez and Lucas (2007).

Theorem 1 *For any* $\mathbf{w} \in \mathbb{R}^J_{++}$ *there is a unique* $\mathbf{p}^*(\mathbf{w})$ *that solves, for all* $j \in J$

$$P_j = \kappa \left(\sum_{\boldsymbol{\ell} \in \mathcal{J}^N} \prod_{n=1}^{N} \left(\left(\left(w_{\ell(n)}\right)^{\gamma_{\ell(n)}} \left(P_{\ell(n)}\right)^{1-\gamma_{\ell(n)}} \right)^{-\theta} T_{\ell(n)} \right)^{\alpha_n \beta_n} \times \prod_{n=1}^{N-1} \left(\tau_{\ell(n)\ell(n+1)}\right)^{-\theta\beta_n} \times \left(\tau_{\ell(N)j}\right)^{-\theta} \right)^{-1/\theta}. \tag{B.2}$$

The function $\mathbf{p}^*(\mathbf{w})$ *has the following properties*

(i) *continuous in* $\mathbf{w}$.

(ii) *each component of* $\mathbf{p}^*(\mathbf{w})$ *is homogeneous of degree one in* $\mathbf{w}$*;*

(iii) *strictly increasing in* $\mathbf{w}$*;*

(iv) *strictly decreasing in* τ_{ij} *for all* $\{i,j\} \in \mathcal{J}^2$ *and strictly increasing in* T_j *for all* $j \in \mathcal{J}$.

(v) $\forall \mathbf{w} \in R^J_{++}$, *bounded between* $\underline{\mathbf{p}^*}(\mathbf{w})$ *and* $\overline{\mathbf{p}^*}(\mathbf{w})$*:*

Proof. Let us set $\tilde{p}_j = \log(P_j)$ and $\tilde{w}_j = \log(w_j)$. For each supply chain $\boldsymbol{\ell} \in \mathcal{J}^N$, let

$$d_{p,i}(\boldsymbol{\ell}) = (1-\gamma_i) \sum_{n:\boldsymbol{\ell}(n)=i} \alpha_n \beta_n < 1 \qquad d_{w,i}(\boldsymbol{\ell}) = \gamma_i \sum_{n:\boldsymbol{\ell}(n)=i} \alpha_n \beta_n < 1$$

Note that for all $i \in \mathcal{J}$, $d_{p,i} \leq 1$ and $d_{w,i} \leq 1$. Now, for all $j \in \mathcal{J}$, define $f_j(\tilde{p}, \tilde{w})$

$$f_j(\tilde{p},\tilde{w}) = \log(\kappa) - \frac{1}{\theta} \log \left(\sum_{\boldsymbol{\ell} \in \mathcal{J}^N} \prod_{n=1}^{N} \exp\left\{ -\theta \alpha_n \beta_n \left[\gamma_{\ell(n)} \tilde{w}_{\ell(n)} + \left(1-\gamma_{\ell(n)}\right) \tilde{p}_{\ell(n)} \right] \right\} T_{\ell(n)}^{\alpha_n\beta_n} \times \Upsilon_{\boldsymbol{\ell}} \right)$$

where $\Upsilon_{\boldsymbol{\ell}} = \prod_{n=1}^{N-1} \left(\tau_{\ell(n)\ell(n+1)}\right)^{-\theta\beta_n} \times \left(\tau_{\ell(N)j}\right)^{-\theta}$.

To establish uniqueness of $\mathbf{p}^*(\mathbf{w})$, we need to show that the Blackwell's sufficiency conditions for the contraction mapping theorem hold. Note that we also need to show that $f(p) = f(p, \tilde{w})$ is a bounded function for all values of $\tilde{w}$. This corresponds to property (v) of $\mathbf{p}^*(\mathbf{w})$, which will be proven below. For the time being, we proceed to prove the other parts of the theorem assuming a unique solution to the system exists.

If there indeed exists a unique solution to $\tilde{p} - f(\tilde{p}, \tilde{w}) = 0$, then homogeneity of degree one in wages (property (ii)) is simple to verify by noting that, given that $\sum_n \alpha_n \beta_n = 1$, if all wages and prices in the right-hand-side of (B.2) are multiplied by a common factor, the price level in the left-hand-side of that equation () is also scaled up or down by the same factor.

To prove differentiability and monotonicity with respect to $\mathbf{w}$, we need to determine the comparative static $\frac{\partial \mathbf{p}}{\partial \mathbf{w}}$. First, note that

$$\frac{\partial f_j(\tilde{p},\tilde{w})}{\partial p_k} = \sum_{\boldsymbol{\ell}\in\mathcal{J}^N} d_{p,k}(\boldsymbol{\ell})\pi_{\boldsymbol{\ell}j}, \tag{B.3}$$

where $\pi_{\boldsymbol{\ell}j}$ is given in (11) in the main text. Then, the Jacobian of the system $\tilde{p} - f(\tilde{p},\tilde{w})$ is given by

$$\frac{\partial\left(\tilde{p} - f\left(\tilde{p},\tilde{w}\right)\right)}{\partial \tilde{p}} = I - A^P,$$

where $\left[A^P\right]_{ij} = \frac{\partial f_i(\tilde{p},\tilde{w})}{\partial p_j}$. Note that matrix A^P is totally positive (this follows from the equation (B.3)), and therefore, by the Perron-Frobenius Theorem, we can bound above the largest eigenvalue of A^P, denoted by $\lambda_{\max}$, by the largest row sum of A^P. More precisely,

$$\begin{aligned}
\lambda_{\max} &\leq \max_k \sum_i \frac{\partial f_k(\tilde{p},\tilde{w})}{\partial \tilde{p}_i} = \max_k \sum_i \left(\sum_{\boldsymbol{\ell}\in\mathcal{J}^N} d_{p,i}(\boldsymbol{\ell})\pi_{\boldsymbol{\ell}k}\right) \\
&= \max_k \left(\sum_{\boldsymbol{\ell}\in\mathcal{J}^N}\left(\sum_{n\in\mathcal{N}}(1-\gamma_{\boldsymbol{\ell}(n)})\alpha_n\beta_n\right)\pi_{\boldsymbol{\ell}k}\right)
\end{aligned}$$

But consider the country with the lowest $\gamma_j = \underline{\gamma}$. And note that

$$\lambda_{\max} \leq (1-\underline{\gamma})\max_k \left(\sum_{\boldsymbol{\ell}\in\mathcal{J}^N}\left(\sum_{n\in\mathcal{N}}\alpha_n\beta_n\right)\pi_{\boldsymbol{\ell}j}\right) = 1-\underline{\gamma}.$$

Because $\lambda_{\max} < 1$, it follows that $I - A^P$ is an M-matrix, and, by properties of M-matrices, the inverse $\left(I - A^P\right)^{-1}$ is totally (weakly) positive. By the implicit function theorem, the Jacobian $\frac{\partial \tilde{p}}{\partial \tilde{w}}$ is given by

$$\frac{\partial \tilde{p}}{\partial \tilde{w}} = \left[I - A^P\right]^{-1} A^W,$$

where A^W is defined as

$$\left[A^W\right]_{ij} = \frac{\partial f_i(\tilde{p},\tilde{w})}{\partial \tilde{w}_j} = \sum_{\boldsymbol{\ell}\in\mathcal{J}^N} d_{w,j}(\boldsymbol{\ell})\pi_{\boldsymbol{\ell}i}.$$

Both A^W and $\left[I - A^P\right]^{-1}$ are totally positive, so $\tilde{p}$ is continuous (property (i)) and monotonically increasing (property (iii)) in $\tilde{w}$.

By analogy, we can show that property (iv) of the theorem also holds by defining $\forall\{i,j\}\in\mathcal{J}^2$, $\tilde{\tau}_{ij} = \log\tau_{ij}$ and $\forall j\in\mathcal{J}$, $\tilde{T}_j = \log T_j$, and also

$$d_{\tau,i}\left(\boldsymbol{\ell}\right) = \sum_{n:\boldsymbol{\ell}(n)=i}\beta_n, \qquad d_{T,i}\left(\boldsymbol{\ell}\right) = -\frac{1}{\theta}\sum_{n:\boldsymbol{\ell}(n)=i}\alpha_n\beta_n.$$

Applying the implicit function theorem to $f(p) = f(p, \tilde{w})$, we get:

$$\forall \{k, j\} \in \mathcal{J}^2 : \qquad \frac{\partial \mathbf{p}}{\partial \tilde{\tau}_{kj}} = \left[I - A^P\right]^{-1} A^{\tau_{kj}},$$

where $A^{\tau_{kj}}$ is $J \times 1$ vector with

$$\left[A^{\tau_{kj}}\right]_i = \frac{\partial f_i(p)}{\partial \tilde{\tau}_{kj}} = \sum_{\boldsymbol{\ell} \in \mathcal{J}} d_{\tau_{kj},i}\left(\boldsymbol{\ell}\right) \pi_{\boldsymbol{\ell} i}.$$

Also,

$$\forall j \in \mathcal{J} : \qquad \frac{\partial \mathbf{p}}{\partial \tilde{T}} = \left[I - A^P\right]^{-1} A^T,$$

where A^T is $J \times J$ matrix with elements

$$\left[A^T\right]_{ij} = \frac{\partial f_i(p)}{\partial T_j} = \sum_{\boldsymbol{\ell} \in \mathcal{J}} d_{T,i}\left(\boldsymbol{\ell}\right) \pi_{\boldsymbol{\ell} i}.$$

Note that, as was shown above, $\left[I - A^P\right]^{-1}$ is totally positive. Then, since for all $i \in \mathcal{J}$ and for all supply chains $d_{T,i}\left(\boldsymbol{\ell}\right) \geq 0$, $f(p)$ is decreasing in T. By analogy, since for all $\{k, j, i\} \in \mathcal{J}^3$, $d_{\tau_{kj},i}\left(\boldsymbol{\ell}^{\boldsymbol{i}}\right)$ is totally positive, $f(p)$ is increasing in τ_{jk}.

As for property (v) on bounds, we can define $\underline{\mathbf{p}^*}\left(\mathbf{w}\right)$ and $\overline{\mathbf{p}^*}\left(\mathbf{w}\right)$ in the following way:

$$\overline{\mathbf{p}^*}\left(\mathbf{w}\right) = \exp\left(f\left(\log\left(\mathbf{p}\right), \tilde{\mathbf{w}}, \mathbf{T}_{\min}, \boldsymbol{\tau}_{\max}\right)\right) \qquad \underline{\mathbf{p}^*}\left(\mathbf{w}\right) = \exp\left(f\left(\log\left(\mathbf{p}\right), \tilde{\mathbf{w}}, \mathbf{T}_{\max}, \boldsymbol{\tau}_{\min}\right)\right),$$

where $\mathbf{T}_{\max}\left(\tau_{\max}\right)$ and $\mathbf{T}_{\min}\left(\tau_{\min}\right)$ are $J \times 1$ $(J \times J)$ vectors (matrices) with all elements equal to the upper bound on labor productivity (trade costs)$T_{\max}$ $\left(\tau_{\max}\right)$ and the lower bound $T_{\min}$ $\left(\tau_{\min}\right)$, respectively. Then, we can note that the set $\mathbf{C}$, defined as

$$\mathbf{C} = \left\{z \in R^J : \quad \log\left(\underline{p^*}_i\left(\mathbf{w}\right)\right) \leq z_i \leq \log\left(\overline{p^*}_i\left(\mathbf{w}\right)\right)\right\}$$

is compact and, by analogy with Alvarez and Lucas (2007), $f(\cdot, \tilde{\mathbf{w}}) : \mathbf{C} \to \mathbf{C}$.

Let us finally tackle the existence and unique of the solution by verifying Blackwell's sufficient conditions for $f(\cdot, \tilde{\mathbf{w}})$ to be a contraction on $\mathbf{C}$. We have already shown that $f(\cdot, \tilde{\mathbf{w}})$ is monotone. We next show that the discounting property also holds. Set $f_i(p) = f_i(p, \tilde{w})$ for any fixed $\tilde{w}$. Then, for $a > 0$ and some $\nu \in (0, 1)$, using a Taylor approximation and the mean-value theorem, we get:

$$\forall i \in \mathcal{J} : \quad f_i(p + a) = f_i(p) + \sum_{k \in \mathcal{J}} a \cdot \frac{\partial f_i\left(p + (1 - \nu)a\right)}{\partial p_k} \leq f_i(p) + a\left(1 - \underline{\gamma}\right)$$

The last inequality follows from the fact that every row sum of A^P can be bounded above by

$$(1 - \underline{\gamma}) \max_k \left(\sum_{\boldsymbol{\ell} \in \mathcal{J}^N} \left(\sum_{n \in \mathcal{N}} \alpha_n \beta_n\right) \pi_{\boldsymbol{\ell} j}\right) = 1 - \underline{\gamma}.$$

Thus, both the monotonicity and discounting properties hold for $f(p) = f(p, \tilde{w})$. Therefore, we can apply the Contraction Mapping Theorem to $f(p, \tilde{w})$, and conclude that there is a unique solution $\mathbf{p}^*(\mathbf{w})$ to the system $\tilde{p} - f(\tilde{p}, \tilde{w})$, and that it satisfies properties (i) through (v). ■

Theorem 2 *There exists $\mathbf{w}^* \in R^{\mathcal{J}}_{++}$ which solves the system of equations*

$$Z(\mathbf{w}^*) = 0.$$

Proof. To show the existence of the equilibrium, we need to verify that the excess demand satisfies the following properties (see Propositions 17.C.1 in Mas-Colell et al., 1995, p. 585):

(i) $Z(\mathbf{w})$ is continuous on $\mathbb{R}^{\mathcal{J}}_{++}$;

(ii) $Z(\mathbf{w})$ is homogeneous of degree 0 in w

(iii) Walras Law: $\mathbf{w} \cdot Z(\mathbf{w}) = 0\ \forall \mathbf{w} \in \mathbb{R}^{J}_{++}$;

(iv) for $k = \max_j L_j > 0$, $Z_i(\mathbf{w}) > -k$ for all $i = 1, ..., n$ and $\mathbf{w} \in \mathbb{R}^{n}_{++}$;

(v) if $w^m \to w^0$, where $w^0 \neq 0$ and $w^0_i \neq 0$ for some i, then

$$\lim_{w^m \to w^0} \left(\max_j \{Z_j(w^m)\} \right) = \infty$$

Let us discuss each of these properties in turn.

(i) **Continuity** of $Z(\mathbf{w})$ on $\mathbb{R}^{\mathcal{J}}_{++}$ follows since $\Pr(\Lambda^n_i, j)$ is a continuous function of $\mathbf{w}$ – for strictly positive wages, each supply chain $\boldsymbol{\ell}$ in $\mathcal{J}^N$ is realized with non-zero probability.

(ii) **Homogeneity of degree zero** follows since $\Pr(\Lambda^n_i, j)$ is homogeneous of degree 0 in $\mathbf{w}$. To show this, note that, from the proof of Theorem 1, the equilibrium price level $\mathbf{p}^*(\mathbf{w})$ is homogeneous of degree 1 in $\mathbf{w}$. Then, both nominator and denominator (i.e., the destination specific term Θ_j) of $\Pr(\Lambda^n_i, j)$ are homogeneous of degree $-\theta$ in $\mathbf{w}$ (remember that $\sum_{n\in\mathcal{N}} \alpha_n \beta_n = 1$). It follows that $\Pr(\Lambda^n_i, j)$ is homogeneous of degree 0 in $\mathbf{w}$, and thus $Z(\mathbf{w})$ is homogeneous of degree 0 in $\mathbf{w}$ as well.

(iii) **Walras Law** follows since the system, $\mathbf{w} \cdot Z(\mathbf{w}) = 0$ is just the set of the general equilibrium conditions. Moreover, by summing up $Z(\mathbf{w})$, we get:

$$\begin{aligned}
\sum_{i\in\mathcal{J}} w_i \cdot Z_i(\mathbf{w}) &= \sum_{i\in\mathcal{J}} \gamma_i \left(\sum_{j\in\mathcal{J}} \sum_{n\in\mathcal{N}} \alpha_n \beta_n \times \Pr\left(\Lambda_i^n, j\right) \times \frac{1}{\gamma_j} w_j L_j \right) - \sum_{i\in\mathcal{J}} \frac{1}{\gamma_i} w_i L_i \\
&= \left(\sum_{n\in\mathcal{N}} \alpha_n \beta_n \times \sum_{j\in\mathcal{J}} \underbrace{\sum_{i\in\mathcal{J}} \Pr\left(\Lambda_i^n, j\right)}_{=1} \times \frac{1}{\gamma_j} w_j L_j \right) - \sum_{i\in\mathcal{J}} \frac{1}{\gamma_i} w_i L_i \\
&= \left(\underbrace{\sum_{n\in\mathcal{N}} \alpha_n \beta_n}_{=1} \times \sum_{j\in\mathcal{J}} \frac{1}{\gamma_j} w_j L_j \right) - \sum_{i\in\mathcal{J}} \frac{1}{\gamma_i} w_i L_i = 0.
\end{aligned}$$

Hence, $\mathbf{w} \cdot Z(\mathbf{w}) = 0$.

(iv) **The lower bound on $Z(\mathbf{w})$**: Since the first term in equation (B.1) is always positive, it follows that $Z(\mathbf{w})$ can be bounded from below by $Z_i(\mathbf{w}) \geq -\frac{1}{\gamma_i} L_i$.

(v) **The limit case**: Suppose $\{w^m\}$ is a sequence such that $w^m \to w^0 \neq 0$, and $w_i^0 = 0$ for some $i \in \mathcal{J}$. In this case, and given that all trade costs parameters are bounded, the probability of the supply chain that is located entirely in country i converges to 1, and the probabilities of realization of all other supply chains converge to 0 (keeping the destination fixed). Let $\Pr\left(i^N, j\right)$ denote the probability of realization of the supply chain for which all stages are located in country i with destination j. Then,

$$\lim_{w^m \to w^0} \left(\max_k \{Z_k(\mathbf{w})\} \right) = \lim_{w^m \to w^0} \left(Z_i(\mathbf{w}) \right)$$

and

$$\begin{aligned}
\lim_{w^m \to w^0} \left(\max_k \{Z_k(\mathbf{w})\} \right) &= \lim_{w^m \to w^0} \left(\frac{1}{w_i} \sum_{j\in\mathcal{J}} \left(\sum_{n\in N} \alpha_n \beta_n \right) \Pr\left(i^N, j\right) \frac{1}{\gamma_j} w_j L_j \right) - \frac{1}{\gamma_i} L_i \\
&= \lim_{w^m \to w^0} \left(\frac{1}{w_i} \sum_{j\in\mathcal{J}} \Pr\left(i^N, j\right) \frac{1}{\gamma_j} w_j L_j \right) - \frac{1}{\gamma_i} L_i \\
&= \lim_{w^m \to w^0} \left(\frac{1}{w_i} \sum_{j\neq i} \Pr\left(i^N, j\right) \frac{1}{\gamma_j} w_j L_j \right) = +\infty.
\end{aligned}$$

In sum, conditions (i) through (v) hold and thus a general equilibrium exists.

■

Theorem 3 *The solution* $\mathbf{w}^* \in R_{++}^{\mathcal{J}}$ *to the system of equations* $Z(\mathbf{w}^*) = 0$ *is unique if the following condition holds:*

$$\frac{2(1-\overline{\gamma})}{\xi^{\theta}(1-\underline{\gamma})} - (1-\underline{\gamma}) - \xi^{2\theta} \geq 0, \quad \textit{where} \quad \xi = \max_{i,j\in\mathcal{J}} \frac{\max_{k\in\mathcal{J}} \tau_{kj}/\tau_{ki}}{\min_{k\in\mathcal{J}} \tau_{kj}/\tau_{ki}} = 1,$$

and where $\overline{\gamma}$ *and* $\underline{\gamma}$ *are the largest and smallest values of* γ_j.

Proof. The proof boils down to verifying that $Z(\mathbf{w})$ has the gross substitutes property in $\mathbf{w}$ under the condition stated in the Theorem (see Proposition 17.F.3 in Mas-Colell et al., 1995, p. 613). More specifically, we need to show that

$$\forall \{i,k\} \in \mathcal{J}^2, i \neq k: \qquad \frac{\partial Z_i}{\partial w_k} > 0.$$

Totally differentiating the equation (B.1) wrt w_k, $k \neq i$, we get:

$$\frac{\partial Z_i(\mathbf{w})}{\partial w_k} = \frac{1}{w_i}\left(\sum_{n\in\mathcal{N}} \alpha_n \beta_n \times \left(\frac{1}{\gamma_k} L_k \Pr(\Lambda_i^n, k) + \sum_{j\in\mathcal{J}} \frac{1}{\gamma_j} w_j L_j \frac{d\Pr(\Lambda_i^n, j)}{dw_k}\right)\right),$$

where

$$\frac{d\Pr(\Lambda_i^n, j)}{dw_k} = \frac{\partial \Pr(\Lambda_i^n, j)}{\partial w_k} + \sum_{l\in\mathcal{J}} \frac{\partial \Pr(\Lambda_i^n, j)}{\partial P_l}\frac{\partial P_l}{\partial w_k}$$

From here, we proceed in three steps:

Step 1:.

Remember that $\Pr(\Lambda_i^n, j) = \sum_{\boldsymbol{\ell}\in\Lambda_i^n} \pi_{\boldsymbol{\ell}j}$, where $\Lambda_i^n = \left\{\boldsymbol{\ell} \in \mathcal{J}^N \mid \ell(n) = i\right\}$. Thus,

$$\frac{\partial \Pr(\Lambda_i^n, j)}{\partial w_k} = \frac{\Pr(\Lambda_i^n, j)}{w_k}\left(\frac{\partial \log(\Pr(\Lambda_i^n, j)\cdot\Theta_j)}{\partial \log(w_k)} - \frac{\partial \log(\Theta_j)}{\partial \log(w_k)}\right). \tag{B.4}$$

Since in equilibrium $\Theta_j = (p_j(\mathbf{w}))^{-\theta}$, we can use the envelope theorem to get

$$\frac{\partial \Pr(\Lambda_i^n, j)}{\partial w_k} = \frac{\theta}{w_k}\left(-\sum_{\boldsymbol{\ell}\in\Lambda_i^n} d_{w,k}(\boldsymbol{\ell})\pi_{\boldsymbol{\ell}j} + \Pr(\Lambda_i^n, j)\frac{\partial \tilde{p}_j}{\partial \tilde{w}_j}\right).$$

Step 2: Bounds on $\frac{\partial \tilde{p}}{\partial \tilde{w}}$.

Note that we can bound the row sums of A^P and $[I - A^P]^{-1}$:

$$(1-\overline{\gamma})\,\mathbf{1} \leq A^P\mathbf{1} \leq \left(1-\underline{\gamma}\right)\mathbf{1},$$

$$\left(1-\underline{\gamma}\right)^{-1}\mathbf{1} \leq \left[I - A^P\right]^{-1}\mathbf{1} \leq (1-\overline{\gamma})^{-1}\,\mathbf{1}, \tag{B.5}$$

where $\overline{\gamma}$ and $\underline{\gamma}$ are the largest and smallest values of γ_j.

For two identical supply chains with different destinations i and j, $\boldsymbol{\ell^i}$ and $\boldsymbol{\ell^j}$ it holds that

$$\forall \{i,j\} \in \mathcal{J}^2: \qquad d_{p,k}(\boldsymbol{\ell^j}) = d_{p,k}(\boldsymbol{\ell^i}), \qquad d_{w,k}(\boldsymbol{\ell^j}) = d_{w,k}(\boldsymbol{\ell^i})$$

$$\forall \{i,j\} \in \mathcal{J}^2: \qquad \pi_{\boldsymbol{\ell} j} = \frac{\left(\tau_{\boldsymbol{\ell}(N)j}/\tau_{\boldsymbol{\ell}(N)i}\right)^{-\theta} \pi_{\boldsymbol{\ell} i}}{\sum\limits_{\tilde{\boldsymbol{\ell}} \in \Lambda} \left(\tau_{\tilde{\boldsymbol{\ell}}(N)j}/\tau_{\tilde{\boldsymbol{\ell}}(N)i}\right)^{-\theta} \pi_{\tilde{\boldsymbol{\ell}} i}}$$

Let's set $\xi = \max_{i,j \in \mathcal{J}} \dfrac{\max_{k \in \mathcal{J}} \tau_{kj}/\tau_{ki}}{\min_{k \in \mathcal{J}} \tau_{kj}/\tau_{ki}} \geq 1$.

$$\forall \{i,j,k\} \in \mathcal{J}^2: \qquad \frac{1}{\xi^\theta} \leq [A^W]_{ij} \cdot \left([A^W]_{kj}\right)^{-1} \leq \xi^\theta$$

Since $\dfrac{\partial \mathbf{p}}{\partial w_j} = \left[I - A^P\right]^{-1} A^W_{[j]}$, where $A^W_{[j]}$ is the jth column of A^W, we can bound the ratio $\dfrac{\partial \tilde{p}_j}{\partial \tilde{w}_k} \Big/ \dfrac{\partial \tilde{p}_i}{\partial \tilde{w}_k}$:

$$\forall \{i,j\} \in \mathcal{J}^2: \qquad \frac{(1-\overline{\gamma})}{\xi(1-\underline{\gamma})} \leq \frac{\partial \tilde{p}_j}{\partial \tilde{w}_k} \Big/ \frac{\partial \tilde{p}_i}{\partial \tilde{w}_k} \leq \frac{\xi(1-\underline{\gamma})}{(1-\overline{\gamma})}.$$

Since all elements of A^W and A^P are less than one,

$$\left[A^W\right]_{jk} \leq \frac{\partial \tilde{p}_j}{\partial \tilde{w}_k} \leq \frac{1}{(1-\overline{\gamma})}. \tag{B.6}$$

Finally we show that for all n and i,

$$\frac{\sum\limits_{\boldsymbol{\ell} \in \Lambda_i^n} d_{w,m}(\boldsymbol{\ell}) \pi_{\boldsymbol{\ell} j}}{\left[A^W\right]_{jk}} \leq \Pr\left(\Lambda_i^n, j\right) \xi^{2\theta} \tag{B.7}$$

Let $\lambda_{\boldsymbol{\ell}}^n$ denote the set of supply chains, identical to $\boldsymbol{\ell} \in J^N$ in all stages except for n (note that there are J chains in $\lambda_{\boldsymbol{\ell}}^n$). With this definition we have

$$\left[A^W\right]_{jk} \geq \sum_{\boldsymbol{\ell} \in \Lambda_i^n} d_{w,m}(\boldsymbol{\ell}) \pi_{\boldsymbol{\ell} j} \left(\frac{\sum_{\tilde{\boldsymbol{\ell}} \in \lambda_{\boldsymbol{\ell}}^n} \pi_{\boldsymbol{\ell} j}}{\pi_{\boldsymbol{\ell} j}}\right)$$

and

$$\frac{\sum\limits_{\boldsymbol{\ell} \in \Lambda_i^n} d_{w,m}(\boldsymbol{\ell}) \pi_{\boldsymbol{\ell} j}}{\left[A^W\right]_{jk}} \leq \frac{\sum\limits_{\boldsymbol{\ell} \in \Lambda_i^n} d_{w,m}(\boldsymbol{\ell}) \pi_{\boldsymbol{\ell} j}}{\sum\limits_{\boldsymbol{\ell} \in \Lambda_i^n} d_{w,m}(\boldsymbol{\ell}) \pi_{\boldsymbol{\ell} j}} \left(\min_{\boldsymbol{\ell} \in \Lambda_i^n} \left(\frac{\sum_{\tilde{\boldsymbol{\ell}} \in \lambda_{\boldsymbol{\ell}}^n} \pi_{\boldsymbol{\ell} j}}{\pi_{\boldsymbol{\ell} j}}\right)\right)^{-1} \tag{B.8}$$

Then, let us bound $\Pr\left(\Lambda_i^n, j\right)$:

$$\Pr\left(\Lambda_i^n, j\right) \geq \left(\max_{\boldsymbol{\ell} \in \Lambda_i^n} \left(\frac{\sum_{\tilde{\boldsymbol{\ell}} \in \lambda_{\boldsymbol{\ell}}^n} \pi_{\boldsymbol{\ell} j}}{\pi_{\boldsymbol{\ell} j}}\right)\right)^{-1} \tag{B.9}$$

Therefore, combining (B.8) and (B.9) we get:

$$\frac{\sum_{\boldsymbol{\ell}\in\Lambda_i^n} d_{w,m}(\boldsymbol{\ell})\pi_{\boldsymbol{\ell}j}}{[A^W]_{jk}} \leq \left(\max_{\boldsymbol{\ell}\in\Lambda_i^n}\left(\frac{\sum_{\tilde{\boldsymbol{\ell}}\in\lambda_{\boldsymbol{\ell}}^n}\pi_{\boldsymbol{\ell}j}}{\pi_{\boldsymbol{\ell}j}}\right)\right)\cdot\left(\min_{\boldsymbol{\ell}\in\Lambda_i^n}\left(\frac{\sum_{\tilde{\boldsymbol{\ell}}\in\lambda_{\boldsymbol{\ell}}^n}\pi_{\boldsymbol{\ell}j}}{\pi_{\boldsymbol{\ell}j}}\right)\right)^{-1}\Pr\left(\Lambda_i^n, j\right)$$

Note that by definition of $\lambda_{\boldsymbol{\ell}}^n$,

$$\left(\frac{\sum_{\tilde{\boldsymbol{\ell}}\in\lambda_{\boldsymbol{\ell}}^n}\pi_{\boldsymbol{\ell}j}}{\pi_{\boldsymbol{\ell}j}}\right)\in\left[\frac{\sum_{k\in\mathcal{J}}\left((c_k)^{-\theta}T_k\right)^{\alpha_n\beta_n}}{\xi^\theta\left((c_i)^{-\theta}T_i\right)^{\alpha_n\beta_n}}, \frac{\xi^\theta\sum_{k\in\mathcal{J}}\left((c_k)^{-\theta}T_k\right)^{\alpha_n\beta_n}}{\left((c_i)^{-\theta}T_i\right)^{\alpha_n\beta_n}}\right],$$

so

$$\frac{\sum_{\boldsymbol{\ell}\in\Lambda_i^n} d_{w,m}(\boldsymbol{\ell})\pi_{\boldsymbol{\ell}j}}{[A^W]_{jk}} \leq \xi^{2\theta}\Pr\left(\Lambda_i^n, j\right).$$

Step 3: To prove the GS property, we need to show that for a fixed destination j, fixed stage n and $m\neq i$

$$\frac{\partial\Pr\left(\Lambda_i^n, j\right)}{\partial w_m}+\sum_{k\in\mathcal{J}}\frac{\partial\Pr\left(\Lambda_i^n, j\right)}{\partial\tilde{p}_k}\frac{\partial\tilde{p}_k}{\partial w_m}\geq 0.$$

By analogy with Step 1,

$$\sum_{k\in\mathcal{J}}\frac{\partial\Pr\left(\Lambda_i^n, j\right)}{\partial\tilde{p}_k}\frac{\partial\tilde{p}_k}{\partial\tilde{w}_m}=\Pr\left(\Lambda_i^n, j\right)\sum_{k\in\mathcal{J}}\frac{\partial\tilde{p}_k}{\partial\tilde{w}_m}\left(\frac{\partial\log\left(\Pr\left(\Lambda_i^n, j\right)\cdot\Theta_j\right)}{\partial\log\left(p_k\right)}-\frac{\partial\log\left(\Theta_j\right)}{\partial\log\left(p_k\right)}\right)$$

$$\sum_{k\in\mathcal{J}}\frac{\partial\pi_{\boldsymbol{\ell}j}}{\partial\tilde{p}_k}\frac{\partial\tilde{p}_k}{\partial\tilde{w}_m}=\theta\pi_{\boldsymbol{\ell}j}\left(-\left(\sum_{k\in\mathcal{J}}d_{p,k}(\boldsymbol{\ell})\frac{\partial\tilde{p}_k}{\partial\tilde{w}_m}\right)+\frac{\partial\tilde{p}_j}{\partial\tilde{w}_m}\right). \tag{B.10}$$

Combining equations (B.4) and (B.10),

$$\frac{d\Pr\left(\Lambda_i^n, j\right)}{d\tilde{w}_k}=\theta\left(2\Pr\left(\Lambda_i^n, j\right)\frac{\partial\tilde{p}_j}{\partial\tilde{w}_m}-\sum_{\boldsymbol{\ell}\in\Lambda_i^n}\pi_{\boldsymbol{\ell}j}\left(\left(\sum_{k\in\mathcal{J}}d_{p,k}(\boldsymbol{\ell})\frac{\partial\tilde{p}_k}{\partial\tilde{w}_m}\right)+d_{w,m}(\boldsymbol{\ell})\right)\right).$$

Let us use the bounds derived in Step 2: from equation (B.5),

$$\frac{d\Pr\left(\Lambda_i^n, j\right)}{dw_k}\geq\theta\left(\frac{\partial\tilde{p}_j}{\partial\tilde{w}_m}\left(\frac{2(1-\overline{\gamma})}{\xi^\theta(1-\underline{\gamma})}\Pr\left(\Lambda_i^n, j\right)-\sum_{\boldsymbol{\ell}\in\Lambda_i^n}\pi_{\boldsymbol{\ell}j}\left(\sum_{k\in\mathcal{J}}d_{p,k}(\boldsymbol{\ell})\right)\right)-\sum_{\boldsymbol{\ell}\in\Lambda_i^n}\pi_{\boldsymbol{\ell}j}d_{w,m}(\boldsymbol{\ell})\right).$$

Finally, invoking equations (B.6) and (B.6), we have:

$$\frac{d\Pr\left(\Lambda_i^n, j\right)}{dw_k}\geq\theta[A^W]_{kj}\Pr\left(\Lambda_i^n, j\right)\left(\frac{2(1-\overline{\gamma})}{\xi^\theta(1-\underline{\gamma})}-\frac{1}{\Pr\left(\Lambda_i^n, j\right)}\sum_{\boldsymbol{\ell}\in\Lambda_i^n}\pi_{\boldsymbol{\ell}j}\left(\sum_{k\in\mathcal{J}}d_{p,k}(\boldsymbol{\ell})\right)-\xi^{2\theta}\right)$$

and thus

$$\frac{d\Pr\left(\Lambda_i^n, j\right)}{dw_k}\geq\theta[A^W]_{kj}\Pr\left(\Lambda_i^n, j\right)\left(\frac{2(1-\overline{\gamma})}{\xi^\theta(1-\underline{\gamma})}-(1-\underline{\gamma})-\xi^{2\theta}\right). \tag{B.11}$$

■

Corollary 1 *Suppose the trade costs have the following form:*

$$(\tau_{ij})^{-\theta} = \rho_i \rho_j.$$

Then the equilibrium is unique if

$$\underline{\gamma}(3-\underline{\gamma}) \geq 2\overline{\gamma} \tag{B.12}$$

Proof. Note that for this specification of trade costs $\xi = 1$, and the RHS of equation (B.11) is positive whenever (B.12) holds. ■

B.3 Introducing Trade Deficits

Let D_j be country j's aggregate deficit in dollars, where $\sum_j D_j = 0$ holds since global trade is balanced. The only difference in the model's equations is that the general equilibrium equation is given by

$$\frac{1}{\gamma_i} w_i L_i = \sum_{j \in \mathcal{J}} \sum_{n \in \mathcal{N}} \alpha_n \beta_n \times \Pr\left(\Lambda_i^n, j\right) \times \left(\frac{1-\gamma_j}{\gamma_j} w_j L_j + w_j L_j - D_j\right).$$

where $w_j L_j - D_j$ is aggregate final good consumption in country j.

B.4 Further Details on Suggestive Evidence

In this Appendix we provide additional details on the suggestive empirical results in section 5. We begin by exploring the robustness of our results in Table 1. For that table, we used 2011 data for 180 countries from the Eora dataset. In Table A.1 we replicate that same table but pooling data from the 19 years for which the Eora dataset is available, namely 1995-2013, while including exporter-year and importer-year fixed effects (rather than the simpler exporter and importer fixed effects in Table A.1). As is apparent from comparing Tables 1 and A.1, the results are remarkably similar, both qualitatively as well as quantitatively. The reason for this is that the estimated elasticities are quite actually quite stable over time, as we have verified by replicating Table 1 year by year (details available upon request).

Tables A.2 and A.3 run the same specifications with the WIOT database using its 2013 and 2016 releases, respectively. The former covers the period 1995-2011 for 40 countries, while the latter covers 2000-2014 for 43 countries. As mentioned in the main text, the results with the 2013 release of the WIOD are generally qualitatively in line with those obtained with the Eora database, and indicate a significantly lower distance elasticity and lower 'home bias' in intermediate-input relative to final-good trade. Nevertheless, the results with the 2016 release of the same dataset are much weaker, and only indicate a lower 'home bias' in intermediate-input relative to final-good trade.

We finally incorporate the scatter plots mentioned in section 5, when describing the results in Table 2. More precisely, the left panel corresponds to the partial correlation underlying column (5) of Table 2 (i.e., partialling out GDP per capita). The right panel is the analogous scatter plot after dropping the Netherlands ('NLD').

Table A.1. Trade Cost Elasticities for Final Goods and Intermediate Inputs (Eora all years)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Distance	-1.118***	-0.824***	-1.153***	-0.854***	-1.224***	-0.910***	-0.797***
	(0.020)	(0.014)	(0.020)	(0.014)	(0.021)	(0.015)	(0.015)
Distance × Input					0.141***	0.113***	0.104***
					(0.005)	(0.006)	(0.006)
Continguity		2.239***		2.254***		2.350***	1.210***
		(0.111)		(0.112)		(0.120)	(0.098)
Continguity × Input						-0.191***	-0.058
						(0.035)	(0.037)
Language		0.481***		0.512***		0.601***	0.515***
		(0.026)		(0.026)		(0.029)	(0.027)
Language × Input						-0.179***	-0.168***
						(0.012)	(0.012)
Domestic							5.826***
							(0.176)
Domestic × Input							-0.656***
							(0.059)
Observations	615,600	615,600	1,231,200	1,231,200	1,231,200	1,231,200	1,231,200
R^2	0.977	0.978	0.967	0.969	0.967	0.969	0.971

Notes: Standard errors clustered at the country-pair level reported. ***, **, and * denote 1, 5 and 10 percent significance levels. All regressions include exporter-year and importer-year fixed effects. Regressions in columns (3)-(7) also include a dummy variable for inputs flows. See Appendix **??** for details on data sources.

Table A.2. Trade Cost Elasticities for Final Goods and Intermediate Inputs (2013 WIOD sample)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Distance	-1.550***	-1.244***	-1.560***	1.243***	-1.587***	-1.265***	-1.081***
	(0.056)	(0.044)	(0.057)	(0.044)	(0.059)	(0.045)	(0.042)
Distance × Input					0.055***	0.045***	0.032*
					(0.014)	(0.017)	(0.017)
Continguity		0.724***		0.750***		0.733***	0.302**
		(0.135)		(0.138)		(0.148)	(0.126)
Continguity × Input						0.033	0.164*
						(0.085)	(0.086)
Language		0.964***		1.002***		1.131***	0.258*
		(0.169)		(0.169)		(0.175)	(0.137)
Language × Input						-0.257**	-0.064
						(0.075)	(0.080)
Domestic							3.634***
							(0.275)
Domestic × Input							-0.787***
							(0.092)
Observations	27,194	27,194	54,380	54,380	54,380	54,380	54,380
R^2	0.981	0.983	0.972	0.974	0.972	0.974	0.978

Notes: Standard errors clustered at the country-pair level reported. ***, **, and * denote 1, 5 and 10 percent significance levels. All regressions include exporter-year and importer-year fixed effects. Regressions in columns (3)-(7) also include a dummy variable for inputs flows. See the Appendix for details on data sources.

Table A.3. Trade Cost Elasticities for Final Goods and Intermediate Inputs (2016 WIOD sample)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Distance	-1.638***	-1.396***	-1.648***	1.395***	-1.656***	-1.396***	-1.210***
	(0.053)	(0.044)	(0.053)	(0.044)	(0.055)	(0.045)	(0.043)
Distance × Input					0.016	0.000	-0.012
					(0.014)	(0.017)	(0.017)
Continguity		0.556***		0.573***		0.603***	0.241**
		(0.122)		(0.123)		(0.139)	(0.121)
Contingity × Input						-0.061	0.061
						(0.092)	(0.094)
Language		0.769***		0.808***		0.883***	0.131
		(0.149)		(0.150)		(0.161)	(0.127)
Language × Input						-0.150**	-0.024
						(0.072)	(0.072)
Domestic							3.453***
							(0.257)
Domestic × Input							-0.785***
							(0.083)
Observations	26,460	26,460	52,920	52,920	52,920	52,920	52,920
R^2	0.982	0.984	0.974	0.975	0.974	0.975	0.978

Notes: Standard errors clustered at the country-pair level reported. ***, **, and * denote 1, 5 and 10 percent significance levels. All regressions include exporter-year and importer-year fixed effects. Regressions in columns (3)-(7) also include a dummy variable for inputs flows. See the Appendix for details on data sources.

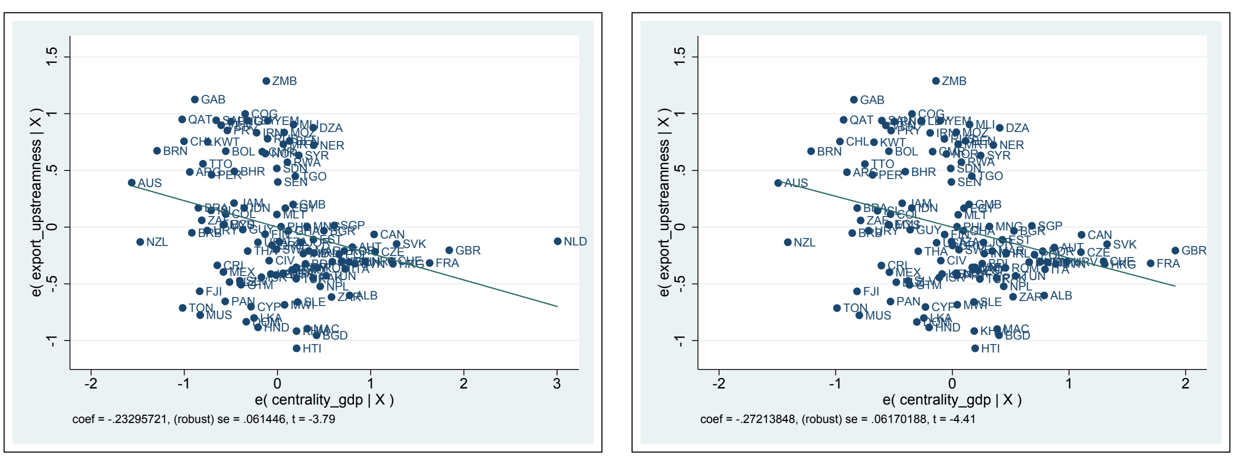

Figure B.2: Partial Correlation between Export Upstreamness and Centrality

B.5 Real Income Gains

Table B.1 reports the real income implications of the three counterfactuals studied in section 7 of the paper for the WIOD sample, and compares them with the numbers that would be obtained in an analogous Eaton and Kortum (2002) framework without sequential production (see the main text for details). Table B.2 presents the same numbers for the Eora sample of countries.

Table B.1: Real Income Gains: WIOD sample

	Autarky		50% Fall		Free Trade	
	EK	GVC	EK	GVC	EK	GVC
Australia	4.9	4.4	23.1	20.6	438.4	403.7
Austria	13.0	14.0	44.3	47.1	607.0	564.3
Belgium	21.4	22.4	62.8	64.0	609.5	618.3
Bulgaria	17.9	19.3	74.9	74.3	1715.9	1855.9
Brazil	3.2	3.4	14.8	15.6	307.3	354.1
Canada	8.0	8.2	27.4	29.0	350.4	371.3
Switzerland	10.2	10.9	41.6	39.1	507.6	424.5
China	4.1	5.3	15.8	18.5	189.4	310.2
Cyprus	13.1	13.9	63.6	60.8	1886.3	2422.6
Czech Republic	21.1	22.8	69.6	70.0	1071.3	932.2
Germany	9.4	10.2	30.9	30.9	242.8	252.2
Denmark	12.6	13.8	50.0	52.1	656.3	640.1
Spain	7.2	7.6	28.0	27.7	405.4	355.1
Estonia	21.5	24.0	87.0	87.4	2115.5	3026.4
Finland	10.1	10.8	44.8	47.1	803.3	816.7
France	7.2	7.6	25.0	25.7	282.4	281.0
Great Britain	6.6	6.8	24.1	24.5	277.7	275.3
Greece	8.3	9.2	34.0	38.1	709.0	763.1
Croatia	11.8	12.9	54.0	55.2	1315.5	1376.0
Hungary	27.8	28.9	83.1	82.1	1058.6	1078.1
Indonesia	5.6	6.1	25.4	29.8	472.1	570.6
India	4.2	4.6	17.0	21.1	326.3	404.5
Ireland	34.0	34.9	89.1	84.6	746.9	795.3
Italy	6.3	6.8	26.0	25.8	344.8	323.9
Japan	4.6	4.9	17.2	17.7	236.2	265.5
South Korea	10.6	11.3	42.2	43.0	492.8	544.1
Lithuania	20.1	22.4	75.8	72.3	1232.2	1491.2
Luxembourg	73.7	75.9	184.1	167.9	3851.8	3935.5
Latvia	14.0	15.5	64.5	67.3	2187.6	2403.5
Mexico	7.5	9.2	26.7	33.4	373.5	445.1
Malta	53.9	52.9	165.1	155.7	5179.7	7635.3
Netherlands	16.0	17.9	53.5	57.5	472.9	512.8
Norway	6.9	8.3	34.1	51.4	520.1	865.0
Poland	11.9	12.6	44.1	43.2	646.3	533.4
Portugal	9.7	10.3	39.6	40.9	779.6	706.4
Romania	10.6	11.3	45.6	47.4	938.6	866.2
Russia	5.4	5.2	24.5	27.9	364.9	468.2
Slovakia	23.4	25.5	79.5	77.3	1342.1	1117.2
Slovenia	18.2	20.3	76.5	71.7	1536.7	1393.2
Sweden	10.0	10.5	40.4	41.1	540.5	496.3
Turkey	7.6	8.2	33.9	33.8	538.4	472.6
Taiwan	15.7	17.9	59.8	70.8	670.8	913.2
USA	3.1	3.3	9.8	10.2	116.0	163.2
Rest of World	11.6	11.1	28.1	26.3	160.1	227.3

Table B.2: Real Income Gains: Eora sample

	Autarky		50% fall		Free Trade	
	EK	GVC	EK	GVC	EK	GVC
Afghanistan	4.1	4.6	17.9	26.5	3128.1	3613.8
Eastern Europe	17.0	18.1	48.4	52.7	675.4	732.9
Algeria	4.6	3.4	28.7	37.9	818.1	1484.6
Western Europe	35.0	37.7	88.9	94.0	1090.1	1377.0
Angola	3.1	1.5	30.1	13.8	1936.6	906.8
Latin America & Caribbean	8.1	8.0	26.1	27.3	794.6	675.5
Argentina	5.4	6.6	24.2	26.6	519.7	554.3
Australia	5.8	7.0	27.1	27.0	490.4	463.8
Central Europe	15.6	17.4	43.0	50.1	417.2	507.7
Central Asia	7.6	8.5	34.4	34.9	1381.0	1316.3
Middle East & North Africa	6.3	7.0	29.9	29.6	506.5	495.6
Bangladesh	3.7	4.5	21.8	24.7	1055.4	888.0
Belgium	28.9	24.9	71.5	84.7	639.3	982.5
Benin	5.2	6.7	26.5	45.2	3223.1	6889.5
South Asia	13.5	22.2	65.2	134.0	4804.1	19628.6
Bolivia	6.7	4.6	47.3	29.6	1556.2	1391.1
Sub-Saharan Africa	9.7	9.8	43.5	45.0	1255.3	1025.0
Brazil	3.1	3.7	15.6	16.0	401.5	442.8
East Asia & Pacific	7.0	8.6	38.9	36.9	891.6	723.5
Burkina Faso	7.9	7.9	23.2	38.8	2733.8	3837.6
Burundi	4.3	6.4	29.9	54.1	3447.3	8806.1
Cambodia	9.8	10.2	50.9	55.8	3133.6	2824.4
Cameroon	3.6	4.6	25.5	30.5	1691.8	2272.2
Canada	8.3	9.8	27.5	33.6	361.5	473.9
Chad	2.0	3.2	21.5	30.7	4484.0	3552.0
Chile	7.7	9.3	40.9	43.2	828.2	886.9
China	5.0	5.8	19.6	20.8	253.9	402.9
Colombia	5.0	7.1	20.9	27.1	567.2	799.9
Cuba	4.5	5.7	21.0	27.3	1113.5	1236.9
Czech Republic	19.0	21.1	62.8	71.2	1042.4	1306.0
Cote dIvoire	3.6	4.1	32.5	26.5	1724.8	1148.0
North Korea	3.0	3.0	39.2	23.1	3986.5	1527.9
DR Congo	5.5	0.9	22.8	5.7	2851.1	780.1
Scandinavia	9.1	10.2	33.6	39.4	392.2	503.3
Dominican Republic	6.0	8.1	29.2	35.9	1322.3	1452.0
Ecuador	6.0	7.4	36.2	38.0	1091.3	1130.7
Egypt	3.1	3.9	16.8	19.7	663.7	752.0
Eritrea	2.7	3.8	23.7	37.3	4009.3	6419.1
Ethiopia	659.8	1.43E+36	192.7	1111.1	1626.1	9477.2
France	8.0	8.0	27.5	28.7	290.9	358.5
Germany	12.3	12.9	37.2	41.8	269.8	404.1
Ghana	3.6	5.0	24.6	30.9	1176.5	1561.1
Greece	10.0	10.6	31.4	36.2	789.6	782.6
Guatemala	5.7	4.7	28.4	22.4	1320.7	918.3
Guinea	6.1	10.8	45.4	76.2	2329.0	9668.8
Haiti	4.1	4.7	27.5	33.0	2801.2	3400.4
Hong Kong	138.5	107.6	142.8	121.8	1860.7	1081.0
India	4.1	4.3	20.8	19.5	400.8	391.0
Indonesia	5.3	6.2	27.0	29.3	482.0	481.3
Iran	6.3	6.4	31.2	28.2	809.7	686.8

	Autarky		50% fall		Free Trade	
	EK	GVC	EK	GVC	EK	GVC
Iraq	1.9	6.7	14.9	49.0	782.7	4424.3
Israel	8.2	6.2	38.9	30.2	795.9	595.0
Italy	7.8	9.2	30.7	35.5	327.5	430.6
Japan	4.4	5.1	18.6	19.3	240.3	280.7
Kazakhstan	5.3	6.0	27.5	34.4	1000.7	1476.1
Kenya	8.8	11.0	33.7	51.1	1380.3	1556.2
Madagascar	6.6	6.2	44.1	38.2	2511.7	2019.8
Malawi	7.1	13.6	41.4	76.9	3648.3	9329.9
Malaysia	21.0	20.0	69.3	73.2	752.5	869.8
Mali	4.8	3.9	24.1	25.4	3022.8	2344.1
Mexico	6.9	10.4	24.9	36.7	369.9	560.5
Morocco	6.6	7.5	32.3	32.9	1007.6	880.0
Mozambique	3.4	4.3	15.4	24.4	1596.0	2753.7
Myanmar	0.0	0.1	2.0	1.3	2775.7	1088.0
Nepal	6.6	7.4	36.3	39.7	2380.9	2131.6
Netherlands	25.6	25.8	65.5	81.4	517.1	850.8
Niger	5.9	7.5	29.9	42.2	2547.8	4952.7
Nigeria	4.2	7.4	20.7	29.4	555.8	1078.0
Pakistan	2.0	3.1	16.6	19.5	851.3	743.2
Peru	4.8	5.6	25.5	27.3	953.4	893.9
Philippines	9.0	12.8	42.1	55.4	613.7	817.4
Poland	10.9	11.7	36.8	38.4	782.8	763.2
Portugal	11.4	12.7	41.3	42.4	876.8	857.3
South Korea	16.0	19.8	60.3	65.0	846.6	1129.7
Romania	11.5	12.9	44.1	48.6	1086.0	1143.7
Russia	3.6	3.7	18.7	20.3	392.5	497.0
Rwanda	6.5	3.1	24.6	22.9	3494.0	2666.9
Saudi Arabia	6.5	7.7	28.0	29.4	620.4	603.8
Senegal	4.5	6.1	24.7	34.4	1575.3	2272.9
Singapore	46.3	47.1	97.3	102.0	981.8	1152.0
Somalia	1.7	1.9	14.5	20.9	6917.6	16193.2
South Africa	7.2	8.2	38.3	41.3	692.8	788.0
South Sudan	0.2	0.4	4.1	5.9	3180.4	1704.8
Spain	8.7	9.3	31.8	32.3	441.9	501.1
Sri Lanka	4.3	6.8	28.6	38.1	953.4	1237.6
Sudan	0.0	0.0	0.6	0.3	1693.6	666.8
Syria	4.6	2.3	33.3	14.6	2124.7	1006.9
Taiwan	10.2	9.6	53.8	39.8	918.7	622.8
Thailand	10.7	12.7	49.9	48.1	781.7	705.2
Tunisia	11.0	10.6	45.6	41.4	1972.2	1225.7
Turkey	8.9	12.5	26.2	36.2	432.6	624.3
Uganda	5.5	6.1	19.4	31.5	2210.0	2587.1
Ukraine	14.0	15.3	50.4	51.5	1556.1	1410.4
UK	10.2	10.7	30.6	33.2	322.4	383.2
Tanzania	17.5	40.7	64.8	164.2	4897.0	16160.2
USA	3.8	4.1	11.5	12.0	135.0	213.5
Uzbekistan	3.4	4.4	24.1	24.6	1186.1	1240.1
Venezuela	3.3	2.0	23.6	19.9	695.7	793.4
Viet Nam	32.8	29.9	78.9	77.9	2251.6	1591.6
Yemen	4.3	6.1	29.8	37.0	1663.5	2162.6
Zambia	5.3	5.8	31.5	35.4	2294.5	2023.9