A Appendix

A.1 Mathematical Proofs

A.1.1 Increasing Trade-Cost Elasticity

We demonstrate that Proposition 1 holds true for arbitrary constant-returns-to-scale production technologies. With that in mind, let the sequential cost function associated with a path of production $\boldsymbol{\ell} = \{\ell(1), \ell(2), ..., \ell(N)\}$ be defined by

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

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

Define $\tilde{p}^{n-1}_{\ell(n)}(\boldsymbol{\ell}) = p^{n-1}_{\ell(n-1)}(\boldsymbol{\ell})\,\tau_{\ell(n-1)\ell(n)}$ to be 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

$$\beta^j_n = \prod_{n'=n+1}^{N} \frac{\partial \ln p^{n'}_{\ell(n')}(\boldsymbol{\ell})}{\partial \ln \tilde{p}^{n'-1}_{\ell(n')}(\boldsymbol{\ell})}, \tag{A.2}$$

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.2), it is then clear that

$$\beta_j^1 \leq \beta_j^2 \leq \cdots \leq \beta_j^N = 1,$$

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

A.1.2 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.

I. 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

$$\boldsymbol{\ell}_n^j = \arg\min_{k \in \mathcal{J}} p_k^n\left(\boldsymbol{\ell}_{n-1}^k\right)\tau_{kj}, \tag{A.3}$$

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.3) 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 $C \to D \to D$.

II. 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

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

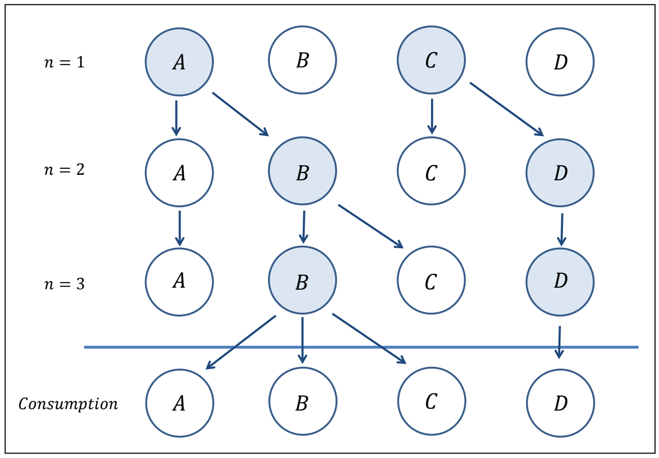

Figure A.1: Dynamic Programming − An Example with Four Countries and Three Stages

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.1.3 Proof of Proposition 3

If there is free trade or τ is constant across all country pairs (including domestically), then all countries source each variety from the same sequence of countries with $\pi_{\boldsymbol{\ell}j} = \pi_{\boldsymbol{\ell}}$ for all $j \in \mathcal{J}$. Analogously, price indices are the same in all markets so that $P_j = P$ for all $j \in \mathcal{J}$. The probability of sourcing a variety through a given sequence is thus

$$\pi_{\boldsymbol{\ell}} = \frac{\prod_{n\in\mathcal{N}} \left(T_{\ell(n)}^n w_{\ell(n)}^{-\gamma\theta}\right)^{1/N}}{\sum_{\boldsymbol{\ell}'\in\mathcal{J}^N} \prod_{n\in\mathcal{N}} \left(T_{\ell'(n)}^n w_{\ell'(n)}^{-\gamma\theta}\right)^{1/N}}.$$

We will now prove that wages are equalized across countries. Note that the total probability of any country being in a given stage n is the same regardless of the destination country and equals

$$\sum_{i\in\mathcal{J}} \Pr\left(\Lambda_i^n\right) = \sum_{i\in\mathcal{J}} \sum_{\boldsymbol{\ell}\in\Lambda_i^n} \frac{\prod_{n'\in\mathcal{N}} \left(T_{\ell(n')}^{n'} w_{\ell(n')}^{-\gamma\theta}\right)^{1/N}}{\Theta} = \sum_{i\in\mathcal{J}} \left(T_i^n w_i^{-\gamma\theta}\right)^{1/N} \times \frac{\prod_{n'\in\mathcal{N}\backslash n} \left(T_{\ell(n')}^{n'} w_{\ell(n')}^{-\gamma\theta}\right)^{1/N}}{\Theta}.$$

Now, suppose that wages are common across countries with $w_j = w$ for all $j \in \mathcal{J}$. Since the probability of any country being at a given stage n needs to equal 1, this implies that

$$\sum_{i\in\mathcal{J}} \left(T_i^n\right)^{1/N} \times \frac{\prod_{n'\in\mathcal{N}\backslash n} \left(T_{\ell(n')}^{n'}\right)^{1/N}}{w^{\gamma\theta}\Theta} = 1 \Rightarrow \frac{\prod_{n'\in\mathcal{N}\backslash n} \left(T_{\ell(n')}^{n'}\right)^{1/N}}{w^{\gamma\theta}\Theta} = \frac{1}{J\overline{T}},$$

where the second line uses our assumption that the geometric mean of T_i^n across countries is constant across stages of production. Let us now plug this into the right-hand side of the general equilibrium equation together with our guess that wages are equalized across countries

$$\begin{aligned} w_i &= \sum_{j\in\mathcal{J}}\sum_{n\in\mathcal{N}} \frac{1}{N} \times \Pr\left(\Lambda_i^n, j\right) \times w = \sum_{n\in\mathcal{N}} \frac{1}{N} \times \frac{\left(T_i^n\right)^{1/N} \times \prod_{n'\in\mathcal{N}\setminus n}\left(T_{\ell(n')}^{n'}\right)^{1/N}}{w^{\gamma\theta}\Theta} \times Jw \\ &= \sum_{n\in\mathcal{N}} \frac{1}{N} \times \frac{\left(T_i^n\right)^{1/N}}{J\overline{\overline{T}}} \times Jw = \frac{1}{N} \times \frac{N\overline{T}}{J\overline{\overline{T}}} \times Jw = w. \end{aligned}$$

Where the third line uses the previous result and where the fourth line uses our assumption that the geometric mean of T_i^n across stages of production is constant across countries. Hence, guessing that wages are equalized across countries delivers a fixed point in those wages. Since the equilibrium is unique, this is the only set of wages satisfying the general equilibrium equation.

To derive the share of goods produced in a domestic supply chain under free trade, rewrite Θ as

$$\Theta = \prod_{n\in\mathcal{N}}\sum_{i\in\mathcal{J}} \left(T_i^n\right)^{1/N} = \left(J \times \overline{T}\right)^N = \left(J \times \frac{1}{N}\sum_{n\in\mathcal{N}}\left(T_j^n\right)^{1/N}\right)^N,$$

for any $j \in \mathcal{J}$. Inserting this into the domestic expenditure share finalizes the proof

$$\pi_{\boldsymbol{j}} = \left(\frac{\text{Geometric Mean}_n\left[\left(T_j^n\right)^{1/N}\right]}{J \times \text{Arithmetic Mean}_n\left[\left(T_j^n\right)^{1/N}\right]}\right)^N.$$

A.1.4 Proof of Proposition 4

If all countries are symmetric, wages are equalized and the domestic expenditure share is

$$\pi_{\boldsymbol{j}} = \frac{1}{\sum_{\boldsymbol{\ell}'\in\mathcal{J}^N}\prod_{n\in\mathcal{N}}\left(\tau_{\ell(n)\ell(n+1)}\right)^{-\theta\beta_n}}.$$

The denominator can be rewritten as

$$\begin{aligned} &\sum_{\ell(1)\in\mathcal{J}} \cdots \sum_{\ell(N)\in\mathcal{J}}\prod_{n\in\mathcal{N}}\left(\tau_{\ell(n)\ell(n+1)}\right)^{-\theta\beta_n}, \\ &= \sum_{\ell(1)\in\mathcal{J}}\sum_{\ell(2)\in\mathcal{J}}\left(\tau_{\ell(1)\ell(2)}\right)^{-\beta_1\theta} \times \cdots \times \sum_{\ell(N)\in\mathcal{J}}\left(\tau_{\ell(N-1)\ell(N)}\right)^{-\theta\beta_{N-1}} \times \left(\tau_{\ell(N)j}\right)^{-\theta\beta_N}, \\ &= \sum_{\ell(N)\in\mathcal{J}}\left(1 + (J-1)\,\tau^{-\beta_1\theta}\right) \times \left(1 + (J-1)\,\tau^{-\beta_2\theta}\right) \times \cdots \times \left(1 + (J-1)\,\tau^{-\beta_{N-1}\theta}\right) \times \left(\tau_{\ell(N)j}\right)^{-\theta\beta_N}, \\ &= \prod_{n=1}^{N}\left(1 + (J-1)\,\tau^{-\beta_n\theta}\right). \end{aligned}$$

Substituting this in the domestic share finishes the proof.

A.1.5 Proof of Proposition 5

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

$$\Pr\left(\Lambda_i^n\right) = \sum_{\boldsymbol{\ell}\in\Lambda_i^n} \pi_{\boldsymbol{\ell}j} = \frac{\sum_{\boldsymbol{\ell}\in\Lambda_i^n} \prod_{m=1}^{N} \left(T_{\ell(m)} \left(c_{\ell(m)}\right)^{-\theta}\right)^{\alpha_m \beta_m} \left(\rho_{\ell(m)}\right)^{\beta_{m-1}+\beta_m}}{\sum_{k\in J} \sum_{\boldsymbol{\ell}\in\Lambda_k^n} \prod_{m=1}^{N} \left(T_{\ell(m)} \left(c_{\ell(m)}\right)^{-\theta}\right)^{\alpha_m \beta_m} \left(\rho_{\ell(m)}\right)^{\beta_{m-1}+\beta_m}}. \tag{A.4}$$

But note that we can decompose this as

$$\Pr\left(\Lambda_i^n\right) = \frac{\left(T_i \left(c_i\right)^{-\theta}\right)^{\alpha_n \beta_n} \left(\rho_i\right)^{\beta_{n-1}+\beta_n} \times \sum_{\boldsymbol{\ell}\in\Lambda_i^n} \prod_{m\neq n} \left(T_{\ell(m)} \left(c_{\ell(m)}\right)^{-\theta}\right)^{\alpha_m \beta_m} \left(\rho_{\ell(m)}\right)^{\beta_{m-1}+\beta_m}}{\sum_{k\in\mathcal{J}} \left(T_k \left(c_k\right)^{-\theta}\right)^{\alpha_n \beta_n} \left(\rho_k\right)^{\beta_{n-1}+\beta_n} \times \sum_{\boldsymbol{\ell}\in\Lambda_k^n} \prod_{m\neq n} \left(T_{\ell(m)} \left(c_{\ell(m)}\right)^{-\theta}\right)^{\alpha_m \beta_m} \left(\rho_{\ell(m)}\right)^{\beta_{m-1}+\beta_m}} \tag{A.5}$$

$$= \frac{\left(T_i \left(c_i\right)^{-\theta}\right)^{\alpha_n \beta_n} \left(\rho_i\right)^{\beta_{n-1}+\beta_n}}{\sum_{k\in\mathcal{J}} \left(T_k \left(c_k\right)^{-\theta}\right)^{\alpha_n \beta_n} \left(\rho_k\right)^{\beta_{n-1}+\beta_n}} \tag{A.6}$$

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

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

Now consider our definition of upstreamness

$$U\left(i\right) = \sum_{n=1}^{N} \left(N-n+1\right) \times \frac{\Pr\left(\Lambda_i^n\right)}{\sum_{n'=1}^{N} \Pr\left(\Lambda_i^{n'}\right)}. \tag{A.7}$$

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 $f_{i'}(n)$ and $f_i(n)$ satisfy the monotone likelihood ratio property in n. As is well known, this is a sufficient condition for $f_{i'}(n)$ to first-order stochastically dominate $f_i(n)$ when $\rho_{i'} > \rho_i$. But then it is immediate that $\mathbb{E}_{f_{i'}}[n] > \mathbb{E}_{f_i}[n]$, and thus the expected value in (A.7), which is simply $N + 1 - \mathbb{E}_{f_i}[n]$, will be lower for country i' than for country i when $\rho_{i'} > \rho_i$. This completes the proof of Proposition 5.

We can finally consider the case with a general path of α_n, but common technology $T_i = T$ across countries. From equation A.6, we have

$$\Pr\left(\Lambda_i^n, j\right) = \frac{\left(c_i\right)^{-\theta\alpha_n\beta_n}\left(\rho_i\right)^{\beta_{n-1}+\beta_n}}{\sum_{k\in\mathcal{J}}\left(c_k\right)^{-\theta\alpha_n\beta_n}\left(\rho_k\right)^{\beta_{n-1}+\beta_n}}.$$

We then have

$$\frac{\Pr\left(\Lambda_i^{n'}\right)/\Pr\left(\Lambda_j^{n'}\right)}{\Pr\left(\Lambda_i^{n}\right)/\Pr\left(\Lambda_j^{n}\right)} = \left(\frac{c_i}{c_j}\right)^{-\left(\alpha_{n'}\beta_{n'}-\alpha_n\beta_n\right)}\left(\frac{\rho_i}{\rho_j}\right)^{\beta_{n'-1}+\beta_{n'}-\beta_{n-1}-\beta_n}$$

Take $n' = n + 1$. Then

$$\frac{\Pr\left(\Lambda_i^{n'}\right)/\Pr\left(\Lambda_j^{n'}\right)}{\Pr\left(\Lambda_i^{n}\right)/\Pr\left(\Lambda_j^{n}\right)} = \left(\frac{c_i}{c_j}\right)^{-\theta\left(\alpha_{n+1}\beta_{n+1}-\alpha_n\beta_n\right)}\left(\frac{\rho_i}{\rho_j}\right)^{\beta_{n+1}-\beta_{n-1}}$$

Let's inspect the exponents more closely. Note $\beta_{n-1} = (1-\alpha_n)\beta_n$, so $\alpha_n\beta_n = \beta_n - \beta_{n-1}$ and

$$\frac{\Pr\left(\Lambda_i^{n'}\right)/\Pr\left(\Lambda_j^{n'}\right)}{\Pr\left(\Lambda_i^{n}\right)/\Pr\left(\Lambda_j^{n}\right)} = \left(\left(\frac{c_i}{c_j}\right)^{-\theta}\right)^{\beta_{n+1}-2\beta_n+\beta_{n-1}}\left(\frac{\rho_i}{\rho_j}\right)^{\beta_{n+1}-\beta_{n-1}}.$$

But

$$\beta_{n+1} - 2\beta_n + \beta_{n-1} < \beta_{n+1} - \beta_{n-1}$$

because $\beta_{n-1} < \beta_n$. This can be iterated starting for $n'' = n'+1$. This result implies that a sufficient condition for

$$\frac{\Pr\left(\Lambda_i^{n'}\right)/\Pr\left(\Lambda_j^{n'}\right)}{\Pr\left(\Lambda_i^{n}\right)/\Pr\left(\Lambda_j^{n}\right)} > 1$$

for $n' > n$ and $\rho_i > \rho_j$ is that $(c_i)^{-\theta}\rho_i$ is larger for more central countries. Unfortunately, the general equilibrium conditions of the model are too complex for us to be able to formally establish that this is indeed the case for all possible parameters values. But, as stated in the main text, we have run millions of simulations and have not found a single case contradicting the claim.

A.2 General Equilibrium under Decentralized Approaches

This Appendix demonstrates the isomorphism between the general equilibrium conditions derived under the lead firm (chain-productivity) formulation in the main text, and the two alternative decentralized approaches outlined in 3.2.

A.2.1 Incomplete Information Approach

We begin with the first approach with stage-specific Fréchet distributions and incomplete information. 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

$$\Pr\left(a_i^n(z)^{\alpha_n\beta_n} \geq a\right) = \exp\left\{-a^{\theta}(T_i)^{\alpha_n\beta_n}\right\}. \tag{A.8}$$

To build intuition, we begin by sketching why and how the approach works for the simple case with only two stages, input production (stage 1) and assembly (stage 2). Later, we will show how the approach naturally generalizes to the case $N > 2$.

With $N = 2$, 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 (A.8) 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 assemblers worldwide, but *not* the productivity of input producers, and thus seek to solve

$$p^F_j(z) = \min_{\ell(2)\in\mathcal{J}}\left\{\left(a^2_{\ell(2)}(z)\, c_{\ell(2)}\right)^{\alpha_2}\mathbb{E}\left[a^1_{\ell^*(1)}(z)\, c_{\ell^*(1)}\tau_{\ell^*(1)\ell(2)}\right]^{1-\alpha_2}\tau_{\ell(2)j}\right\}. \tag{A.9}$$

If retailers could observe the particular realizations of input producers, the expectation in (A.9) would be replaced by the realization of $a^1_{\ell(1)}(z)\, c_{\ell(1)}\tau_{\ell(1)\ell(2)}$ in all $\ell(1) \in \mathcal{J}$, and characterizing the optimal choice would be complicated because it would depend on the product of the distributions $a^2_{\ell(2)}(z)$ and $a^1_{\ell(1)}(z)$,which is not Fréchet under (A.8). Given our incomplete information assumption, however, the expectation in (A.9) does not depend on the particular realizations of upstream productivity draws, and this allows us to apply the well-know properties of the *univariate* Fréchet distribution in (A.8) to characterize the problem of retailers.

To see this, 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

$$\pi_{\boldsymbol{\ell}j} = \underbrace{\frac{\left(T_{\ell(1)}\right)^{1-\alpha_2}\left(c_{\ell(1)}\tau_{\ell(1)\ell(2)}\right)^{-\theta(1-\alpha_2)}}{\sum_{k\in\mathcal{J}}(T_k)^{1-\alpha_2}\left(c_k\tau_{k\ell(2)}\right)^{-\theta(1-\alpha_2)}}}_{\Pr(\ell(1)|\ell(2))} \times \underbrace{\frac{\left(T_{\ell(2)}\right)^{\alpha_2}\left(\left(c_{\ell(2)}\right)^{\alpha_2}\tau_{\ell(2)j}\right)^{-\theta}\left(\mathcal{E}_{\ell(2)}\right)^{-\theta}}{\sum_{i\in\mathcal{J}}(T_i)^{\alpha_2}\left((c_i)^{\alpha_2}(\tau_{ij})\right)^{-\theta}(\mathcal{E}_i)^{-\theta}}}_{\Pr(\ell(2))}. \tag{A.10}$$

A bit less trivially, but also exploiting well-known properties of the Fréchet distribution, it can 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 scalar $\varsigma > 0$. This allows us to reduce (A.10) to

$$\pi_{\boldsymbol{\ell}j} = \frac{\left(T_{\ell(1)}\right)^{1-\alpha_2}\left(c_{\ell(1)}\tau_{\ell(1)\ell(2)}\right)^{-\theta(1-\alpha_2)}\left(T_{\ell(2)}\right)^{\alpha_2}\left(\left(c_{\ell(2)}\right)^{\alpha_2}\tau_{\ell(2)j}\right)^{-\theta}}{\sum_{k\in\mathcal{J}}\sum_{i\in\mathcal{J}}\left(T_k\right)^{1-\alpha_2}\left(c_k\tau_{ki}\right)^{-\theta(1-\alpha_2)}\left(T_i\right)^{\alpha_2}\left(\left(c_i\right)^{\alpha_2}\left(\tau_{ij}\right)\right)^{-\theta}}. \tag{A.11}$$

It should be clear that this expression is identical to (9) – plugging in (10) – 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 (8), by $\Pr\left(p_j^F(\boldsymbol{\ell}, z) \leq p\right) = 1 - \exp\left\{-\tilde{\Theta}_j p^\theta\right\}$, where $\tilde{\Theta}_j$ is the denominator in (A.11), and is the analog of Θ_j in (10) 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.

We next generalize this result 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 above. 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 notation above, define for any $s > 0$ the expectation

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

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 (1), we can write

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

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 $n-1$, $\mathcal{E}_{\ell_z^j(n)}^{n-1}[\cdot]$, 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 $\mathcal{E}_{\ell_z^j(n)}^{n-1}[\cdot]$ 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 (1), the problem faced by a stage $n+1$ producer in j can be written as

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

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

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

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:

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

where $\varsigma = \prod_{n=1}^N \Gamma\left(1 + \frac{1-\sigma}{\beta_n\theta} \right)^{\frac{1}{1-\sigma}}$. This expression is identical to (11) up to a scalar (which is irrelevant for all equilibrium conditions and that could be 'neutralized' by an appropriate rescaling of the stage-specific Fréchet distributions).

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

$$\pi_{\boldsymbol{\ell}j} = \Pr\left(\ell_z^j(N) = \ell(N) \middle| \ell_z^{\ell(N)}(N-1) = \ell(N-1)\right) \times$$

$$\times \prod_{n=2}^{N-1} \Pr\left(\ell_z^{\ell(n+1)}(n) = \ell(n) \middle| \ell_z^{\ell(n)}(n-1) = \ell(n-1)\right) \times \Pr\left(\ell_z^{\ell(2)}(1) = \ell(1)\right)$$

$$= \Pr\left(\ell_z^j(N) = \ell(N)\right) \times \prod_{n=1}^{N} \Pr\left(\ell_z^{\ell(n+1)}(n) = \ell(n)\right)$$

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

which is identical to equation (9) in the main text obtained in the 'randomness-in-the-chain' formulation of technology.

A.2.2 Oberfield Approach

We next turn to the second decentralized approach inspired by work of Oberfield (2018). To ease the notation, let us define

$$Z_{\ell(n)}^n = \left(a_{\ell(n)}^n\right)^{-\alpha_n},$$

so that we can write equation (1) as

$$p_{\ell(n)}^n = \frac{1}{Z_{\ell(n)}^n}\left(c_{\ell(n)}\right)^{\alpha_n}\left(p_{\ell(n-1)}^{n-1}\tau_{\ell(n-1)\ell(n)}\right)^{1-\alpha_n}.$$

A key conceptual difference with this approach is that the efficiency level $Z_{\ell(n)}^n$ is now assumed to be buyer-seller specific (or *match* specific). In particular, a firm producing stage n in location $\ell(n)$ meets a certain number of potential sellers of stage $n-1$ in each location $\ell(n-1)$, with each of these potential sellers being associated with a distinct 'match' productivity of combining the good completed up to stage $n-1$ with the labor and materials at stage n. This buyer-seller specific productivity is drawn from a Pareto distribution with shape parameter θ and lower bound $\underline{Z}_{\ell(n)}^n$. Below, we will focus on the limiting case in which $\underline{Z}_{\ell(n)}^n \to 0$. Given all the available match-specific productivities and production costs, each stage-n producer (or buyer) chooses the supplier offering the lowest cost for the good produced at stage $n-1$. The number of available potential suppliers in each sourcing country $\ell(n-1)$ varies across producers, and the precise number $m_{\ell(n-1)\ell(n)}^n$ of potential suppliers based in country $\ell(n-1)$ available to a given firm producing stage n in country $\ell(n)$ is assumed to follow a Poisson distribution with arrival rate $\left(T_{\ell(n)}\right)^{\alpha_n}\left(\underline{Z}_{\ell(n)}^n\right)^{-\theta}$. For $n=1$, and for the time being, we assume that productivity in location $\ell(1)$ is fixed at $Z_{\ell(1)}^1 = \left(T_{\ell(1)}\right)^{1/\theta}$, though we will relax this assumption below.

We now derive the distribution of final good prices in country j when sourcing goods through an arbitrary supply chain $\boldsymbol{\ell}$. To build intuition, let us first study the case with two stages ($N=2$). Consider the distribution of prices that a stage 2 producer in country $\ell(2)$ can offer to consumers in country j if stage 1 output is bought from country $\ell(1)$ and the highest matched-pair productivity with suppliers in that country is $\hat{Z}_{\ell(2)}^2$. This distribution is given by

$$\begin{aligned}G_j^2\left(p\,|\ell\left(1\right),\ell\left(2\right)\right) &= \Pr\left(p\leq\frac{1}{\hat{Z}^2_{\ell(2)}}\left(c_{\ell(2)}\right)^{\alpha_2}\left(p^1_{\ell(1)}\tau_{\ell(1)\ell(2)}\right)^{1-\alpha_2}\tau_{\ell(2)j}\right),\\ &= \Pr\left(\hat{Z}^2_{\ell(2)}\leq\tilde{Z}\left(p\right)\right),\end{aligned}$$

where $\tilde{Z}\left(p\right)=\left(c_{\ell(2)}\right)^{\alpha_2}\left(c_{\ell(1)}\left(T_{\ell(1)}\right)^{-1/\theta}\tau_{\ell(1)\ell(2)}\right)^{1-\alpha_2}\tau_{\ell(2)j}/p$.

Now remember that the stage-2 producer has various potential suppliers in each country $\ell\left(1\right)$, so for the price to be higher than p, or for $\max_{\mu=1,\ldots m_{\ell(1)\ell(2)}}\left\{Z^{2,\mu}_{\ell(2)}\right\}=\hat{Z}^2_{\ell(2)}<\tilde{Z}\left(p\right)$, we need $Z^{2,m}_{\ell(2)}<\tilde{Z}\left(p\right)$ for all the draws μ associated with all the potential suppliers $m_{\ell(1)\ell(2)}$ that a specific firm has. Since both the number of suppliers and productivity of each set of suppliers is stochastic, we can obtain the overall distribution of prices invoking the formula for the Poisson probability density function and also plugging in the cumulative density function for the Pareto distribution:

$$\begin{aligned}&G_j\left(p\,|\ell\left(1\right),\ell\left(2\right)\right)\\ &=\sum_{m=0}^{\infty}\prod_{\mu=1}^{m}\Pr\left(Z^{2,\mu}_{\ell(2)}\leq\tilde{Z}\left(p\right)\right)\times\Pr\left(m_{\ell(1)\ell(2)}=m\right),\\ &=\sum_{m=0}^{\infty}\prod_{\mu=1}^{m}\left(1-\left(\frac{\underline{Z}^2_{\ell(2)}}{\tilde{Z}\left(p\right)}\right)^{\theta}\right)\times\frac{\left(\left(T_{\ell(2)}\right)^{\alpha_2}\left(\underline{Z}^2_{\ell(2)}\right)^{-\theta}\right)^m\exp\left\{-\left(T_{\ell(2)}\right)^{\alpha_2}\left(\underline{Z}^2_{\ell(2)}\right)^{-\theta}\right\}}{m!},\\ &=\exp\left\{-p^{\theta}\left(T_{\ell(1)}c^{-\theta}_{\ell(1)}\tau^{-\theta}_{\ell(1)\ell(2)}\right)^{1-\alpha_2}\left(T_{\ell(2)}c^{-\theta}_{\ell(2)}\right)^{\alpha_2}\tau^{-\theta}_{\ell(2)j}\right\}.\end{aligned}\tag{A.13}$$

This is the same expression we obtain in the "Fréchet-in-the-chain" formulation in the main text.

Now let us extend these results to the case with $N=3$, and consider the problem of producers of the final assembly stage $n=3$ in country $\ell\left(3\right)$. For such a producer, the distribution of prices it can offer to consumers in j when stage-2 inputs are bought from country $\ell\left(2\right)$ is more involved than before because it now depends on the product of the distribution of buyer-seller productivity draws $Z^3_{\ell(3)}$ and the upstream input prices $p^2_{\ell(2)}$ that each input seller itself sells at (that is, influenced by the buyer-seller productivity that the stage-2 seller has with its own input suppliers). However, note that the buyer-seller productivity draws at stage-3 are independent of the upstream productivity draws. Instead, what is crucial to take into account is the fact that stage-3 producers that get more stage-2 matches will get, on average, both a better buyer-seller productivity draw but also a better stage-2 input price. Thus, we can split the problem into two parts. We first obtain the expected price distribution conditional on a buyer-seller relationship and then we obtain the price distribution by characterizing the distribution of optimal matches.

Define the distribution of stage-3 prices in j from a given supply chain $\boldsymbol{\ell}=\left\{\ell\left(1\right),\ell\left(2\right),\ell\left(3\right)\right\}$ conditional on a specific buyer-seller relationship characterized by $Z^3_{\ell(3)}$ as

$$\begin{aligned}F_j\left(p\left|\boldsymbol{\ell},Z^3_{\ell(3)}\right.\right)&=\Pr\left(p\leq\frac{1}{Z^3_{\ell(3)}}\left(c_{\ell(3)}\right)^{\alpha_3}\left(p^2_{\ell(2)}\left(\boldsymbol{\ell}\right)\tau_{\ell(2)\ell(3)}\right)^{1-\alpha_3}\tau_{\ell(3)j}\right),\\ &=\exp\left\{-\Theta_{\ell(1)\ell(2)}\left(\frac{pZ^3_{\ell(3)}}{\tau^{1-\alpha_3}_{\ell(2)\ell(3)}c^{\alpha_3}_{\ell(3)}\tau_{\ell(3)j}}\right)^{\theta/(1-\alpha_3)}\right\},\end{aligned}$$

where $\Theta_{\ell(1)\ell(2)}=\left(T_{\ell(1)}c^{-\theta}_{\ell(1)}\tau^{-\theta}_{\ell(1)\ell(2)}\right)^{1-\alpha_2}\left(T_{\ell(2)}c^{-\theta}_{\ell(2)}\right)^{\alpha_2}$ and where we have invoked our above distribution

(A.13). As in the $N = 2$ stage case, with $N = 3$ the distribution of prices along chain $\boldsymbol{\ell}$ will be determined by the fact that each producer at the assembly stage chooses the upstream supplier that offers the best combination of buyer-seller productivity and input prices. That is

$$\begin{aligned}
&G_j\left(p\left|\ell(1), \ell(2), \ell(3)\right.\right) \\
&= \sum_{m=0}^{\infty} \prod_{\mu=1}^{m} \int_{\underline{Z}^3_{\ell(3)}}^{\infty} F_j\left(p \left| \boldsymbol{\ell}, Z^{3,\mu}_{\ell(3)}\right.\right) \Pr\left(Z^{3,\mu}_{\ell(3)} = Z\right) dZ \times \Pr\left(m_{\ell(2)\ell(3)} = m\right), \\
&= \exp\left\{ -\left(T_{\ell(3)}\right)^{\alpha_3} \left(\underline{Z}^3_{\ell(3)} - \int_{\underline{Z}^3_{\ell(3)}}^{\infty} F_j\left(p\left|\boldsymbol{\ell}, Z\right.\right) \frac{\theta}{Z^{\theta+1}} dZ \right) \right\},
\end{aligned}$$

where we used the fact that $Z^{3,\mu}_{\ell(3)}$ is a Pareto random variable with lower bound $\underline{Z}^3_{\ell(3)}$ and shape parameter θ. Now, define $\chi(p)$ such that $F_j\left(p\left|\boldsymbol{\ell}, Z^3_{\ell(3)}\right.\right) = \exp\left\{-\chi(p)\left(Z^3_{\ell(3)}\right)^{\theta/(1-\alpha_3)}\right\}$, and solve the above integral by taking the limit when $\underline{Z}^3_{\ell(3)} \to 0$ and using a change of variable $\zeta(p) = \chi(p) Z^{\theta/(1-\alpha_3)}$ to obtain

$$\begin{aligned}
\int_0^{\infty} F_j\left(p\left|\boldsymbol{\ell}, Z\right.\right) \frac{\theta}{Z^{\theta+1}} dZ &= \frac{\chi(p)^{1-\alpha_3}}{1-\alpha_3} \int_0^{\infty} \exp\left\{-\zeta(p)\right\} \zeta(p)^{\alpha_3 - 1} d\zeta(p), \\
&= \frac{\chi(p)^{1-\alpha_3}}{1-\alpha_3} \Gamma(\alpha_3),
\end{aligned}$$

where $\Gamma(\cdot)$ is the gamma function. Plugging this back in (and remember that we took the limit $\underline{Z}^3_{\ell(3)} \to 0$) we obtain that

$$\begin{aligned}
&G_j\left(p\left|\ell(1), \ell(2), \ell(3)\right.\right) \\
&= \exp\left\{ -\left(T_{\ell(3)}\right)^{\alpha_3} \frac{\chi(p)^{1-\alpha_3}}{1-\alpha_3} \Gamma(\alpha_3) \right\}, \\
&= \exp\left\{ -p^{\theta} \times \prod_{n=1}^{3} \left(c^{-\theta}_{\ell(n)} T_{\ell(n)}\right)^{\alpha_n \beta_n} \times \prod_{n=1}^{2} \left(\tau_{\ell(n)\ell(n+1)}\right)^{-\theta\beta_n} \times \left(\tau_{\ell(3)j}\right)^{-\theta} \times \frac{\Gamma(\alpha_3)}{1-\alpha_3} \right\}.
\end{aligned}$$

where notation is such that $\beta_n \equiv \prod_{m=n+1}^{N} (1-\alpha_m)$ and $\alpha_1 = 1$. This last expression is the exact same expression we obtain in the main text for $\Pr\left(p^F_j(\boldsymbol{\ell}, z) \geq p\right)$ in the $N = 3$ case except for the last scalar term involving the gamma function term. Nevertheless, this scalar term is irrelevant for the main equilibrium conditions in the model.

We have derived this result for stages $n = 3$ and $n = 2$, but it should be clear that the above derivations would work for any two stages n and $n-1$, as long as the distribution of production costs in upstream stage $n-1$ is Fréchet distributed. This has two implications. First, our assumption above that, for $n = 1$, productivity in location $\ell(1)$ is fixed at $Z^1_{\ell(1)} = \left(T_{\ell(1)}\right)^{1/\theta}$ can be relaxed and we can instead assume that $Z^1_{\ell(1)}$ is Fréchet distributed with shape parameter θ and scale parameter $T_{\ell(1)}$. Second, one can use induction to conclude from our results above that, for a general N, we obtain

$$\Pr\left(p^F_j(\boldsymbol{\ell}, z) \geq p\right) = \exp\left\{ -p^{\theta} \times \prod_{n=1}^{N} \left(c^{-\theta}_{\ell(n)} T_{\ell(n)}\right)^{\alpha_n \beta_n} \times \prod_{n=1}^{N-1} \left(\tau_{\ell(n)\ell(n+1)}\right)^{-\theta\beta_n} \times \left(\tau_{\ell(N)j}\right)^{-\theta} \times \tilde{\varsigma} \right\},$$

where $\tilde{\varsigma}$ is a positive scalar that is irrelevant for all equilibrium conditions and that can be 'neutralized' by an appropriate rescaling of the stage-specific Poisson distributions. It should be apparent that this expression coincides with equation (8) in the main text, up to this immaterial scalar $\tilde{\varsigma}$.

Finally, it remains to be show that this decentralized solution not only delivers the same distribution of final-good prices, but also the same GVC trade shares as in expression (9) in the main text. But this is implied by our previous derivations related to the decentralized approach with incomplete information. In particular, fixing a downstream stage n, the distribution of upstream costs at $n-1$ is again Fréchet distributed, so applying the law of total probability in the same manner as in (A.12) above, it is straightforward to re-derive equation (9) in the main text. And, to reiterate, the scalar $\tilde{\varsigma}$ is irrelevant for these equilibrium conditions.

A.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.

A.4 Graphical Description of Multi-Stage Production

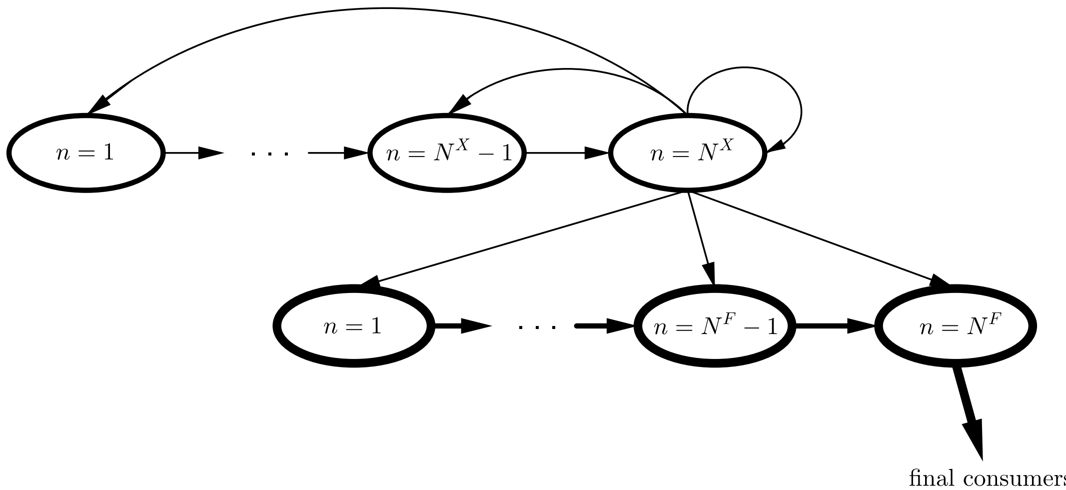

Figure A.2: Multi-stage production with separate intermediate input and final good supply chains.

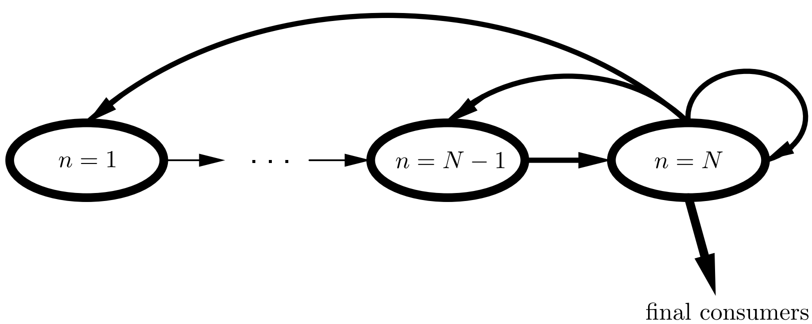

Figure A.3: Multi-stage production with common intermediate input and final good supply chains.

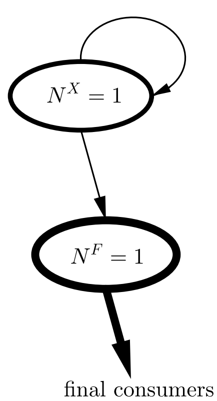

Figure A.4: Single-stage production with separate intermediate input and final good technology (Alexander 2017).

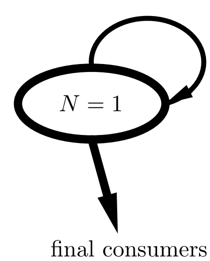

Figure A.5: Single-stage production with common intermediate input and final good technology (Eaton and Kortum 2002).

A.5 Estimation Results

	γ_j					T_j^X					T_j^F				
N	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
AUS	0.54	0.65	0.74	0.79	0.83	7.96	7.97	8.03	8.13	8.29	3.52	4.88	5.01	4.96	4.86
AUT	0.53	0.63	0.73	0.78	0.83	1.42	1.04	0.72	0.55	0.43	2.64	2.07	1.74	1.53	1.40
BEL	0.53	0.62	0.72	0.79	0.84	1.67	0.95	0.65	0.48	0.36	2.01	1.64	1.45	1.33	1.26
BGR	0.36	0.48	0.58	0.65	0.70	0.02	0.01	0.00	0.00	0.00	0.03	0.02	0.01	0.00	0.00
BRA	0.56	0.68	0.76	0.81	0.84	0.10	0.08	0.04	0.03	0.02	0.37	0.12	0.06	0.04	0.03
CAN	0.58	0.68	0.77	0.83	0.87	6.28	5.35	4.59	4.23	4.03	3.44	3.40	3.25	3.08	2.94
CHE	0.53	0.62	0.72	0.78	0.82	9.21	8.00	7.85	7.46	6.92	10.9	12.4	14.5	16.5	17.5
CHN	0.33	0.45	0.55	0.62	0.67	0.16	0.13	0.07	0.05	0.04	0.35	0.12	0.08	0.05	0.04
CZE	0.44	0.53	0.64	0.71	0.76	0.15	0.06	0.03	0.01	0.01	0.22	0.13	0.08	0.05	0.04
DEU	0.54	0.65	0.74	0.80	0.83	3.10	3.16	2.40	1.98	1.66	5.57	4.95	4.69	4.47	4.29
DNK	0.57	0.64	0.73	0.79	0.83	3.01	1.55	1.14	0.86	0.66	5.39	5.08	4.42	4.03	3.76
ESP	0.52	0.63	0.72	0.77	0.80	0.57	0.44	0.27	0.20	0.16	1.25	0.78	0.56	0.44	0.38
FIN	0.51	0.60	0.69	0.74	0.78	1.29	0.70	0.51	0.39	0.32	1.99	2.08	1.61	1.35	1.19
FRA	0.55	0.66	0.74	0.79	0.82	1.87	1.93	1.39	1.12	0.94	4.07	3.07	2.59	2.28	2.07
GBR	0.56	0.67	0.75	0.80	0.84	3.49	3.29	2.59	2.23	2.01	3.32	3.13	2.81	2.57	2.39
GRC	0.58	0.66	0.74	0.78	0.81	0.08	0.03	0.01	0.01	0.01	0.24	0.14	0.07	0.05	0.03
HRV	0.46	0.57	0.68	0.74	0.78	0.02	0.01	0.00	0.00	0.00	0.03	0.03	0.01	0.01	0.00
HUN	0.52	0.59	0.70	0.78	0.83	0.05	0.01	0.00	0.00	0.00	0.09	0.05	0.03	0.01	0.01
IDN	0.53	0.65	0.73	0.79	0.82	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00
IND	0.53	0.65	0.73	0.77	0.80	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00
IRL	0.62	0.67	0.79	0.88	0.95	3.03	0.93	0.56	0.33	0.19	2.97	2.29	2.20	2.22	2.31
ITA	0.51	0.62	0.71	0.76	0.80	0.91	0.82	0.56	0.45	0.38	1.82	1.28	1.01	0.85	0.75
JPN	0.52	0.64	0.72	0.77	0.80	1.32	1.88	1.39	1.16	1.02	6.56	3.66	3.11	2.73	2.47
KOR	0.42	0.53	0.63	0.69	0.74	0.55	0.50	0.34	0.26	0.21	1.72	0.93	0.76	0.64	0.56
LTU	0.46	0.57	0.69	0.76	0.81	0.03	0.01	0.00	0.00	0.00	0.02	0.04	0.02	0.01	0.01
LUX	0.32	0.51	0.67	0.79	0.88	0.96	2.65	3.89	4.76	5.19	0.18	1.61	2.80	4.28	5.90
MEX	0.59	0.70	0.77	0.81	0.84	0.05	0.03	0.01	0.01	0.01	0.38	0.11	0.06	0.04	0.03
NLD	0.60	0.69	0.80	0.88	0.92	5.95	3.74	3.18	2.85	2.62	2.89	2.83	2.83	2.81	2.80
NOR	0.62	0.73	0.83	0.88	0.91	31.8	32.5	40.1	43.7	46.1	14.3	21.2	22.5	23.2	23.2
POL	0.48	0.59	0.68	0.74	0.78	0.20	0.10	0.05	0.03	0.02	0.32	0.18	0.11	0.08	0.06
PRT	0.54	0.63	0.71	0.77	0.80	0.13	0.06	0.03	0.02	0.01	0.27	0.15	0.08	0.05	0.04
ROU	0.49	0.59	0.68	0.74	0.78	0.04	0.02	0.01	0.00	0.00	0.07	0.03	0.02	0.01	0.01
ROW	0.44	0.57	0.67	0.73	0.77	0.06	0.03	0.01	0.01	0.00	0.06	0.02	0.01	0.01	0.00
RUS	0.55	0.69	0.79	0.85	0.89	0.63	0.45	0.29	0.23	0.20	0.06	0.04	0.03	0.02	0.02
SVK	0.44	0.52	0.64	0.72	0.77	0.12	0.04	0.02	0.01	0.00	0.20	0.15	0.08	0.06	0.04
SVN	0.38	0.53	0.65	0.73	0.79	0.05	0.03	0.01	0.00	0.00	0.06	0.08	0.04	0.02	0.02
SWE	0.55	0.65	0.74	0.80	0.83	3.64	2.43	1.98	1.66	1.41	4.30	4.39	3.98	3.71	3.52
TUR	0.53	0.63	0.71	0.76	0.80	0.09	0.04	0.02	0.01	0.01	0.15	0.08	0.04	0.03	0.02
TWN	0.55	0.64	0.75	0.83	0.88	0.79	0.35	0.22	0.17	0.14	0.31	0.24	0.17	0.13	0.11
USA	0.57	0.69	0.77	0.82	0.85	9.15	18.7	17.0	16.6	16.6	17.8	16.5	16.7	16.4	15.9

Table A.1: Estimation Results - Asymmetric Parameterizations.

	γ_j		T_j	
N	1	2	1	2
AUS	0.52	0.88	4.79	3.48
AUT	0.55	0.87	2.14	0.57
BEL	0.54	0.83	1.92	0.45
BGR	0.61	0.95	0.10	0.00
BRA	0.57	0.99	0.15	0.01
CAN	0.55	0.92	3.43	1.65
CHE	0.52	0.81	9.44	6.35
CHN	0.33	0.57	0.18	0.03
CZE	0.48	0.73	0.23	0.02
DEU	0.55	0.87	3.90	1.65
DNK	0.59	0.92	5.04	1.83
ESP	0.54	0.88	0.77	0.14
FIN	0.54	0.87	2.01	0.59
FRA	0.56	0.93	2.58	0.93
GBR	0.55	0.91	2.95	1.24
GRC	0.63	1.00	0.16	0.01
HRV	0.70	1.00	0.22	0.01
HUN	0.61	0.91	0.14	0.01
IDN	0.55	0.93	0.01	0.00
IND	0.56	0.97	0.00	0.00
IRL	0.63	0.92	3.89	0.94
ITA	0.51	0.85	1.15	0.29
JPN	0.54	0.93	2.43	1.08
KOR	0.44	0.71	0.79	0.21
LTU	0.71	1.00	0.58	0.10
LUX	0.45	0.81	8.02	42.5
MEX	0.64	1.00	0.11	0.00
NLD	0.56	0.86	2.83	0.81
NOR	0.61	0.98	22.3	17.8
POL	0.50	0.79	0.24	0.02
PRT	0.57	0.93	0.22	0.01
ROU	0.53	0.84	0.06	0.00
ROW	0.43	0.72	0.05	0.00
RUS	0.50	0.84	0.22	0.02
SVK	0.56	0.85	0.47	0.05
SVN	0.62	1.00	0.68	0.25
SWE	0.56	0.88	4.04	1.50
TUR	0.54	0.88	0.11	0.01
TWN	0.51	0.80	0.43	0.06
USA	0.58	1.00	11.2	15.3

Table A.2: Estimation Results - Symmetric Parameterizations.

B Supplementary Online Appendix: Not for Publication

B.1 Partial Equilibrium Model: An Example

In this Appendix, we illustrate some of the salient and distinctive features of our partial model of sequential production in section 2 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 (4), 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 B.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 scale these international trade costs up or down by a shifter s (so starting from τ_{ij}, we instead use $\tilde{\tau}_{ij}(s) = 1 + s \times (\tau_{ij} - 1)$).[41] 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 B.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 B.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 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.[42] A second noteworthy aspect, apparent from the upper right panel of Figure B.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 B.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

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

[42] As we show in Online Appendix B.1 of the working paper version of our paper (Antràs and de Gortari, 2017), 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.

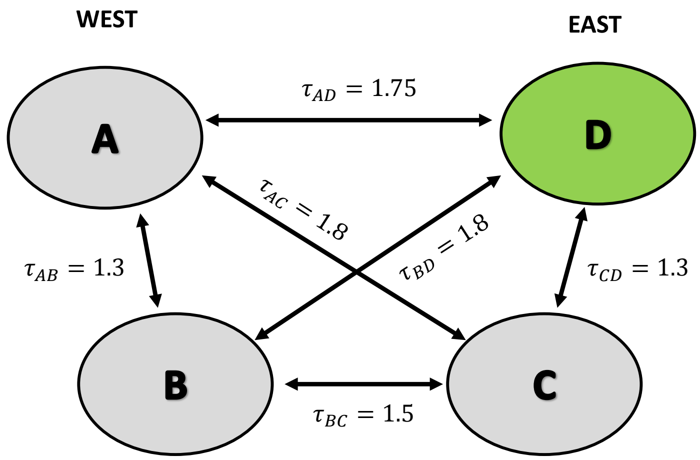

Figure B.1: An Example with Four Countries

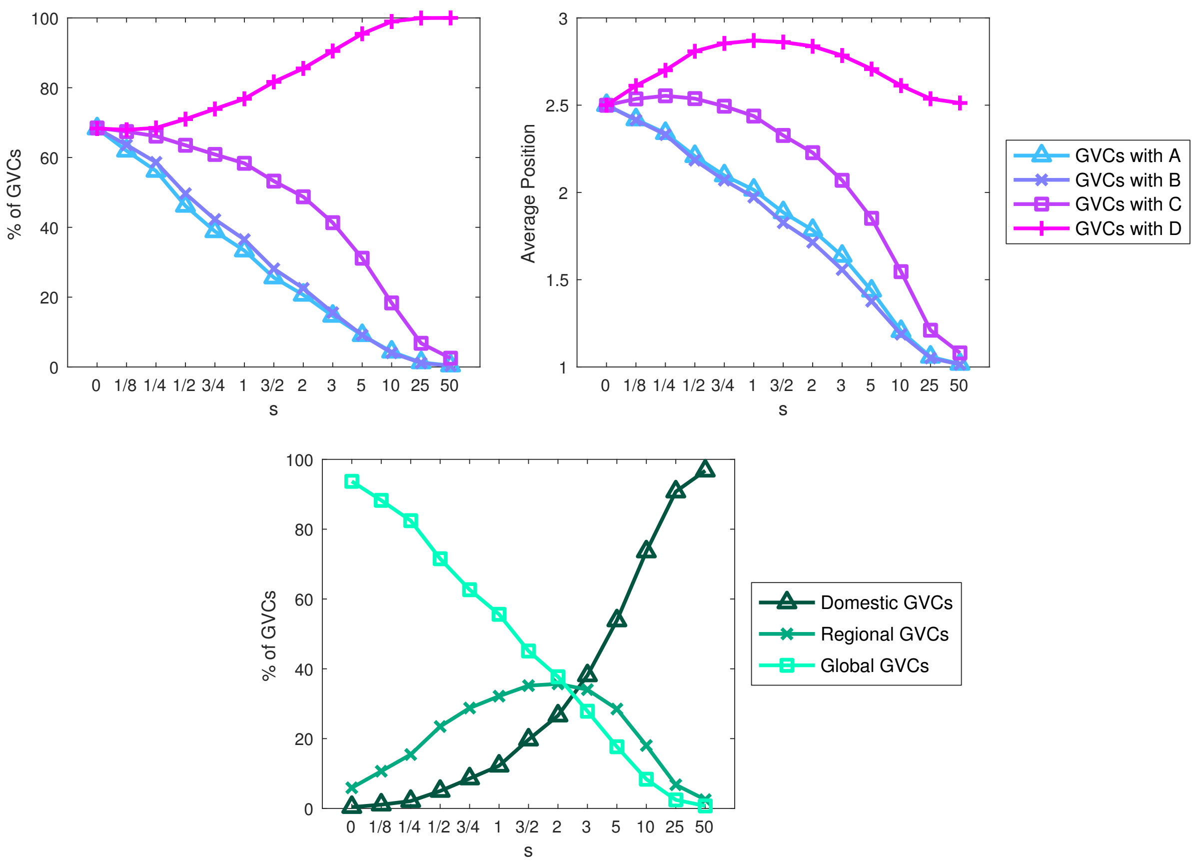

Figure B.2: Some Features of Optimal Production Paths be predicted by an analogous model without sequentiality (see the Online Appendix B.1 in Antràs and de Gortari, 2017). The reason for this is the compounding effect of trade costs, which other things equal makes it costly to offshore intermediate production stages in chains in which D has comparative advantage in the most upstream *and* downstream stages.

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. Because the core of our estimation focuses on a one-sector version of our model with variation in value added shares γ_i across countries, we develop our proofs for that more general case. 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 6 ***(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 7 ***(Excess demand)*** *The excess demand function $\mathbf{Z}(\mathbf{w})$ is defined as*

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

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

Definition 8 ***(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 (9) in the main text. Then, the Jacobian of the system $\tilde{p} - f(\tilde{p}, \tilde{w})$ is given by

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

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

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

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

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

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

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

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 $d_{T,i}(\boldsymbol{\ell}) \geq 0$, $f(p)$ 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 $\underline{\mathbf{p}^*}(\mathbf{w})$ and $\overline{\mathbf{p}^*}(\mathbf{w})$ in the following way:

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

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

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

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:*

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

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,

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

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

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

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

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

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

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

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

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,

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

Let $\lambda^n_{\boldsymbol{\ell}}$ 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 $\lambda^n_{\boldsymbol{\ell}}$). With this definition we have

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

and

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

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

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

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

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

Note that by definition of $\lambda^n_{\boldsymbol{\ell}}$,

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

so

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

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,

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

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

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

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

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

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

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

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

and thus

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

■

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

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

Then the equilibrium is unique if

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

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 Further Details on Suggestive Evidence

In this Appendix, we provide additional details on the suggestive empirical results in section 4.5. 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 B.1, we re-run the same specifications but focusing on manufacturing flows (rather than overall flows). Interestingly, the differential home bias in final-good production in Table 1 of the draft is no longer observed when focusing on manufacturing flows. Nevertheless, the response of trade flows to distance continues to be significantly higher for final goods than for intermediate inputs.

In Table B.2, we revert to the full Eora sample, but instead of focusing on 2011 data, we pool 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 1). As is apparent from comparing Tables 1 and B.2, 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 B.3 and B.4 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 **??**, 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 B.1. Trade Cost Elasticities for Final Goods and Input Flows (Manufacturing)

	Total Flows		Final-Good Flows		Input Flows	
	(1)	(2)	(3)	(4)	(5)	(6)
Distance	-1.162***	-0.793***	-1.275***	-0.893***	-1.119***	-0.755***
	(0.018)	(0.015)	(0.019)	(0.016)	(0.018)	(0.014)
Contiguity		1.243***		1.283***		1.246***
		(0.099)		(0.108)		(0.096)
Language		0.473***		0.580***		0.422***
		(0.027)		(0.031)		(0.026)
Domestic		4.299***		4.188***		4.336***
		(0.183)		(0.206)		(0.175)
Observations	32,041	32,041	32,041	32,041	32,041	32,041
R^2	0.973	0.976	0.959	0.963	0.977	0.979

Notes: Standard errors are clustered at the country-pair level. ***, **, and * denote 1, 5, and 10 percent significance levels. All regressions include exporter and importer fixed effects.

Table B.2. Trade Cost Elasticities for Final Goods and Input Flows (Eora all years)

	Total Flows		Final-Good Flows		Input Flows	
	(1)	(2)	(3)	(4)	(5)	(6)
Distance	-1.118***	-0.716***	-1.242***	-0.812***	-1.065***	-0.680***
	(0.020)	(0.013)	(0.021)	(0.015)	(0.019)	(0.013)
Contiguity		1.170***		1.189***		1.173***
		(0.088)		(0.096)		(0.086)
Language		0.401***		0.504***		0.358***
		(0.024)		(0.028)		(0.023)
Domestic		5.480***		5.800***		5.197***
		(0.159)		(0.180)		(0.151)
Observations	615,600	615,600	615,600	615,600	615,600	615,600
R^2	0.977	0.980	0.958	0.963	0.976	0.980

Notes: Standard errors are clustered at the country-pair level. ***, **, and * denote 1, 5, and 10 percent significance levels. All regressions include exporter-year and importer-year fixed effects.

Table B.3. Trade Cost Elasticities for Final Goods and Input Flows (2013 WIOD sample)

	Total Flows		Final-Good Flows		Input Flows	
	(1)	(2)	(3)	(4)	(5)	(6)
Distance	-1.550***	-1.072***	-1.579***	-1.021***	-1.541***	-1.110***
	(0.056)	(0.041)	(0.064)	(0.044)	(0.053)	(0.042)
Contiguity		0.370***		0.394***		0.375***
		(0.117)		(0.128)		(0.118)
Language		0.212		0.270*		0.181
		(0.145)		(0.139)		(0.156)
Domestic		3.141***		3.710***		2.771***
		(0.268)		(0.287)		(0.271)
Observations	27,194	27,194	27,186	27,186	27,194	27,194
R^2	0.981	0.986	0.969	0.978	0.978	0.982

Notes: Standard errors are clustered at the country-pair level. ***, **, and * denote 1, 5, and 10 percent significance levels. All regressions include exporter-year and importer-year fixed effects.

Table B.4. Trade Cost Elasticities for Final Goods and Input Flows (2016 WIOD sample)

	Total Flows		Final-Good Flows		Input Flows	
	(1)	(2)	(3)	(4)	(5)	(6)
Distance	-1.638***	-1.222***	-1.641***	-1.142***	-1.654***	-1.289***
	(0.053)	(0.042)	(0.059)	(0.042)	(0.050)	(0.045)
Contiguity		0.266**		0.292**		0.251**
		(0.108)		(0.116)		(0.111)
Language		0.129		0.197		0.091
		(0.126)		(0.125)		(0.134)
Domestic		2.950***		3.537***		2.584***
		(0.249)		(0.263)		(0.257)
Observations	26,460	26,460	26,460	26,460	26,460	26,460
R^2	0.982	0.986	0.973	0.980	0.978	0.981

Notes: Standard errors are clustered at the country-pair level. ***, **, and * denote 1, 5, and 10 percent significance levels. All regressions include exporter-year and importer-year fixed effects.

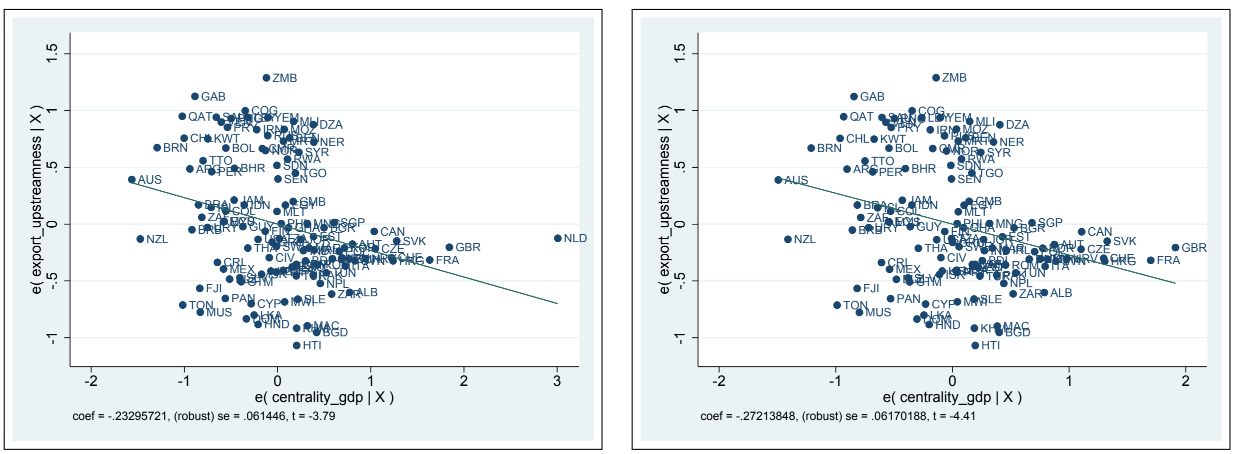

Figure B.3: Partial Correlation between Export Upstreamness and Centrality

B.4 Estimation Algorithm

We estimate our model's parameters using the mathematical programming with equilibrium constraints (MPEC) approach of Su and Judd (2012). We implement our model numerically in Matlab using the KNITRO optimizer developed by Byrd et al. (2006). This approach requires providing the optimizer with the analytical gradient and Jacobian matrix of partial first derivatives in order to deliver precise parameter estimates. We now provide the formulas for these partial derivatives – a tool that will prove useful in future implementations of our model. For the sake of simplicity, we provide the algebra for the model with separate supply chains for intermediate inputs and final goods in the single-industry case. The multi-industry formulas are analogous.

B.4.1 MPEC Estimation Algorithm

The MPEC algorithm requires defining the estimation problem as a minimization problem with two parts. First, the objective function is given by a loss function capturing the distance between a set of targeted and estimated moments. Second, the constraints include both relationships between the estimated parameters of interest and the equilibrium equations defining the model's endogenous variables. In contrast to nested fixed point algorithms, the MPEC approach does not solve for an equilibrium in each parameter iteration but yields much faster performance by only ensuring that the equilibrium constraints are satisfied in the very last iteration.

Endogenous Variables We have three sets of endogenous variables yielding a total of $2\left(N^X - 1\right) + 2\left(N^X - 1\right) + 5J + 2J^2$ variables. First, the parameters $\left\{\alpha_n^X, \beta_n^X, T_j^X, \alpha_n^F, \beta_n^F, T_j^F, \gamma_j\right\}$. Second, the general equilibrium variables $\left\{w_j, P_j^X\right\}$ with P_j^X the unit price of a CES bundle of finished intermediate input varieties in country j. Third, the estimated input-output moments $\{X_{ij}, F_{ij}\}$.

Objective Function We minimize the following loss function

$$Q = \sum_{i\in\mathcal{J}}\sum_{j\in\mathcal{J}} \omega_{ij}^X \left(X_{ij} - \hat{X}_{ij}\right)^2 + \sum_{i\in\mathcal{J}}\sum_{j\in\mathcal{J}} \omega_{ij}^F \left(F_{ij} - \hat{F}_{ij}\right)^2,$$

with ω_{ij}^X and ω_{ij}^F the weights on the targeted moments and $\hat{X}_{ij}$ and $\hat{F}_{ij}$ the targeted data.

Linear Constraints We impose general equilibrium and three normalizations as linear constraints

$$\begin{aligned} w_j L_j &= \sum_{i\in\mathcal{J}} (X_{ji} + F_{ji}) - \sum_{i\in\mathcal{J}} X_{ij},\ \forall j, \\ \sum_{j\in\mathcal{J}} w_j L_j &= 100, \\ \sum_{j\in\mathcal{J}} T_j^X &= 100, \\ \sum_{j\in\mathcal{J}} T_j^F &= 100. \end{aligned}$$

Nonlinear Constraints We impose three sets of nonlinear constraints: i) the relation between the sequential production parameters, ii) the price index definition, iii) the input-output variables definition:

$$\begin{aligned} \beta_n^X &= \left(1 - \alpha_{n+1}^X\right) \beta_{n+1}^X,\ \forall n < N^X, \\ \beta_n^F &= \left(1 - \alpha_{n+1}^F\right) \beta_{n+1}^F,\ \forall n < N^F, \\ P_j^X &= \Gamma\left(1 + \frac{1-\sigma}{\theta}\right)^{\frac{1}{1-\sigma}} \left[\Theta_j^X\right]^{-\frac{1}{\theta}},\ \forall j, \\ F_{ij} &= \sum_{\boldsymbol{\ell}\in\Lambda_i^N} \pi_{\boldsymbol{\ell}j}^F \times (w_j L_j - D_j),\ \forall i, j, \\ X_{ij} &= \sum_{n=1}^{N^F-1} \sum_{\boldsymbol{\ell}\in\Lambda_i^n\cap\Lambda_j^{n+1}} \beta_n^F \times \sum_{k\in\mathcal{J}} \pi_{\boldsymbol{\ell}k}^F \times (w_k L_k - D_k), \\ &+ \sum_{n=1}^{N^X-1} \sum_{\boldsymbol{\ell}\in\Lambda_i^n\cap\Lambda_j^{n+1}} \beta_n^X \times \sum_{k\in\mathcal{J}} \pi_{\boldsymbol{\ell}k}^X \times \frac{1-\gamma_k}{\gamma_k} w_k L_k, \\ &+ \sum_{\boldsymbol{\ell}\in\Lambda_i^N} \pi_{\boldsymbol{\ell}j}^X \times \frac{1-\gamma_j}{\gamma_j} w_j L_j,\ \forall i, j, \end{aligned}$$

where $\Gamma(\cdot)$ is the gamma function, D_j is country j's (exogenous) gross deficit. Note that Θ_j^X, $\pi_{\boldsymbol{\ell}j}^X$, and $\pi_{\boldsymbol{\ell}j}^X$ are auxiliary variables depending only on problem's endogenous variables.

B.4.2 Auxiliary Partial Derivatives

Remember from (18) the marginal costs $c_j = (w_j)^{\gamma_j} \left(P_j^X\right)^{1-\gamma_j}$. The marginal cost's partial derivatives equal

$$\frac{\partial c_j}{\partial w_j} = \gamma_j \frac{c_j}{w_j}, \qquad \frac{\partial c_j}{\partial P_j^X} = \left(1-\gamma_j\right) \frac{c_j}{P_j^X}, \qquad \frac{\partial c_j}{\partial \gamma_j} = c_j \ln\left(\frac{w_j}{P_j^X}\right).$$

Throughout the rest of this subsection we drop the X and F superscripts for notational simplicity but all derivations apply for both cases. Define the auxiliary variable

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

so that $\Theta_j = \sum_{\boldsymbol{\ell} \in \mathcal{J}^N} \Theta_j(\boldsymbol{\ell})$. The partial derivatives of Θ_j equal

$$\frac{\partial \Theta_j}{\partial \alpha_m} = \beta_m \sum_{\boldsymbol{\ell} \in \mathcal{J}^N} \ln\left(T_{\ell(m)} c_{\ell(m)}^{-\theta}\right) \Theta_j(\boldsymbol{\ell}),$$

$$\frac{\partial \Theta_j}{\partial \beta_m} = \sum_{\boldsymbol{\ell} \in \mathcal{J}^N} \left(\alpha_m \ln\left(T_{\ell(m)} c_{\ell(m)}^{-\theta}\right) + \ln\left(\tau_{\ell(m)\ell(m+1)}^{-\theta}\right)\right) \Theta_j(\boldsymbol{\ell}),$$

$$\frac{\partial \Theta_j}{\partial T_i} = \frac{1}{T_i} \sum_{n=1}^{N} \alpha_n \beta_n \times \sum_{\boldsymbol{\ell} \in \Lambda_i^n} \Theta_j(\boldsymbol{\ell}),$$

$$\frac{\partial \Theta_j}{\partial \gamma_i} = -\frac{\theta}{c_i} \frac{\partial c_i}{\partial \gamma_i} \sum_{n=1}^{N} \alpha_n \beta_n \times \sum_{\boldsymbol{\ell} \in \Lambda_i^n} \Theta_j(\boldsymbol{\ell}),$$

$$\frac{\partial \Theta_j}{\partial w_i} = -\frac{\theta}{c_i} \frac{\partial c_i}{\partial w_i} \sum_{n=1}^{N} \alpha_n \beta_n \times \sum_{\boldsymbol{\ell} \in \Lambda_i^n} \Theta_j(\boldsymbol{\ell}),$$

$$\frac{\partial \Theta_j}{\partial P_i^X} = -\frac{\theta}{c_i} \frac{\partial c_i}{\partial P_i^X} \sum_{n=1}^{N} \alpha_n \beta_n \times \sum_{\boldsymbol{\ell} \in \Lambda_i^n} \Theta_j(\boldsymbol{\ell}).$$

Similarly, the partial derivatives of $\pi_{\boldsymbol{\ell}j}$ can be written in terms of these partial derivatives as

$$\begin{aligned}
\frac{\partial \pi_{\boldsymbol{\ell}j}}{\partial \alpha_m} &= \beta_m \ln\left(T_{\ell(n)} c_{\ell(m)}^{-\theta}\right) \times \pi_{\boldsymbol{\ell}j} - \frac{\pi_{\boldsymbol{\ell}j}}{\Theta_j}\frac{\partial \Theta_j}{\partial \alpha_m}, \\
\frac{\partial \pi_{\boldsymbol{\ell}j}}{\partial \beta_m} &= \left(\alpha_m \ln\left(T_{\ell(m)} c_{\ell(m)}^{-\theta}\right) + \ln\left(\tau_{\ell(m)\ell(m+1)}^{-\theta}\right)\right) \times \pi_{\boldsymbol{\ell}j} - \frac{\pi_{\boldsymbol{\ell}j}}{\Theta_j}\frac{\partial \Theta_j}{\partial \beta_m}, \\
\frac{\partial \pi_{\boldsymbol{\ell}j}}{\partial T_i} &= \frac{1}{T_i}\sum_{n=1}^{N} \alpha_n \beta_n \times 1\left[\ell(n) = i\right] \pi_{\boldsymbol{\ell}j} - \frac{\pi_{\boldsymbol{\ell}j}}{\Theta_j}\frac{\partial \Theta_j}{\partial T_i}, \\
\frac{\partial \pi_{\boldsymbol{\ell}j}}{\partial \gamma_i} &= -\frac{\theta}{c_i}\frac{\partial c_i}{\partial \gamma_i}\sum_{n=1}^{N} \alpha_n \beta_n \times 1\left[\ell(n) = i\right] \pi_{\boldsymbol{\ell}j} - \frac{\pi_{\boldsymbol{\ell}j}}{\Theta_j}\frac{\partial \Theta_j}{\partial \gamma_i}, \\
\frac{\partial \pi_{\boldsymbol{\ell}j}}{\partial w_i} &= -\frac{\theta}{c_i}\frac{\partial c_i}{\partial w_i}\sum_{n=1}^{N} \alpha_n \beta_n \times 1\left[\ell(n) = i\right] \pi_{\boldsymbol{\ell}j} - \frac{\pi_{\boldsymbol{\ell}j}}{\Theta_j}\frac{\partial \Theta_j}{\partial w_i}, \\
\frac{\partial \pi_{\boldsymbol{\ell}j}}{\partial P_i^X} &= -\frac{\theta}{c_i}\frac{\partial c_i}{\partial P_i^X}\sum_{n=1}^{N} \alpha_n \beta_n \times 1\left[\ell(n) = i\right] \pi_{\boldsymbol{\ell}j} - \frac{\pi_{\boldsymbol{\ell}j}}{\Theta_j}\frac{\partial \Theta_j}{\partial P_i^X}.
\end{aligned}$$

Writing these formulas in the computer can be challenging and prone to typos, so it is useful to note that the following relation should hold for any parameter x: $\sum_{\boldsymbol{\ell}\in\mathcal{J}^N} \partial \pi_{\boldsymbol{\ell}j}/\partial x = 0$.

B.4.3 Gradient and Jacobian Matrix

Objective Function Gradient The derivatives for the objective function equal

$$\frac{\partial Q}{\partial X_{ij}} = 2\omega_{ij}^X \left(X_{ij} - \hat{X}_{ij}\right), \qquad \frac{\partial Q}{\partial F_{ij}} = 2\omega_{ij}^F \left(F_{ij} - \hat{F}_{ij}\right).$$

Jacobian Matrix of Nonlinear Constraints First-Order Partial Derivatives Call $C(x)$ the nonlinear constraint defining variable x such that $C(x^*) = 0$ at the solution. For example, let $C\left(\beta_n^X\right) = \beta_n^X - \left(1 - \alpha_{n+1}^X\right)\beta_{n+1}^X$. The Jacobian elements for the nonlinear constraints equal the following.

- Sequential production parameters:

$$\frac{\partial C(\beta_n)}{\partial \beta_n} = 1, \qquad \frac{\partial C(\beta_n)}{\partial \beta_{n+1}} = -\left(1 - \alpha_{n+1}\right), \qquad \frac{\partial C(\beta_n)}{\partial \alpha_{n+1}} = \beta_{n+1}.$$

  These hold for both intermediate inputs and final goods.

- Price indices:

$$\frac{\partial C\left(P_j^X\right)}{\partial x} = 1\left[x = P_j^X\right] + \frac{1}{\theta}\frac{P_j^X}{\Theta_j^X}\frac{\partial \Theta_j^X}{\partial x},$$

  where x is any of the endogenous variables.

- Final good flows:

$$\frac{\partial C(F_{ij})}{\partial x} = 1\left[x = F_{ij}\right] - \sum_{\boldsymbol{\ell}\in\Lambda_i^N} \frac{\partial \pi_{\boldsymbol{\ell}j}^F}{\partial x} \times \left(w_j L_j - D_j\right) - 1\left[p = w_j\right] \sum_{\boldsymbol{\ell}\in\Lambda_i^N} \pi_{\boldsymbol{\ell}j}^F \times L_j,$$

  where x is any of the endogenous variables.

- Intermediate input flows:

$$\begin{aligned}
\frac{\partial C\left(X_{ij}\right)}{\partial X_{ij}} =&1, \\
\frac{\partial C\left(X_{ij}\right)}{\partial x} =&-\sum_{n=1}^{N^X-1}\sum_{\boldsymbol{\ell}\in\Lambda_i^n\cap\Lambda_j^{n+1}}\sum_{k\in\mathcal{J}}\left(\beta_n^X\frac{\partial\pi_{\boldsymbol{\ell}k}^X}{\partial x}+1\left[x=\beta_n^X\right]\pi_{\boldsymbol{\ell}k}^X\right)\times\frac{1-\gamma_k}{\gamma_k}w_kL_k-\sum_{\boldsymbol{\ell}\in\Lambda_i^N}\frac{\partial\pi_{\boldsymbol{\ell}j}^X}{\partial x}\times\frac{1-\gamma_j}{\gamma_j}w_jL_j \\
&-1\left[x=P_h^X\right]\sum_{n=1}^{N^F-1}\sum_{\boldsymbol{\ell}\in\Lambda_i^n\cap\Lambda_j^{n+1}}\beta_n^F\times\sum_{k\in\mathcal{J}}\frac{\partial\pi_{\boldsymbol{\ell}k}^F}{\partial x}\times\left(w_kL_k-D_k\right),\ x=\alpha_m^X,\beta_m^X,T_h^X,P_h^X, \\
\frac{\partial C\left(X_{ij}\right)}{\partial x} =&-\sum_{n=1}^{N^F-1}\sum_{\boldsymbol{\ell}\in\Lambda_i^n\cap\Lambda_j^{n+1}}\sum_{k\in\mathcal{J}}\left(\beta_n^F\frac{\partial\pi_{\boldsymbol{\ell}k}^F}{\partial x}+1\left[x=\beta_n^F\right]\pi_{\boldsymbol{\ell}k}^F\right)\times\left(w_kL_k-D_k\right),\ x=\alpha_m^F,\beta_m^X,T_h^F, \\
\frac{\partial C\left(X_{ij}\right)}{\partial w_h} =&-\sum_{n=1}^{N^X-1}\sum_{\boldsymbol{\ell}\in\Lambda_i^n\cap\Lambda_j^{n+1}}\beta_n^X\times\left(\sum_{k\in\mathcal{J}}\frac{\partial\pi_{\boldsymbol{\ell}k}^X}{\partial w_h}\times\frac{1-\gamma_k}{\gamma_k}w_kL_k+\pi_{\boldsymbol{\ell}h}^X\times\frac{1-\gamma_h}{\gamma_h}L_h\right), \\
&-\sum_{\boldsymbol{\ell}\in\Lambda_i^N}\left(\frac{\partial\pi_{\boldsymbol{\ell}j}^X}{\partial w_h}\times w_j+1\left[h=j\right]\pi_{\boldsymbol{\ell}j}^X\right)\frac{1-\gamma_j}{\gamma_j}L_j, \\
&-\sum_{n=1}^{N^F-1}\sum_{\boldsymbol{\ell}\in\Lambda_i^n\cap\Lambda_j^{n+1}}\beta_n^F\times\left(\sum_{k\in\mathcal{J}}\frac{\partial\pi_{\boldsymbol{\ell}k}^F}{\partial w_h}\times\left(w_kL_k-D_k\right)+\pi_{\boldsymbol{\ell}h}^F\times L_h\right), \\
\frac{\partial C\left(X_{ij}\right)}{\partial \gamma_h} =&-\sum_{n=1}^{N^X-1}\sum_{\boldsymbol{\ell}\in\Lambda_i^n\cap\Lambda_j^{n+1}}\beta_n^X\times\left(\sum_{k\in\mathcal{J}}\frac{\partial\pi_{\boldsymbol{\ell}k}^X}{\partial \gamma_h}\times\frac{1-\gamma_k}{\gamma_k}w_kL_k-\pi_{\boldsymbol{\ell}h}^X\times\left(\frac{1}{\gamma_h}\right)^2w_hL_h\right), \\
&-\sum_{\boldsymbol{\ell}\in\Lambda_i^N}\left(\frac{\partial\pi_{\boldsymbol{\ell}j}^X}{\partial \gamma_h}\times\frac{1-\gamma_j}{\gamma_j}-1\left[h=j\right]\pi_{\boldsymbol{\ell}j}^X\left(\frac{1}{\gamma_j}\right)^2\right)w_jL_j, \\
&-\sum_{n=1}^{N^F-1}\sum_{\boldsymbol{\ell}\in\Lambda_i^n\cap\Lambda_j^{n+1}}\beta_n^F\times\sum_{k\in\mathcal{J}}\frac{\partial\pi_{\boldsymbol{\ell}k}^F}{\partial \gamma_h}\times\left(w_kL_k-D_k\right).
\end{aligned}$$

B.5 Multi-Industry Loss Function

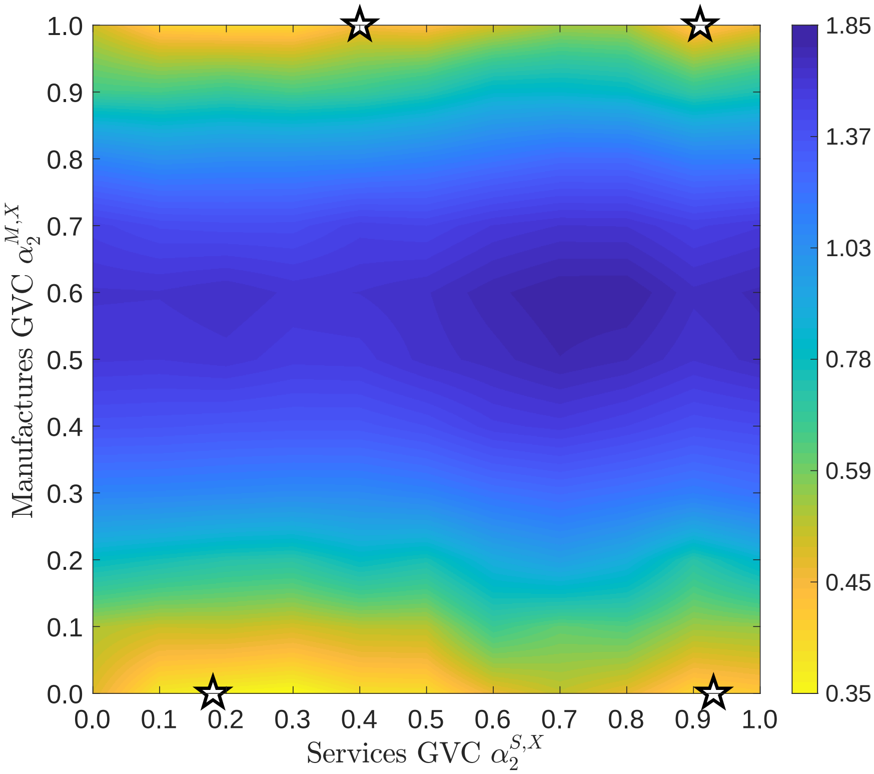

Figure B.4: Multi-Industry Loss Function and Local Minima for $N = 2$.