Production Clustering and Offshoring

By Vladimir Tyazhelnikov*

I introduce a model of international production that allows the production chain to be of any length or number of sourcing countries, while imposing only weak assumptions on the structure of production and trade costs. The production process does not have to be perfectly sequential, and the final goods can be made from any number of independent subchains. I show that in this model, allocation decisions at different stages of production are interdependent, which generates a new channel of proximity-concentration trade-off. The presence of trade costs makes firms cluster their production in certain countries while trade liberalization allows firms to fragment their production more and exploit productivity differences between countries more efficiently. Clustering patterns depend on the characteristics of the production structure, with stronger clustering associated with longer and less connected structures. Clustering intensity in upstream stages of production is generally higher and less affected by exogenous changes in production structure.
JEL: C61, D24, F12, F23, L14
Keywords: Global value chains, clustering, offshoring, fragmentation, sequential production

Production processes are becoming more global — trade in parts and components accounts for approximately two thirds of overall world trade (Johnson and Noguera (2012)). One of the prevalent modes of global production, global value chains (GVCs), has recently attracted considerable attention. The key property of GVCs is that they represent sequential production — intermediate goods are used in the production of other intermediate goods. This paper studies the implications of this sequentiality.

In this paper, I introduce a model of offshoring in which a firm solves such a problem for an arbitrarily long production chain. The main novelty of the model is a new channel of proximity-concentration trade-off: due to the presence of trade costs, a firm has to organize its production in clusters, even though some parts

* Vladimir Tyazhelnikov: The University of Sydney, Office 561, Social Sciences Building, The University of Sydney, NSW, 2006, Australia, vladimir.tyazhelnikov@sydney.edu.au. I am grateful to Robert Feenstra, Katheryn Russ and Deborah Swenson for their suggestions and support. I also thank Murali Agastya, Kirill Borusyak, Lorenzo Caliendo, Jaerim Choi, Oleg Itskhoki, Edwin Jiang, Réka Juhász, Andreas Moxnes, James Morley, Dmitry Mukhin, Stephen Redding, John Romalis, Ina Simonovska, Alan Taylor, and Mingzhi Xu for helpful comments, as well as seminar participants at UC Davis, Warwick Economics PhD Conference 2016, Sonoma State University, European Trade Study Group 2016, Iowa State University, University of Oslo, The University of Sydney, Western Social Sciences Association 2017, Australian National University, Conference on Global Production (National University of Singapore), Australasian Trade Workshop, Society for Advanced Economic Theory, Moscow International Economics Conference, and University of Melbourne.

within these clusters may be cheaper to produce in another location. Decreased trade costs allow a firm to fragment its production more and exploit productivity differences between countries more efficiently.[1]

These predictions are consistent with a sequential production model (the "snake") proposed by Baldwin and Venables (2013). They show that assumptions about production structure matter greatly; the snake generates qualitatively different predictions on trade flows compared to a simultaneous production model (the so-called "spider") . Snakes and spiders are two limiting cases of perfectly sequential and nonsequential technologies, but these two cases can be too restrictive. I am the first to allow firms to have a more general class of technology, which I call "trees", that nests both spider and snake technologies. A tree can have more than one sequential production subchain. These subchains represent the production technology of complex intermediate parts assembled to become a final good.

One example of tree technology is the Boeing 787 Dreamliner. This is an extremely complex product that consists of 2.3 million parts[2], with Boeing having production facilities in Europe, Canada, Latin America, Asia, and Australia as well as hundreds of suppliers in each of those regions.[3] For instance, 35% of the Dreamliner's parts are manufactured in Japan by Boeing and its 150 suppliers.[4] Moreover, many of Boeing's suppliers also produce and source their inputs internationally, so Boeing's largest supplier, the American company Spirit Aerosystems, has production facilities in Scotland, Ireland, France, Malaysia, and Morocco and hundreds of suppliers worldwide.[5]

I show that in the presence of such sequentiality, the problem of choosing optimal production locations becomes nontrivial. Most of the previous literature on GVCs introduced particular structures to obtain the closed form solution to this problem; these restrictions include a number of stages and particular structures of production and trade costs. These restrictions made a closed form solution possible by shutting down the interdependence of stages of production, hence ignoring the clustering effect that I am studying in this paper. The allocation problem is complex because it cannot be broken down by a sequence of independent decisions at every stage: location decisions at each stage affect costs associated with production decisions on the up- and downstream stages.

I solve the allocation problem by employing an optimal control algorithm based on the Bellman optimality principle. Conditional on the production location in the downstream stage, a decision at the current stage becomes a simple problem. It turns out that this solution to the central planner's problem is the same as

[1]Production clusters of complementary intermediate parts are easily observable in the managerial practices of multinationals. Frigant and Lung (2002) describe the strategy of modular production and its prevalence in the car manufacturing market. Other examples of production clustering are electronics (Baldwin and Clark (2000)) and the bicycle industry (Galvin and Morkel (2001)).

[2]http://787updates.newairplane.com/787-Suppliers/World-Class-Supplier-Quality

[3]http://www.boeing.com/global

[4]http://www.boeing.com/resources/boeingdotcom/company/key_orgs/boeing-international/pdf/japanbackgrounder.pdf

[5]https://www.spiritaero.com/company/overview

the allocation in decentralized equilibrium, where independent firms are looking for the lowest-price supplier, the reason being that the price of the intermediate goods in the upstream stage and in its downstream destination is a sufficient statistic to make an optimal decision at every stage.

Next, I describe firm behavior in snake and tree models. Trade liberalization leads to lower prices and higher welfare in most offshoring models, which is not surprising — even if a firm does not change its production decisions, inputs that a firm is sourcing from abroad are now cheaper. I show that in the presence of sequentiality, changes in optimal allocation necessarily lead to an increase in production efficiency, regardless of whether the new optimal path corresponds to higher or lower spending on transportation. I attribute these gains in production efficiency to higher fragmentation.

My model generates other interesting outcomes. For example, depending on the destination country of the final good, a firm can choose different optimal paths (for example, Ford Europe and Ford USA). The idea is that there is a trade-off between production efficiency and proximity to consumer markets. If shipping costs of the final good are high enough, a firm would prefer to perform the later stages of production in the destination countries.

Another implication of production clustering is that trade liberalization between two countries has an ambiguous effect on offshoring to third countries. The mechanism here is similar to "Bridge Production" from Ramondo and Rodrıguez-Clare (2013); if a firm chooses to reallocate production of one part to as a result of bilateral trade liberalization, it may also choose to produce adjacent parts in nearby low-cost countries, thus forming a new regional cluster.

Finally, my model is consistent with reshoring, a well-documented phenomenon (Sirkin et al. (2011), Wu and Zhang (2011)), in which a previously offshored part is once again produced domestically. A firm facing high trade costs may choose to offshore a large cluster of production activities. With lower costs, a firm can afford to have smaller clusters and may choose to reshore some parts previously produced in the large cluster. It follows that one should be careful interpreting the impact of reshoring on domestic employment and wages; reshoring in my model is driven by a fall in trade costs and, hence, is accompanied by the offshoring of parts previously produced domestically. This mechanism is similar to that in Harms, Lorz and Urban (2012), which my model generalizes.

To establish the link between production structure and clustering patterns, I rely on the analysis of "complete trees", a particular kind of tree that has a fractal structure and is characterized by three parameters: length, the number of upstream parts at every stage (order), and the number of countries where production is possible. Complete trees allow me to establish the link between these three parameters and the intensity of clustering without loss of generality — I show that any tree can be represented as a complete tree.

In a series of Monte Carlo simulations, I show that in a snake model, most up- and downstream stages are associated with the same level of clustering, which is higher than for all internal stages. The reason this occurs is that parts at the most up- and downstream stages have only one down- and upstream neighbor correspondingly, while all other stages have two neighbors, which makes it less likely that a part at this internal stage will be colocated with either of its neighbors.

With an increase in the order of a tree, a symmetry between the most up- and downstream stages is violated; the most upstream stage still clusters more than any other stages for all values of trade costs, while the most downstream stage clusters more than internal stages for low values of trade costs and less when trade costs are high. Two mechanisms drive this nonmonotone ranking: on the one hand, the smaller number of neighbors makes clustering easier, with the most downstream part having more neighbors than most upstream parts but fewer than the internal ones; on the other hand, parts with upstream neighbors that easily cluster with them can prioritize clustering with their downstream neighbors.

The latter mechanism explains why in longer trees, the intensity of clustering for upstream stages barely changes, while downstream stages are affected considerably. The allocation problem at the most upstream stage is always similar, but while moving downstream, clustering propagates imperfectly, thus leading to less clustering in the most downstream stage in longer trees.

Finally, I show that while reshoring rarely occurs in a snake and never occurs in a spider, it can be a common phenomenon in other tree-like structures, affecting, in some cases, up to 40% of all produced parts. Reshoring is more likely to occur in longer trees, with higher chances of occuring upstream for shorter trees and downstream in the case of longer trees. Higher tree order is also associated with higher chances of reshoring, with the most downstream stages affected the most. The probability of reshoring is also increasing in the number of countries as with the higher number of cost draws for each part, there is a higher chance that there will be cost draws that would lead to reshoring.

An important limitation of my approach is that I assume constant returns to scale, so the unit costs of production in a given country are independent of the total value added in this country. In case this assumption is relaxed, costs associated with production decisions depend not only on the production location of up- and downstream parts, as in the case of constant returns to scale, but also on the production locations of every single part instead, which would make the application of the recursive algorithm infeasible.[6]

This paper contributes to the rich literature on global value chains – which includes Antràs and Chor (2013), Costinot, Vogel and Wang (2013), Fally and Hillberry (2015), Johnson and Moxnes (2019), Razhev (2015), and Harms, Lorz and Urban (2012), by introducing a quantifiable model of offshoring with weak assumptions on production and transportation costs, the number of stages of production, and the number of countries. Importantly, this paper moves beyond

[6] Jiang and Tyazhelnikov (2020) relax this assumption and allow the location of each part to directly affect the costs associated with the production of every other part. They overcome the computational constraints with an algorithm introduced by Jia (2008) and further refined by Arkolakis and Eckert (2017).

the dichotomy of snakes and spider models by introducing and analyzing more general tree technology.

My model is similar in spirit to Tintelnot (2017) and Antras, Fort and Tintelnot (2017). Both papers introduce quantifiable models with complex problems of firms that do not have a closed form solution but rather can be solved with the help of a numerical algorithm. In Tintelnot (2017) the combination of fixed costs of opening a new plant and variable shipping costs drives the proximity-concentration trade-off. Antras, Fort and Tintelnot (2017) assume that intermediate inputs are imperfect substitutes or complements, and a firm incurs fixed costs of sourcing inputs from every country. There are no fixed costs in my model; I focus on firms' allocation problems that arise in those cases when technology exhibits at least some degree of sequentiality.

The clustering mechanism and the recursive algorithm described in this paper are also close in spirit to the contemporaneous work by Antràs and De Gortari (2020). There are, however, two main distinctions. First, Antràs and De Gortari (2020) focus on the "centrality-downstreamness nexus," a mechanism that leads to more centrally located countries specializing in more downstream production stages under the assumption of iceberg trade costs. In contrast, I focus on the clustering effect — lower trade costs are associated with higher fragmentation, whether downstream or upstream in the value chain; I am also agnostic about assumptions on trade costs. Second, to obtain closed form solutions, Antràs and De Gortari (2020) make assumptions on cost distribution; in their baseline characterization, they assume that the total cost of the whole production chain, rather than each individual component, follows a Fréchet distribution, which degenerates the clustering mechanism.[7]

The rest of the paper is organized as follows: Section I describes a firm's cost minimization problem and the algorithms to solve it. Section II describes properties of firms' behavior. Section III. Section IV concludes.

I. Firm's Problem

In this section, I define the firm's cost minimization problem and introduce a recursive algorithm which solves it. For clarity of exposition, I start with the snake model and specific trade costs and introduce a simple version of the recursive algorithm. Then I extend the problem for the more general case of tree technology and present the extended version of the algorithm. Further, I discuss the properties of this algorithm, introduce an extension for iceberg trade costs, and present an alternative algorithm based on linear programming.

A. Snake

[7] A model by Johnson and Moxnes (2019) can also generate clustering effects, the authors, however, limit their analysis to the case of 2 stages, thus, also closing this channel.

SETUP. — There is one firm that produces a final good from N intermediate parts. Each part can be produced in one of K countries. I follow Eaton and Kortum (2002) and assume that parts are perfect substitutes, which is not critical for the functioning of the model, but allows for a more intuitive representation of clustering.[8] Production costs are country- and stage-specific and are equal to $a_{i,j}$ where i is a stage of production and j is the country of production. The parts have to be produced in a given order determined by the reversed numeration of the stages, i.e., stage 1 corresponds to the most downstream part and stage N to the most upstream. Every time the firm chooses to produce an upstream part k in a different country, it pays trade costs $\tau T(j,k)$, where T is the matrix of trade costs with $T(j,j)$=0, $T(j,k) > 0 | j \neq k$, and τ is the trade cost scale parameter.[9] I assume that the most downstream stage 1 is distribution and can only happen in the final good's destination market. Firm's production decisions and costs may then depend on the destination market, so I will present this problem separately for each destination market d.

The firm then minimizes its per-unit costs, specific for the destination market d, which I call marginal costs MC_d

$$MC_d = \min_{\{c_i\}_{i=2}^{N},} \sum_{i=1}^{N} \left(\sum_{k=1}^{K} \mathbf{1}(c_i = k)\, a_{i,k} + \tau T(c_i, c_{i+1}) \right), \tag{1}$$

where $c_i = k$ if the firm chooses to produce part i in country k, and $c_1 = d$.[10]

Note that the firm cannot break this problem into $N-1$ independent subproblems for every stage, as the decision at the current stage affects all subsequent decisions. The main feature of this model is a trade-off between clustering production to minimize trade costs and exploiting productivity differences between countries.

Figure 1 illustrates this trade-off with a simple example, with two countries and five stages. Black dots represent a firm's choice to produce a part in a given country, the dotted line represents the border between two countries, and arrows represent the firm's optimal path. In this example, the firm's optimal choice is a function of trade costs: for high values of τ, the firm chooses to produce the whole good in the East. With lower trade costs, it may make sense to offshore

[8] Note that parts produced in different countries are not necessarily identical; the cost of production in each country can be interpreted as quality-adjusted. I find this deviation from the Armington assumption quite realistic for industries such as automanufacturing and electronics. When a car producer sources input, for example, wheels, it cares about the wheels' price and quality but does not benefit from a larger variety of wheels sourced from different locations.

[9] In this paper, by trade costs I mean the costs of offshoring. In a narrow sense, they are the costs associated with shipping intermediate goods between countries.

[10] In Section I.B I extend expression 1 for the more general case of complete trees. The nature of the problem remains the same and the main difference is that in case of trees, each part is not uniquely characterized by the stage it is produced at. Instead, each part is described by two indexes; stage of production, as in the snake model, and an index within the stage.

a large cluster to the West. Finally, when τ is even lower, the firm breaks its large production cluster in the West and reshores part 3 to the East. I provide a numerical example consistent with this in Proposition 5.

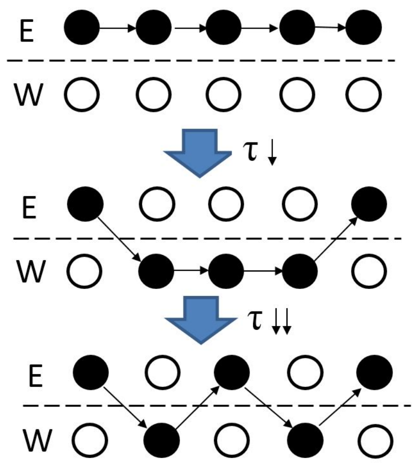

FIGURE 1. NONMONOTONIC IMPACT OF TRADE LIBERALIZATION

ALGORITHM. — The firm's choice set in (1) includes K^N paths, thus solving problem (1) by calculating marginal costs for each path is often infeasible, even for moderate values of N and K. As this problem does not have a closed form solution, the literature has constrained itself to particular modeling assumptions that allow for closed form solutions. I instead propose an optimal control algorithm that can efficiently solve this problem in its general formulation.

The cost minimization problem (1) can be rewritten in the form of the Bellman equation:

$$V_i\left(c_{i-1}\right)=\min_{c_i\in K}\left\{\sum_{k=1}^{K}\mathbf{1}\left(c_i=k\right)a_{i,k}+\tau T\left(c_i,c_{i-1}\right)+V_{i+1}\left(c_i\right)\right\}, \tag{2}$$

where $K=\{1,...,K\}$ and $V_{N+1}\left(c_N\right)=0$. The value of the value function $V_i\left(c_{i-1}\right)$ represents the optimal costs of a subchain that includes part i and all the upstream parts. This value depends on the production location of the downstream part c_{i-1}, corresponding optimal decision of part c_i and values corresponding to the optimal decisions in upstream stages $V_{i+1}\left(c_i\right)$.

This problem can be solved recursively. The algorithm determines for each of K countries at stage $N-1$ the optimal production location in stage N. The total costs in both stages of these optimal choices are written in the value function in stage $N-1$. The process is repeated at stage $N-2$ for each of K countries to choose where to locate production at stage $N-1$, given the value of the value function in every country at stage $N-1$. The same is done in every stage, until there are K value functions for $c_1 = d$ ($d \in K$), that represent optimal trajectories of final goods produced for market d. This path minimizes the costs according to the Bellman principle of optimality. Given that there are just K values of a value function to be stored at every stage of production, and at every stage the algorithm chooses the minimum of these K values for each value of state variable c_i, the number of operations an algorithm has to perform is $K \times N$.

B. *Tree Production Structure*

In the previous section, I focused on the snake model, assuming that all parts have to be produced in some exogenously given natural order. This assumption is one of the few popular approaches in the offshoring literature to model a production technology of a complex good. Other popular choices are spiders (nonsequential models), where the order of production does not matter (Antras and Helpman (2004), Feenstra and Hanson (1996) and others), and the models where it is costly to produce each part in a different country from the location of the headquarters, as in Grossman and Rossi-Hansberg (2008, 2012).

I now introduce "tree" technology that exhibits features of both snakes and spiders. Production of the final good can be represented as a set of sequential subchains assembled into more complex intermediate goods. Both snake and spider technologies are particular cases of this tree technology. A tree with one subchain is a snake, and a tree where all subchains are of length 1 is a spider. I illustrate the taxonomy of these technological assumptions in Figure 2.

Tree technology relies on only one restrictive assumption: sequentially produced intermediate parts can be assembled together but cannot be disassembled.[11] If this assumption is violated and one intermediate part is required to produce $N > 1$ intermediate parts at the downstream stage, the problem can still be tractable under a weaker assumption. In particular, if this single upstream part can be produced in different locations depending on which of N downstream parts uses it as an input, the original production structure can be re-written so that the single part is now represented as N parts with potentially different production

[11] This assumption seems reasonable: if the firm chose to assemble some parts, why would it disassemble them and reassemble them later? Surprisingly, this kind of behavior occurs in international trade. A classic example is a tariff on light trucks, also known as the chicken tax. In 1963 the United States introduced a 12% tax on light trucks in response to France and Germany increasing their tariffs on U.S. chicken. This tariff has not changed since, and car producers use loopholes to avoid it. For example, Ford imports its Transit Connect model from its plant in Turkey with rear seats and back windows so that this vehicle is classified as a wagon and is not subject to the chicken tax. These seats are shredded after Transit Connect crosses the border and windows are replaced with metal panels (http://www.wsj.com/articles/SB125357990638429655). Nevertheless, this example is an anecdote rather than a widespread pattern in international trade.

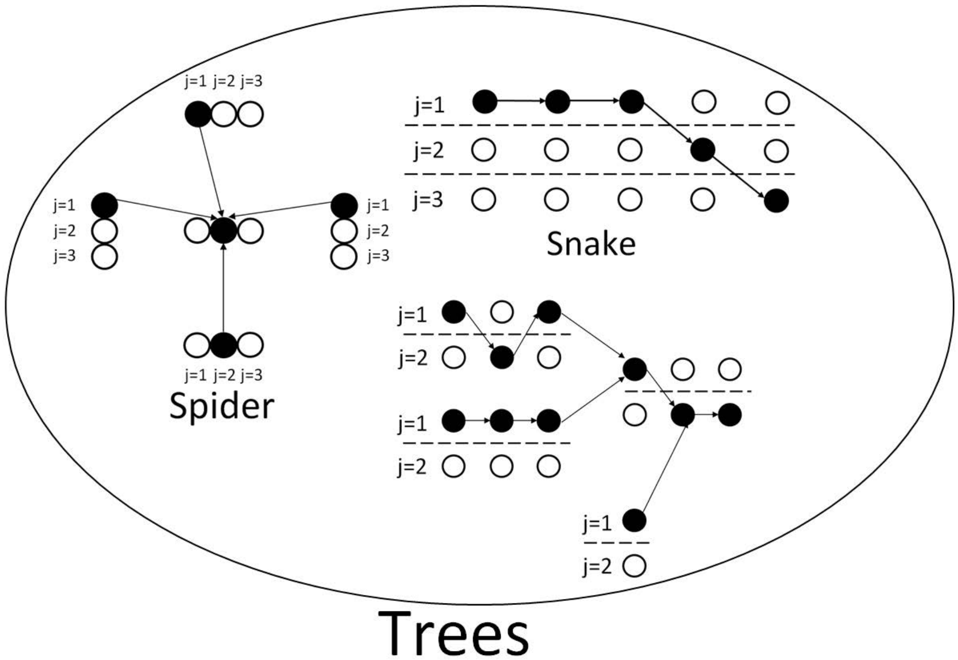

Figure 2. Taxonomy of Technological Assumptions

locations but with identical costs of production. This new production structure is a tree and.[12]

Technology. — I assume that the firm can have an arbitrary number of subchains that can be combined at any stage of production. Figure 3 illustrates the two-country example of such technology: some intermediate goods produced sequentially have to be assembled in the East or the West at a given stage. Note that stages do not uniquely identify parts, as more than one subchain can be produced at the same stage. I call a "node" a stage of production at a given branch, and then there is a one-to-one mapping between nodes and parts. An assembly node is a node where the firm combines two or more upstream parts in one complex part; the most up- and downstream nodes are "terminal" and the rest "nonterminal" nodes. A pair of nodes are colocated when parts corresponding to these nodes are produced in the same country. I impose no restrictions on the number of subchains and the number of assembly nodes.

I assume that intermediate assembly is costly and that these costs of assembly can differ by location. As intermediate parts can be combined but cannot be dis-

[12]This approach is similar to one proposed in Sraffa (1975), where he discusses the existence of an equilibrium in a general model, which can describe a wide range of production processes.

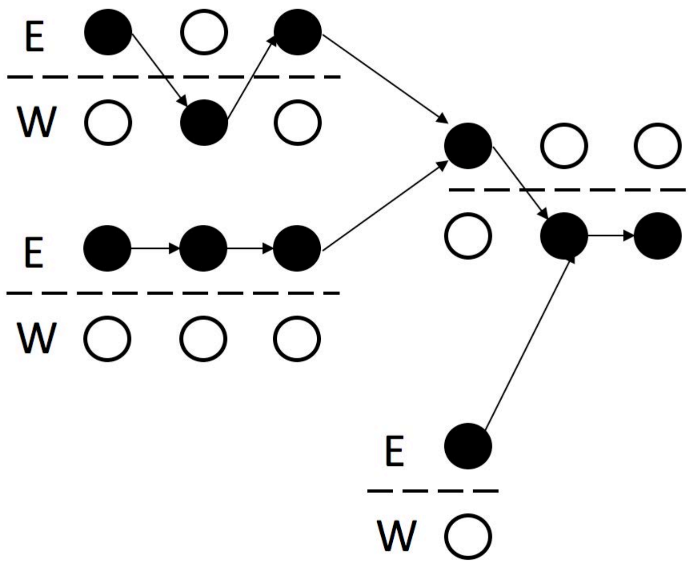

Figure 3. Tree Technology

assembled, the production structure will then look like a tree: there are multiple subchain branches that join at the subassembly points, becoming one final-good trunk in the end.

Complete Trees. — A tree is a very general organizational structure that describes a wide range of production processes including snakes and spiders. To make statements on the link between clustering patterns and general properties of trees, I impose a specific structure, which I call complete trees. A complete tree of order M and length N is a tree with N production stages, and at each stage of production, each intermediate good uses M parts from the upstream stage.

An important feature of complete trees is that any tree can be represented as a complete tree; thus a snake is a complete tree of order $M = 1$, and a spider is a complete tree of length $N = 2$ and order $M > 1$. An arbitrary tree can be represented as a complete tree of order M and length N, where M is the largest number of nodes connected to a single node, and N is the length of the longest subchain in such a tree. The number of nodes in a resulting complete tree is larger than the number of nodes in the original tree; each node in a complete tree that does not have a counterpart in the original tree has zero production costs in each country. The optimal allocation in the original tree and its general counterpart will then correspond to the same optimal value and same optimal locations for nondegenerate nodes.

Such a representation is illustrated in Figure 4. The left diagram represents an incomplete tree with the three most upstream parts required to produce a good at the second stage, and two second-stage goods are assembled into the final good at the most downstream stage. Each part can be produced in any of K countries, which is not reflected in the figure for the sake of clarity. The largest number of connected nodes in this example is 3, and the length of the longest subchain is also 3, so the corresponding complete tree will be of order 3 and length 3. The right diagram represents such a tree, with white circles representing degenerate nodes with zero production costs.

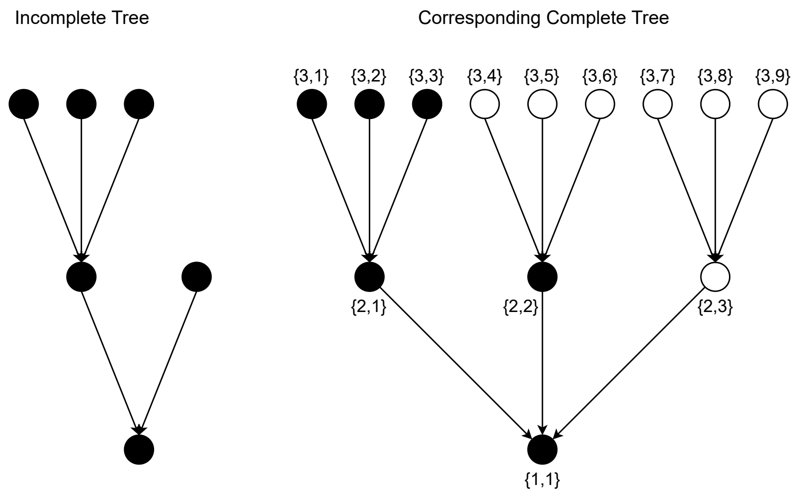

Figure 4. Tree Technology

Complete trees have two major advantages. First, they are characterized by three parameters: order, length, and the number of countries. This allows me to introduce some regularity to trees and link clustering forces to the general properties of the trees. The second advantage is that complete trees' fractal nature simplifies the notation of nodes and stages of production and allows for a universal optimization procedure, which will immediately work with any complete tree.

Each node $\{i, b\}$ is characterized by the stage of production $i \in \{1, ..., N\}$ with larger values corresponding to upstream stages and a number within a stage of production $b = \left\{1, ..., M^{i-1}\right\}$. A given node $\{i, b\}$ is connected to M upstream nodes indexed $\{i+1, M \times (b-1)+1\}, ..., \{i+1, M \times b\}$ for $i < N$ and to one downstream node $\left\{i-1, \left\lceil \frac{b}{M} \right\rceil\right\}$ for $i > 1$. This enumeration is illustrated in panel (b) of Figure 4.

To preserve the fractal nature of complete trees, I assume that both the final assembly and distribution tasks take place at the most downstream stage in the destination country. I relax this assumption in Section II.C. Below is the cost function for complete trees:

$$(3) MC_d = \min_{\{c_{i,b}\}_{i=2}^{N}} \sum_{i=1}^{N} \sum_{b=1}^{M^{i-1}} \left(\sum_{k=1}^{K} \mathbf{1}\left(c_{i,b}=k\right) a_{i,b,k} + \tau T\left(c_{i,b}, c_{i-1,\lceil \frac{b}{M} \rceil}\right) \right), \quad c_{1,1}=d$$

Expressions (1) and (3) look different because of the more complex indexing structure of a tree. The main idea remains the same — the costs of production of every part depend on its location choice and production location of the downstream part.

The potential disadvantage of such an approach is that some incomplete trees with a moderate number of nodes may have a corresponding complete tree with a large number of nodes, which makes the algorithm computationally burdensome. In this case, an allocation problem can be solved without switching to the complete tree. I provide an enumeration of nodes and the Bellman equation for this case in Appendix A2.

Algorithm for a Tree. — The Bellman equation corresponding to tree problem (3) is:

$$(4) \quad V_{i,b}\left(c_{i-1,\lceil \frac{b}{M} \rceil}\right) = \min_{c_{i,b}\in K} \left\{ \sum_{k=1}^{K} \mathbf{1}\left(c_{i,b}=k\right) a_{i,b,k} + \tau T\left(c_{i,b}, c_{i-1,\lceil \frac{b}{M} \rceil}\right) + \sum_{l=M(b-1)+1}^{M\times b} V_{i+1,l}\left(c_{i,l}\right) \right\},$$

where $c_{1,1}=d$ and $V_{N+1,b}\left(c_{N,b}\right)=0$ for $\forall b \in \left\{1,...,M^{N-1}\right\}$. For nonassembly nodes, the Bellman equation (4) is similar to equation (2): every subchain problem is solved similarly to the baseline chain problem. The difference arises when two or more subchains are combined. In this case, the value functions for each chain just add up. The underlying intuition is that all the decisions before the assembly point are made conditional on the assembly's downstream location. When the location of the assembly is chosen, it should minimize the sum of costs of all subchains being assembled.

Note that there are separate value functions (with index $\{i,b\}$) for different branches. As subchains are assembled, the value functions add up, and a new value function associated with a new joint branch appears. In the last stage, when the final good is produced, there is one branch (trunk) left, which is associated with a value of the single value function with index $\{i,b\}=\{1,1\}$.

Optimality, Existence, and Uniqueness. —

Proposition 1: *The recursive procedure described by the equation (4) always has a solution and finds all allocations that are the solutions to (3).*

PROOF:

The existence follows from the fact that MC is defined for any set of choice variables $\{c_{i,b}\}_{\{i,b\}=\{1,1\}}^{\{N,M^{N-1}\}}$ and each choice variable can take a finite number (K) of values.

To show optimality, let optimal path B $\{\tilde{c}_{i,b}\}_{\{i,b\}=\{1,1\}}^{\{N,M^{N-1}\}}$ be a solution to (3) that is different from path A found from (4) and consider 3 possible scenarios:

1) If for a most upstream node $\{N,b\}$, $\tilde{c}_{N,b} \neq \arg\min_{c_{N,b}\in K}\left\{\sum_{k=1}^{K}\mathbf{1}\left(c_{N,b}=k\right)a_{N,b,k}+\tau T\left(c_{N,b},\tilde{c}_{N-1,\lceil\frac{b}{M}\rceil}\right)\right\}$, then path B' that coincides with B for stages $1,...N-1$ and with

$$c_{N,b}=\arg\min_{c_{i,b}\in K}\left\{\sum_{k=1}^{K}\mathbf{1}\left(c_{i,b}=k\right)a_{i,b,k}+\tau T\left(c_{i,b},c_{i-1,\lceil\frac{b}{M}\rceil}\right)\right\}$$

will lead to lower value of MC, which contradicts optimality of B.

2) Let $v<N$ be the most upstream stage at which for some node $\{v,b\}$,

$$\tilde{c}_{v,b}\neq\arg\min_{c_{v,b}\in K}\left\{\sum_{k=1}^{K}\mathbf{1}\left(c_{v,b}=k\right)a_{v,b,k}+\tau T\left(c_{v,b},c_{v-1,\lceil\frac{b}{M}\rceil}\right)+\sum_{l=M(b-1)+1}^{M\times b}V_{v+1,l}\left(c_{v,l}\right)\right\},$$

then path B' that coincides with B except for node $\{v,b\}$ and all the nodes upstream to $\{v,b\}$. For node $\{v,b\}$ and its upstream neighbors they are the solutions to (4) with $c_{v-1,\lceil\frac{b}{M}\rceil}=\tilde{c}_{v-1,\lceil\frac{b}{M}\rceil}$.

3) If

$$\tilde{c}_{i,b}=\arg\min_{c_{i,b}\in K}\left\{\sum_{k=1}^{K}\mathbf{1}\left(c_{i,b}=k\right)a_{i,,bk}+\tau T\left(c_{i,b},c_{i-1,\lceil\frac{b}{M}\rceil}\right)+\sum_{l=M(b-1)+1}^{M\times b}V_{i+1,l}\left(c_{i,l}\right)\right\},$$

for $\forall i=\{1,..N\}$, $b=\left\{1,...,M^{i-i}\right\}$ then A and B coincide.

The optimality proof just finds the most upstream stage where an alternative path disagrees with the one found from the recursive procedure and then by construction of the recursive procedure, an upstream subchain, will lead to a smaller value of marginal costs, which contradicts the optimality of the alternative path.[13]

Finally, note that the Bellman equation (4) has a minimum operator at every stage, meaning that either there is a unique optimal choice at every stage, or, if an optimal choice is not unique, the firm is indifferent to the two or more sourcing choices. Consequently, the only possible situation when the optimal allocation is not unique is when all nonunique solutions correspond to the same value of total costs. In the stochastic formulation I introduce in Section III, the probability of this outcome converges to zero, and, for simplicity, from now on, I will assume that the optimal path is unique.

Induction Order. — All the algorithms described above started from the most upstream part and moved forward downstream; I call this logic forward induction. Interestingly, the snake allocation problem (1) can be solved with more intuitive backward induction.

This backward induction algorithm relies on the same optimal control algorithm but reverses the direction of forward induction. In this case, the firm is solving the problem of allocating N parts in M countries to minimize its production costs. It has technological restrictions on the order of production, but it does not have any terminal conditions on the production location of the first or the last stages. Neither does the order of production imply any direction; if the firm produces part i in a location different than $i-1$, it has to pay trade costs τ. It works, however, in both directions; if the firm produces part $i-1$ in a location different than i, it incurs the same costs τ. The Bellman equation (2) can then be rewritten as:

$$V_i\left(c_{i+1}\right)=\min_{c_i \in K}\left\{\sum_{k=1}^{K} \mathbf{1}\left\{c_i=k\right\} a_{i,k}+\tau T\left(c_i, c_{i+1}\right)+V_{i-1}\left(c_i\right)\right\},$$

with $c_1=d$, $V_0\left(c_1\right)=0$, and $\tau T\left(c_N, c_{N+1}\right)=0$.

Note that such reversal of the induction order is only possible for the snake case and specific trade costs. If the trade costs are iceberg, the value of these costs depends on the value of the product. In forward induction, the value of the product was equal to the value of the value function as it included parts that had already been produced by the time the border was being crossed, while the value function of the backward induction algorithm included the costs of parts that would be produced after the intermediate good was shipped.

This backward induction algorithm cannot be directly applied to the tree case because it uses the production location from the upstream stage as a state variable

[13] When dynamic programming is applied to the analysis of GVCs, there is usually no discounting between the stages; thus standard contraction mapping results do not hold. Kikuchi et al. (2019) offer an alternative set of conditions that ensure the existence and uniqueness of the dynamic program solution for a large class of recursive problems without a discount factor.

for the given node. In case part i is assembled with L_i intermediate parts, each of which can be produced in one of K countries, it gives K^{L_i} possible values of the state variable, which can be very large even for a moderate number of countries and assembled parts.

Decentralized Solution. — The forward induction logic suggests that the central planner's problem and a decentralized solution will lead to the same allocation. As noted above, the value of the value function at every stage is equal to the cost of the intermediate good. In other words, the price of an intermediate good at a given point is a sufficient statistic, and the firm does not need to know the whole previous path to make an optimal decision.

The decentralized solution can then be represented as the following environment. There are independent firms at every node in every country. Each of these firms chooses the lowest-cost supplier, which can offer them the lowest price of an intermediate product that includes shipping costs and depends on the costs of production and the costs of inputs that each supplier uses. As a result, each potential supplier makes an optimal choice but does not necessarily have a buyer. In the end, only firms that have buyers survive, and the optimal path is characterized by a sequence of surviving firms.[14]

Hence, predictions of the model do not depend on the boundaries of the firm, and they hold for any exogenously given organizational structure.

C. Iceberg Trade Costs

When trade costs depend on the value of the intermediate good, the firm's problem becomes more complicated: to make an optimal decision, the firm must know the value of an intermediate good.

Under the assumption of iceberg trade costs, every time an intermediate good crosses the border between countries i and j, a fraction $\frac{1}{\tau}$ of it "melts". In other words, the firm has to ship τ units of an intermediate good from country i to receive one unit of an intermediate good in country j. As the firm minimizes its per-unit costs, I reformulate the problem as follows: the firm produces 1 unit of each intermediate part until it chooses to cross the border. Whenever a border crossing between countries i and j takes place, the firm has to produce τ times more of parts on all upstream stages, and hence, the transportation costs it incurs are $\left(\tau T\left(c_{i,b}, c_{i-1\nu_{i,b}}\right) - 1\right)\chi_{i,b}$, where $\chi_{i,b}$ is the cost of an intermediate good crossing the border.

The costs of the firm will then depend on the quantity of intermediate inputs it has to produce and can be written as:

[14] In this interpretation, my model becomes a stylized model of endogenous network formation. See Lim (2018), Oberfield (2018), and Tintelnot et al. (2018) for more complex models.

$$MC_d = \min_{\{c_{i,b}\}} \sum_{i=1}^{N} \sum_{b=1}^{M^{i-1}} \sum_{k=1}^{K} \mathbf{1}\left(c_{i,b}=k\right) a_{i,b,k}\chi_{i,b},\ c_{1,1}=d \tag{5}$$

$$\chi_{i,b} = \chi_{i-1,\lceil \frac{b}{M} \rceil} \left(\tau T\left(c_{i,b}, c_{i-1,\lceil \frac{b}{M} \rceil}\right)+1\right),\ \chi_{1,1}=1. \tag{6}$$

The expression for the Bellman equation is then straightforward:

$$V_{i,b}\left(c_{i-1,\lceil \frac{b}{M} \rceil}\right) = \min_{c_{i,b}\in K}\left\{\tau T\left(c_{i,b}, c_{i-1,\lceil \frac{b}{M} \rceil}\right)\left[\sum_{k=1}^{K}\mathbf{1}\left(c_{i,b}=k\right) a_{i,b,k} + \sum_{l=M(b-1)+1}^{M\times b} V_{i+1,l}\left(c_{i,l}\right)\right]\right\}, \tag{7}$$

with $c_{1,1}=d$ and $V_{N+1,b}\left(c_{N,b}\right)=0$ for $\forall b\in\left\{1,...,M^{N-1}\right\}$.

The only difference between this Bellman equation and equation (4) is that trade costs now depend on the value of the intermediate good at stage $\{i,b\}$, so the trade costs term is now multiplicative rather than additive as in equation (4).

D. *Comparison with Linear Programming*

In the case of specific trade costs, the location decision for each node affects only costs associated with location decisions for the adjacent up- and downstream nodes. An alternative way to solve the allocation problem is to use zero-one integer linear programming (LP). The minimization problem requires an alternative formulation where choice variables are not production locations at each node $\{i,b\}$ $c_{i,b}$ but a set of dummies $\xi_{i,b,c,k,l}$ that code the position of a pair of connected nodes $\{i,b\}$ and $\{i-1,c\}$ for $i>2$ each possible pairwise production locations k and l in each of these nodes. Each pair of nodes is then coded by $K\times K$ dummy variables and the total number of pairs of connected nodes $\frac{M^N-M}{M-1}-1$. The production location of the most downstream stage $i=1$ is represented by K dummy variables κ_{1k}. So the total number of variables is equal to $K^2\left(\frac{M^N-M}{M-1}-1\right)+K$

Note that each nonterminal node is coded twice by these variables — once as an upstream node in a pair and as a downstream node another time — so there should be constraints that would ensure that the production location of the same node is consistent among the variables it is included in. These constraints can be written as:

$\sum_{k=1}^{K}\kappa_{1,k}=1$, $\sum_{l=1}^{K}\xi_{2,b,1,k,l}=\kappa_{1,k}$, $\sum_{l}^{K}\xi_{i,b,c,k,l}=\sum_{m}^{K}\xi_{i-1,c,d,m,n}$

The first constraint ensures that the most downstream part is produced in one country only. The second set of constraints regulates the consistency in production locations between the most- and second downstream stages and the third set of constraints does the same for the remaining pairs of stages. The total number of constraints is $1 + K\left(\frac{M^{N-1}-1}{M-1} - 1\right)$

The firm's marginal costs can then be expressed as:

$$MC = \sum_{k=1}^{K} \kappa_{1,k} a_{1,k} + \sum_{i=2}^{N} \sum_{b=1}^{M^{i-1}} \sum_{k=1}^{K} \sum_{l=1}^{K} \xi_{i,b,\lceil \frac{b}{M} \rceil,k,l} a_{i,b,k} + \tau \sum_{i=2}^{N} \sum_{b=1}^{M^{i-1}} \sum_{k=1}^{M} \left(1 - \xi_{i,b,\lceil \frac{b}{M} \rceil,k,k}\right)$$

so the LP problem can be written down as:

$$\min_{\kappa_{1,k},\xi_{i,b,c,k,l}} MC \text{ s.t.}$$

$$\sum_{k=1}^{K} \kappa_{1,k} = 1,\ \kappa_{1,k} \in \{0,1\}$$

$$\sum_{l=1}^{K} \xi_{2,b,1,k,l} = \kappa_{1,k},\ \forall k \in K,\ \forall b \in \{1,...,M\},\ \xi_{2,b,1,k,l} \in \{0,1\}$$

$$\sum_{l}^{K} \xi_{i,b,c,k,l} = \sum_{m}^{K} \xi_{i-1,c,d,m,n},\ \forall i = \{3,..,N\},\ \forall k,n \in K,\ \forall b \in \left\{1,...,M^{i-1}\right\},\ c = \left\lceil \frac{b}{M} \right\rceil,\ \xi_{i,b,c,k,l} \in \{0,1\}$$

Note that this approach would not be possible for the case of iceberg trade costs in which location decisions made at any stage affects the cost of an intermediate good and hence the total transportation costs on all downstream stages.

For the case of specific trade costs, this approach allows me to solve the allocation problem but it has to operate with a large number of variables and constraints, while the recursive approach of the Bellman equation allows me to break down the problem into a sequence of simple and easy to solve subproblems.

In Figure 5 I present the ratio of computational speeds of LP and recursive algorithms for different values of N, M, and K.[15] I use $N = 5$, $M = 2$, and $K = 2$ as a baseline specification and show the comparative statics of the computational speed ratio with respect to M in panel (a), with respect to N in panel (b), and with respect to K in panel (c). Computational speed depends on a particular cost draw, so for each set of parameters I used 1000 cost draws and used the median computational time by each method.

From Figure 5 one can see that the recursive algorithm strictly dominates the LP algorithm. First, the relative computational time ranges between a factor of 16 for a short snake and 700 for more complex trees. In addition, the LP algorithm

[15]Preparing the problem for the LP solver requires multiple operations with large matrixes and takes significant time, up to 5 times more than the algorithm itself requires. I use *intlinprog* package for Matlab on a computer with 16GB RAM.

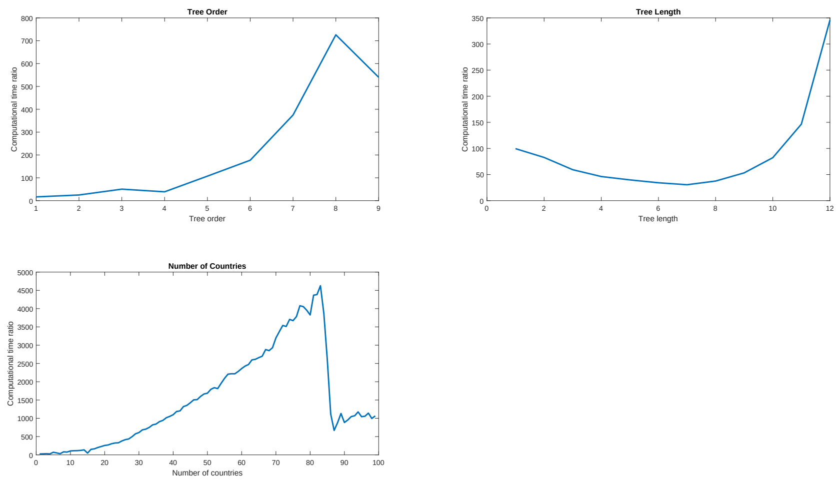

Figure 5. Computational Speed of Linear Programming and Recursive Algorithms

uses a considerable amount of operative memory, so on a computer with 16 GB of RAM, it works for only $N < 13$ and $M < 10$ when the only limitation of a recursive algorithm is the maximum size of the matrix that stores the data on production costs.

II. Properties of the Model

There is no simple analytical solution for the general cases of chain and tree technology; however, some properties of firms' behavior can be derived. In this section, I introduce theoretical results that hold for any chain and tree problem. I primarily focus on comparative statics with respect to τ, a single parameter reflecting openness for trade. Changes in τ can be interpreted as multilateral trade liberalization, or more generally, as a proportionate decrease in offshoring costs.

A. Multilateral Trade Liberalization

Proposition 2: *In the optimum the firm's total costs are nondecreasing in trade costs.*[16]

[16] Here and further, I state propositions and theorems with weak monotonicity for two reasons: first, if trade costs are so high that offshoring is impossible, some of the comparative statics related to offshoring do not work; and second, the firm has a finite number of optimal choices, and then the firm's optimal choice cannot change with any infinitesimal change in a parameter value. To handle the first problem,

PROOF:

See Appendix A1

This proposition is straightforward: when the firm faces lower trade costs, if it does not change its production decision it will face the same or lower total production costs. Then, there is no way a new optimal path is more costly than the old one. Note that the proof does not use any assumption on a production structure and relies only on the firm's revealed preferences argument. This means that this result will hold for a large class of firm problems.

Now I decompose the costs of the firm. Let function $NTMC\left(Y\right) \equiv \sum_{i=1}^{N} \sum_{b=1}^{M^{i-1}} \left(a_{i,b,k} \mathbf{1}\left\{c_{i,b} = k\right\}\right)$ be a value of nontransport marginal costs for path Y. Let $TTMC \equiv MC - NTMC$ be a value of total trade costs that can be represented as $TTMC = \tau TQ$. I call TQ transportation quantity, as it reflects transportation schedules independent of the trade cost price shifter τ. For specific trade costs, one can think of TQ as the total number of miles a transportation ship traveled, and τ as the price of gas.[17] For iceberg trade costs, TQ can be interpreted as the total costs of goods that crossed the border or trade volumes.

LEMMA 1: *The transportation quantity is a nonincreasing function of τ.*

PROOF:

See Appendix A1

The intuition behind this proposition for the case of specific trade costs is the following: when the price of gas decreases, the firm has no incentives to decrease the number of miles traveled, even though total transportation expenses can increase or decrease.[18] Similarly, for iceberg trade costs, when the "melting rate" τ decreases, the firm has no incentives to trade less, even though the total volume of "melted" goods can increase or decrease.

Proposition 3: *Provided there is some offshoring, the firm's optimal total costs are increasing in trade costs.*

PROOF:

See Appendix A1

Proposition 4: *If the firm changes its unique optimal path due to a decrease in τ, then the nontransportation costs of production ($NTMC$) decrease.*

it is enough to assume that trade costs are not too high and offshoring is possible. The second problem disappears when a large number of firms are taken into consideration: with a continuum of firms, changes in parameter values lead to a change in the optimal path for at least some of the firms.

[17]In case transportation costs are similar for all country pairs $T_{ij} = T_{kl}$ for $\forall i, j, k, l \in \{1, ..., M\}, i \neq j, k \neq l$, τ can be interpreted as the number of border crossings.

[18]In the case of similar transportation costs between all country pairs, Lemma 1 indicates that the number of border crossings is a nonincreasing function of τ. Then by defining a cluster as a sequence of parts produced in the same country, the following statement is true: The average size of a cluster is a nondecreasing function of τ. It simply follows from the fact that the average size of a cluster is equal to $s = \frac{N}{m+1}$, where m is the number of border crossings. As by Lemma 1, m is nonincreasing in τ and, hence, s is nondecreasing in τ.

PROOF:

See Appendix A1

Proposition 4 is the key proposition that represents gains from fragmentation. Not surprisingly, the firm increases its total productivity when trade costs decrease: if the firm is engaged in offshoring, it pays less for transportation. Proposition 4, however, reveals another channel through which productivity increases: the firm's optimal path depends on trade costs. With a change in trade costs, the firm can choose a different production structure that would lead to higher production efficiency. Similar to Proposition 2, this result does not rely on the sequentiality of the production structure.

Baqaee and Farhi (2019) make a related point and decompose gains from trade in a general equilibrium framework with international production by two channels — technology and reallocation channels. They further derive a sufficient statistic for this decomposition, which is not possible in the present framework due to the complex nature of the optimization problem.

Proposition 5: *Production in a given country can depend on trade costs non-monotonically. Reshoring is possible.*

PROOF:

Consider the following numerical example with snake technology, two countries and five stages of production: $a_1 = \{4, 4, 4, 4, 4\}$, $a_2 = \{10, 2, 5, 2, 10\}$. Then

1) If $\tau \geq 1.5$, $c = \{1, 1, 1, 1, 1\}$
2) If $1.5 > \tau \geq 0.5$, $c = \{1, 2, 2, 2, 1\}$
3) If $0.5 > \tau \geq 0$, $c = \{1, 2, 1, 2, 1\}$

Figure 1 illustrates Proposition 5. Black dots represent the firm's choice to produce a part in a given country. The dotted line represents the border between the two countries. Arrows represent the firm's optimal path. With high trade costs, the firm chooses to produce the whole good in the first country. With decreased trade costs, it chooses to offshore a large cluster to the second country. When trade costs decreased even further, the firm fragmented its production more and reshored the third part back to the first country.

Reshoring is only possible for the number of stages larger than 2. In case $N = 2$, the tree degenerates to a spider, where the clustering mechanism is inactive. The production location problem for the destination market c_1 can be solved independently for each node $\{2, b\}$: $c_{2,b} = \arg\min \{\mathbf{1}(c_{2,b} = k)\, a_{2,b,k} + \tau T(c_{2,b}, c_{1,1})\}$ for specific trade costs and
$c_{2,b} = \arg\min \{\mathbf{1}(c_{2,b} = k)\, a_{2,b,k}\, (1 + \tau T(c_{2,b}, c_{1,1}))\}$ for iceberg trade costs. $c_{2,b}$ then exhibits monotone behavior with $c_{2,b} = c_{1,i}$ for high values of τ and $c_{2,b} = \arg\min \{\mathbf{1}(c_{2,b} = k)\, a_{2,b,k}\}$ for small values of τ.

B. Limiting Cases

Generally, there is no closed form solution for the marginal costs of production of the final good in problems (1) and (3). The reason is not a drawback of some modeling assumptions:[19] the interdependence of production on different stages both generates clustering effects and makes a problem harder to solve. In Section I, I show that the interdependence of decisions on different stages of production leads to a complicated solution; the optimal path depends on the value of $M \times N + 1$ parameters: i.e. the costs of production and trade costs. The solutions for limiting cases, however, are trivial:

1) If $\tau = 0$, $MC = \sum_{i=1}^{N} \sum_{b=1}^{M^{i-1}} \min_{k \in K} \{a_{i,b,k}\}$

2) If $\tau = \infty$, $MC = \min_{k \in K} \left\{ \sum_{i=1}^{N} \sum_{b=1}^{M^{i-1}} a_{i,b,k} \right\}$.

In the free trade case, the firm simply chooses to produce each part in the cheapest location. In case $\tau = \infty$, offshoring of parts is impossible, and the firm chooses the cheapest location to produce the whole good. Note that $\min_{k \in K} \left\{ \sum_{i=1}^{N} \sum_{b=1}^{M^{i-1}} a_{i,b,k} \right\} \geq \sum_{i=1}^{N} \sum_{b=1}^{M^{i-1}} \min_{k \in K} \{a_{i,b,k}\}$ with an equality sign only in case there exists such country k that $a_{i,b,j} \geq a_{i,b,k}$ for $\forall i, b, j$. These two limiting cases represent two states of the Ricardian world: when countries specialize in the production of parts and in the production of final goods. When trade costs decrease, more opportunities arise to exploit productivity differences between countries:

Proposition 6: *For any* $\tau \in (0, \infty)$, $\sum_{i=1}^{N} \sum_{b=1}^{M^{i-1}} \min_{k \in K} \{a_{i,b,k}\} \leq MC(a, \tau) \leq \min_{j \in K} \left\{ \sum_{i=1}^{N} \sum_{b=1}^{M^{i-1}} a_{i,b,k} \right\}$.

PROOF:

Follows directly from Proposition 2.

In particular, it means that every firm has limited potential to gain from offshoring: gains in production efficiency of the firm are limited by $\min_{k \in K} \left\{ \sum_{i=1}^{N} \sum_{b=1}^{M^{i-1}} a_{i,b,k} \right\} - \sum_{i=1}^{N} \sum_{b=1}^{M^{i-1}} \min_{k \in K} \{a_{i,b,k}\}$.

[19]One reasonable way to simplify the problem is to assume that the firm does not know its costs at any given stage until it builds a plant. This assumption makes the problem trivial; the firm will produce all the parts in the country that has lower *ex ante* costs. If shipment costs differ for some countries, the firm can face a trade-off between production efficiency and proximity to consumer markets described in Section II.C; however, under this assumption the clustering mechanism becomes redundant.

C. Final Good Shipping

In Section I, I assumed that the final good is assembled in the destination market. I did so to preserve the fractal nature of complete trees. In this subsection, the firm can assemble the final good in a country different than the destination country d. To differentiate the cost minimization problem and the proximity to consumer market consideration, I introduce two different kinds of trade costs: shipping costs of intermediate goods τ and costs of shipping of the final good τ^F. The firm can choose different optimal paths for the production of final goods with different destination countries.[20]

I model shipping and distribution as an additional stage 0 and interpret costs associated with the remaining stages as production costs. The firm would have the following cost minimization problem for each destination country:

$$MC_d = \min_{\left\{c_{i,b}\right\}_{\{i,b\}=1,1}^{N,M^{N-1}},} \sum_{i=1}^{N} \sum_{b=1}^{M^{i-1}} MC_{i,b} + \tau^F T^F \left(c_{1,1}, d\right) \tag{8}$$

$$MC_{i,b} = \sum_{k=1}^{K} \left(\mathbf{1}\left(c_{i,b} = k\right) a_{i,b,k} + \tau T\left(c_{i,b}, c_{i-1,\lceil \frac{b}{M} \rceil}\right)\right), \tag{9}$$

where d is the index of the destination country, $T^F(i,j)$, and τ^F is a final good shipment-costs matrix and multiplier respectively. I introduce τ^F for two main reasons: first, tariffs on final and intermediate goods can be different, and second, the costs of offshoring may include other factors in addition to the costs of transportation and tariffs.

The firm chooses an optimal path to minimize the sum of production costs and shipment costs. Depending on the relative size of τ and τ^F, the firm will assign different weights to considerations on the production costs and costs of shipping to the destination markets.

Proposition 7: *For sufficiently large τ^F, optimal paths with different destinations are different.*

PROOF:

For $\tau^F > \sum_{k=1}^{K} \sum_{i=1}^{N} \sum_{b=1}^{M^{i-1}} a_{i,b,k}$ the firm chooses to produce the final part in the destination country.

Note that only one optimal path minimizes the costs of production (or the firm is indifferent between more than one path).

Corollary:

For a sufficiently large τ^F, an optimal path does not minimize the marginal costs of production.

[20] In the present model, the firm does not face a complicated export-platform problem, as in Tintelnot (2017), because in the present model, there are no fixed costs of opening a plant and the firm solves the problem of reaching the consumer markets independently for each destination country.

D. Effects of FTA

This section shows that bilateral trade liberalization in a multicountry framework can lead to unexpected results.

Proposition 8: *In the case of three countries, trade liberalization between two countries may increase production in a third country.*

PROOF:

A numerical example can be provided. Assume there are three countries, country 1 is far from countries 2 and 3, and countries 2 and 3 are close: $\tau^0_{12} = \tau^0_{21} = 2$, $\tau_{13} = \tau_{31} = 2$, $\tau_{23} = \tau_{32} = 0.5$. The final good consists of three parts and their costs of production are equal: $a_{11} = 2$, $a_{12} = 7$, $a_{13} = 2$ for country 1; $a_{21} = 8$, $a_{22} = 5$, $a_{23} = 8$ for country 2; and $a_{31} = 8$, $a_{32} = 8$, $a_{33} = 2$ for country 3. Then the cost minimizing decision would be to produce all parts in the first country. Now if trade costs between countries 1 and 2 decrease $\tau^1_{12} = \tau^1_{21} = 1$, then the optimal decision is to produce the first part in country 1, second in country 2 and third in country 3. Therefore, a decrease in trade costs between countries 1 and 2 increases production in country 3.

The third country benefits because, before trade liberalization, the production costs of part 3 in country 3 were low, but not low enough to make offshoring of this part to country 3 profitable due to high trade costs. With the decrease in trade costs, part 2 became offshored to country 2, but as parts 2 and 3 became adjacent, trade costs between countries 1 and 3 no longer matter, and the firm faces lower trade costs between countries 2 and 3. I provide the following example: with high trade costs, a U.S. firm chooses not to offshore its production. With the decrease in trade costs with Malaysia, this firm may want to offshore some production stages there. However, as these parts are offshored, there may be an advantage to offshore adjacent parts to Indonesia, which is close to Malaysia.

Similar to Proposition 5, the FTA result is driven by the clustering mechanism and works for $N > 2$ only. For spider ($N = 2$), the sourcing location choice for each node $2b$ is made independently: $c_{2,b} = \arg\min \left\{ \mathbf{1}\left(c_{2,b} = k\right) a_{2,bk} + \tau_{c_{2,b}c_1} T\left(c_{2,b}, c_{1,1}\right) \right\}$ for specific trade costs and $c_{2,b} = \arg\min \left\{ \mathbf{1}\left(c_{2,b} = k\right) a_{2,bk} \left(1 + \tau_{c_{2,b}c_{1,1}} T\left(c_{2,b}, c_{1,1}\right)\right) \right\}$ for iceberg trade costs. If bilateral trade liberalization affects sourcing decisions, it increases trade between countries that joined the FTA and hence decreases production in the third country.

E. Total Costs Distribution

All the problems described in Section I took production costs as given. It is unlikely, however, to obtain the data on costs in each particular stage of production for each firm. Moreover, to solve this problem, one would need to know not only the actual costs of production but also opportunity costs of production of these parts in other countries. To further analyze the link between tree structure and clustering, in the next section, I follow Yi (2010), Ramondo and Rodrıguez-Clare (2013), and Johnson and Moxnes (2019) among others, and assume that costs of production at each stage in each country follow some random variable. The standard assumption is the Fréchet distribution, popular because it leads to the closed form solution of many models. Fréchet, however, is not the only possible choice; in this section, I am not making distributional assumptions to make my analysis as general as possible.

Every firm draws $M^{N-i} \times K$ matrix A_i of costs for each stage i from some distributions $F_j(a)$, $j \in K$. Here I assume that these draws are i.i.d. Facing the sequence of cost matrixes $A \equiv \{A_1, ..., A_K\}$ and trade costs τ, the firm solves its problem and has optimal marginal costs that depend on production and transportation costs $MC(A, \tau)$. As elements of A are random variables, $MC(A, \tau)$ is a random variable as well; the distribution of optimal marginal costs is then a function of parameters θ_j of distribution $F_j(a)$, $j \in M$: $G_{MC}(\theta, \tau)$, where $\theta = \{\theta_1, ..., \theta_M\}$.

Proposition 9: *If $\tau_1 < \tau_0$, random variable $MC(\theta, \tau_0)$ weakly*[21] *first-order stochastically dominates $MC(\theta, \tau_1)$.*

PROOF:

By Proposition 2, $MC(A, \tau)$ is nondecreasing in τ. This means that for any given matrix set of draws A $Pr(MC(A, \tau_0) < x) \leq Pr(MC(A, \tau_1) < x)$. At the same time, by definition of $MC(\theta, \tau_0)$, random variables $MC(\theta, \tau_0)$ and $MC(A, \tau_0)$ follow the same distribution. This means that $Pr(MC(A, \tau) < x) = Pr(MC(\theta, \tau_1) < x)$; hence, $Pr(MC(\theta, \tau_0) < x) \leq Pr(MC(\theta, \tau_1) < x) \Rightarrow MC(\theta, \tau_0)$ first order stochastically dominates $MC(\theta, \tau_1)$

Corollary: Distribution $MC(\theta, \tau)$ first-order stochastically dominates $MC(\theta, \tau_0)$ and is dominated by $MC(\theta, \tau_1)$ for $\tau \in (\tau_0, \tau_1)$.

In particular, $MC(\theta, \tau)$ is bounded by two well defined distributions: $\sum_{i=1}^{N} \sum_{b=1}^{M^{i-1}} \min_{k \in K} \{a_{i,b,k}\}$ and $\min_{k \in K} \left\{ \sum_{i=1}^{N} \sum_{b=1}^{M^{i-1}} a_{i,b,k} \right\}$.

Figure 6 illustrates Proposition 9. I simulate 10,000,000 firms with $N = 5$, $M = 2$, $K = 2$, and the Fréchet distribution of the draws with the shape parameter 4.12 (Simonovska and Waugh (2014)). Note that the distribution of firms' marginal costs does not just shift as trade becomes cheaper, it changes its shape.

III. Simulations

This section analyzes how trade patterns depend on the tree order, its length, and the number of countries. Due to the complexity of the problem, there is no

[21] Weak dominance appears in the case of large τ_0 and τ_1 such that there is no offshoring. In this case, changes in trade costs do not affect the productivity of the firms.

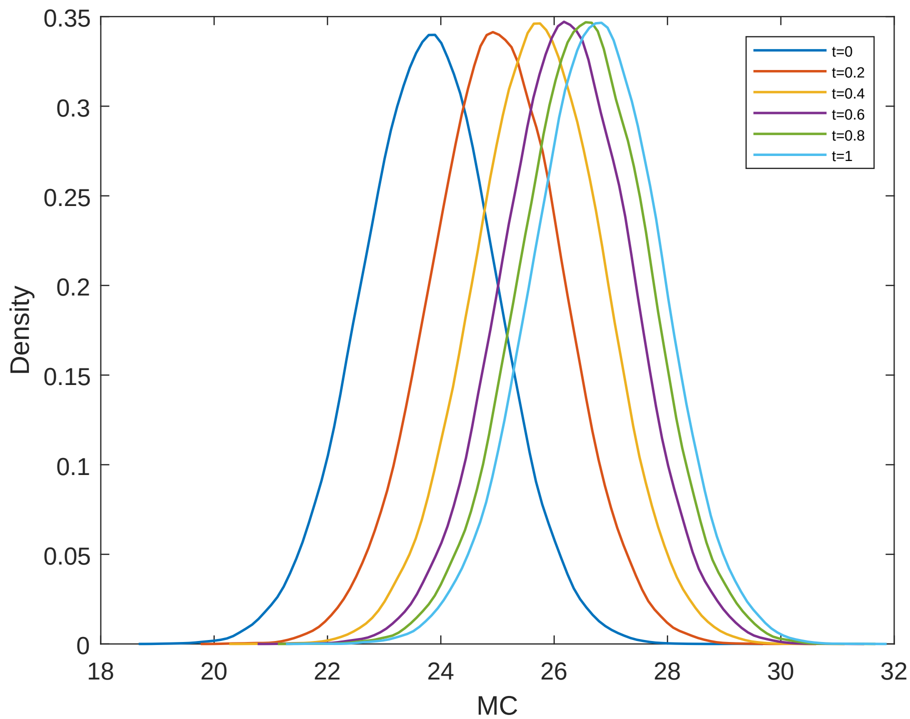

Figure 6. Evolution of Firms' Marginal Costs

closed form solution, so I rely on Monte Carlo simulations. To make my results comparable with the rest of the literature, I employ the Frećhet distribution of production costs $a_{i,b,k}$ in each node in each country with a common shape parameter of 4.12 (Simonovska and Waugh (2014)).

I choose a tree with $N = 5$, $M = 2$, and $K = 2$ as a baseline formulation because it defines the smallest tree that generates all the clustering patterns discussed in this section. Direct application of (3) will generate K optimal paths, conditional on the destination market; for the sake of clarity, I use (8) with $\tau^F = 0$ instead and allow parts in the most downstream production stage 1 to be produced in any country.

As a measure of clustering between stages of production i and $i-1$, I use the share of corresponding colocated nodes $S_i = \frac{\sum_{b=1}^{M^{i-1}} \mathbf{1}\left(c_{i,b}=c_{i-1,\lceil \frac{b}{M} \rceil}\right)}{M^{i-1}}$. This simple measure allows me to directly compare different trees, specific and iceberg trade costs, and analyze clustering between nonadjacent nodes. I provide the corresponding results for trade elasticities in Appendix A4.[22]

In this section, I focus on analyzing specific trade costs because, with iceberg trade costs, effective trade costs are much larger for downstream stages, especially for long trees; these stage-specific trade costs distort clustering behavior and make it more difficult to analyze. This effect is similar to the centrality-downstreamness nexus results in Antràs and De Gortari (2020), so I do not concentrate on it.[23] For each figure with specific trade costs, I provide its counterpart with iceberg trade costs in Appendix A3.

I will consider clustering for both pairs of nodes that are directly connected and not connected. To keep track of various types of relationships, I use family terms. I say that directly connected nodes exhibit parent-child or filial relationships, with the unique downstream node being a parent and multiple upstream nodes the children. Nodes that belong to the same stage and share a common parental node are sister nodes and so forth.

For a tree of length M, there will then be $M-1$ measures of clustering for each pair of adjacent stages. A pair that includes the most upstream stage is referred to as the upstream pair, and the next one is referred to as the second upstream pair. I follow this logic for all the pairs except the one that includes the most downstream node, which I call the most downstream pair.

A. *Order of a Tree*

Figure 7 shows how tree order shapes clustering patterns. As a baseline, I take a tree of length $M = 5$, $K = 2$ countries, and simulate $S = 100,000$ firms. Different panels correspond to the order of the tree. The top left figure corresponds to the case of $M = 1$ pure snake. In this case, the strength of clustering is identical for the most up- and downstream nodes and for the second and third upstream nodes. This happens because, in the case of specific trade costs, up- and downstream nodes are symmetric in the sense that each is adjacent to only one other node. As I discussed in Section I.B, both forward and backward induction algorithms allow finding the optimal allocation, and the problem itself looks similar in either direction.

The second and third most upstream stages also exhibit the same level of clustering for the same reason. Most up- and downstream stages have a higher

[22]An important difference between S_i and elasticities is that to compute the latter, for a given value of trade costs, a firm has to change its optimal allocation, which may happen to a small fraction of firms or not happen at all. As a result, elasticities are noisier measures of clustering which are not defined for some values of trade costs.

[23]While in general, it is reasonable to expect that more expensive downstream goods will be subject to higher trade costs, this is not necessarily the case. In the case of physical shipment, some upstream components such as steel and metal parts can be expensive to ship compared to the shipment of, for example, downstream electronic components that are lightweight and do not take up much space. Many countries do not charge tariffs on processing trade. Finally, even in relatively small trees, the cost of the intermediate good increases quickly for downstream stages, with the ratio of costs of the most down- and upstream stages approximately equal to M^{N-1}, which for larger trees is unlikely to reflect the actual trade costs ratio.

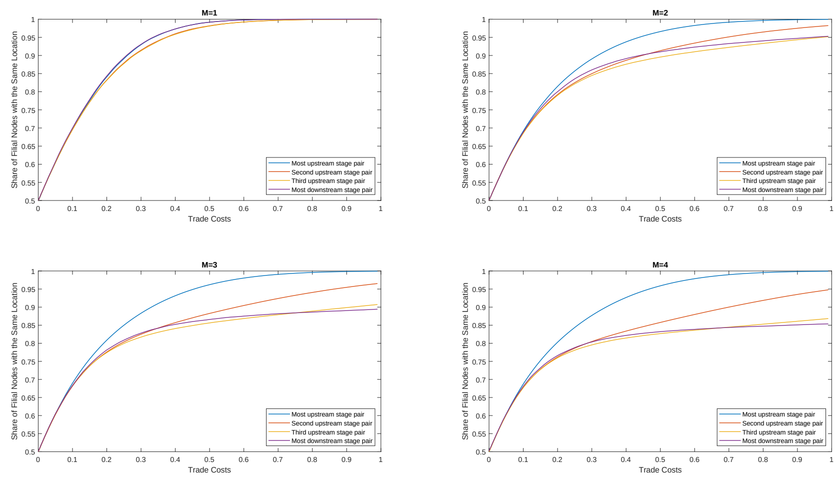

Figure 7. Clustering and Tree Order

strength of clustering because there is only one adjacent node for each, which drives clustering decisions, while the second and third nodes have to account for both their parents' and children's locations.

This situation changes for $M = 2$, as the symmetry is violated. The most upstream nodes, for instance, have only one parent and no children, which makes it easy for this node to cluster with its parent. At the same time, the most downstream node has $M > 1$ children. If all of the intermediate goods in these children nodes are produced in different countries, then the most downstream node can cluster with no more than one child. This asymmetry holds for all the nonterminal nodes as well — each of them has M children and just one parent.

As most upstream nodes have only one parent and no children, colocating with their parental node is easier, so for all values of M, they exhibit higher levels of clustering. The most downstream node does not have a parent, so it has one less adjacent node than nonterminal nodes, making it easier for the most downstream node to cluster for low values of trade costs. This advantage, however, becomes less significant as the total number of children nodes increases as the tree order increases.

For larger values of trade costs, the second and then third upstream nodes cluster more than the most downstream node. The reason is that most upstream nodes are easy to cluster and, for large enough values of trade costs, they automatically cluster with their parent. If the production location in this parent node changes, so will production locations for its children. This second to the most upstream node can then focus on clustering with its parent. In other words, the more children nodes cluster with their parent node, the easier it becomes for this parent node to cluster with its parent; thus high clustering on upstream production stages increases clustering on downstream stages.

Interestingly, as illustrated in Figure 8, iceberg trade costs invert the clustering patterns discussed above. Downstream nodes face higher trade costs and thus have higher incentives to cluster. Unlike in the case of specific trade costs, a higher tree order makes clustering stronger for downstream stages and weaker for upstream stages. The effects described for the specific trade costs are still present — intermediate inputs have more upstream neighbors in trees of higher order. At the same time, for a tree with a given length and cost distribution, the average cost ratio of the most downstream to the most upstream part increases with the speed of a polynomial of $M-1$-th degree. As for larger M, the cost of a good increases further along the value chain, as do the trade costs, which creates more clustering pressure for the downstream stages.

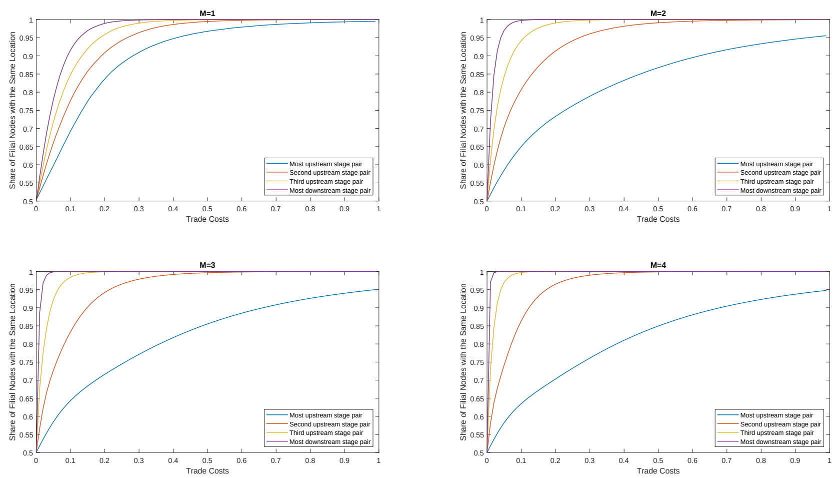

FIGURE 8. CLUSTERING AND TREE ORDER: ICEBERG TRADE COSTS

B. Length of a Tree

Figure 9 illustrates the link between the length of a tree and clustering for $M = 2$, $K = 2$, and $S = 100,000$. The effect of tree length is the strongest for the most downstream stage and is effectively zero for the most upstream stage. The reason is the mechanism of clustering propagation described in Section III.A, that is, the problem of the most upstream nodes does not directly depend on the order or the tree's length. Each stage further downstream, however, depends on the clustering strength on all the upstream stages. Thus clustering propagates more in longer trees.

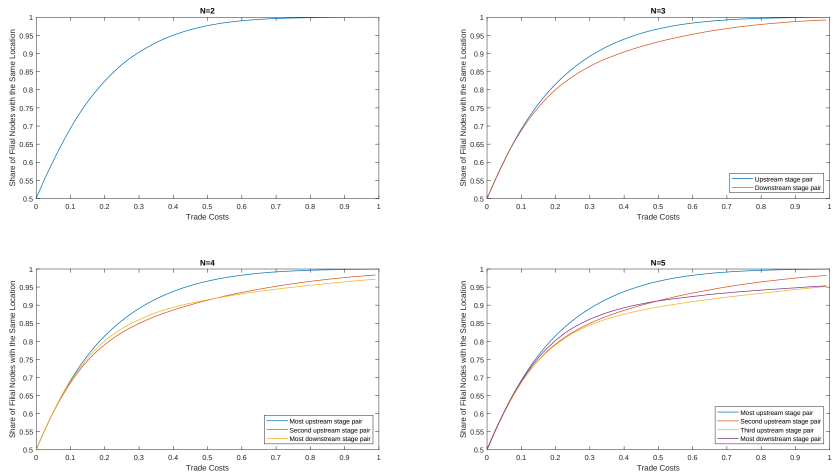

FIGURE 9. CLUSTERING AND TREE LENGTH

C. Number of Countries

Figure 10 illustrates the effect of changes in the number of countries for $N = 5$, $M = 2$, and $S = 100,000$. Note that for $\tau = 0$, when clustering does not occur, the average share of filial nodes with the same location is no longer equal to 0.5 and, due to symmetry, is instead equal to $\frac{1}{K}$. The larger number of countries uniformly decreases the strength of clustering at all stages, which is natural - with the larger number of countries, there are more cost draws for each node, with a chance of low-cost realization in one of the countries, thus there are more incentives to produce in a country different from the adjacent nodes.

A greater number of countries pushes clustering down at all stages, but the effect is stronger for downstream stages because, beyond the direct effect, their children nodes are less likely to cluster, thus reducing incentives to cluster with their parent nodes.

D. Indirect Clustering

Clustering takes place because trade costs arise when adjacent nodes are not colocated. At the same time, due to clustering, nodes that are not directly con-

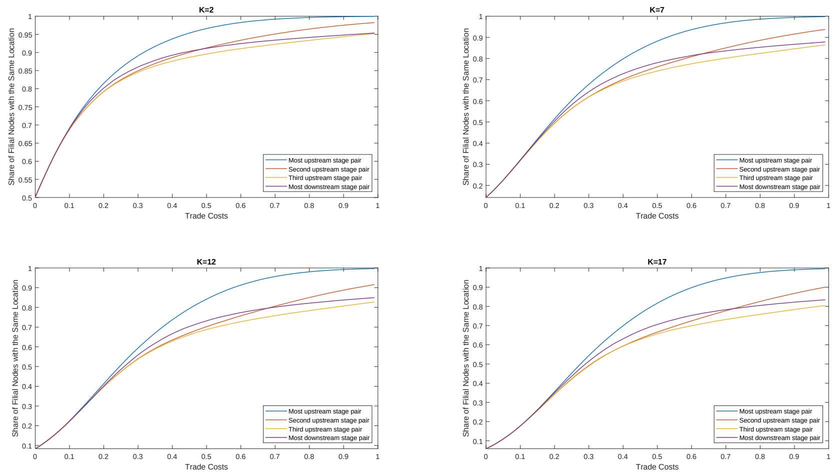

Figure 10. Clustering and the Number of Countries

nected may have a high probability of being colocated. In this section, I consider different cases of indirect clustering. I start with measuring how often nodes at stages i and $i+2$ are co-produced, or grandchildren-grandparents relationship. As one generation in between is skipped, there are one less of such pairs of stages, so for $N = 5$, there are three such pairs.

The results are presented in Figure 11 for $N = 5$, $M = 2$, $K = 2$, and $S = 100,000$. One can see that the grandparents-grandchildren relationship is associated with the smaller, but still significant, share of colocated nodes. Even for great-grandchildren, this share remains significant, indicating that clustering propagates through the network.

Sister nodes are the nodes that have a common parent. The corresponding panel indicates that sister nodes are also often colocated. This occurs because each of the sister nodes tends to cluster with their common parent, so if the filial relationship is strong enough, then parts from the sister nodes will also be coproduced. A high share of coproduced sister nodes makes colocation of their parent node with them cheaper, in turn making it easier for the parent node to colocate with its parent. Stage-specific shares of colocated sister nodes are the reason why changes in trade costs have a heterogeneous effect on clustering strength at different stages of production.

Cousin nodes belong to the same stage of production but do not have a common parent. By construction of the complete trees, these nodes have a common grandparent. Just as grandparent-granchildren node pairs tend to be colocated, so do cousin nodes; however this link is weaker than that between grandchildren and grandparents and sisters. Finally, the niece-aunt relationship represents a pair of an arbitrary node and a sister node of its parent. This relationship exhibits the largest difference between up- and downstream stage pairs because it represents a combination of sister and filial relationships, both of which tend to generate more clustering for upstream stage pairs.

Overall, even nodes that are not directly connected exhibit strong colocation patterns.

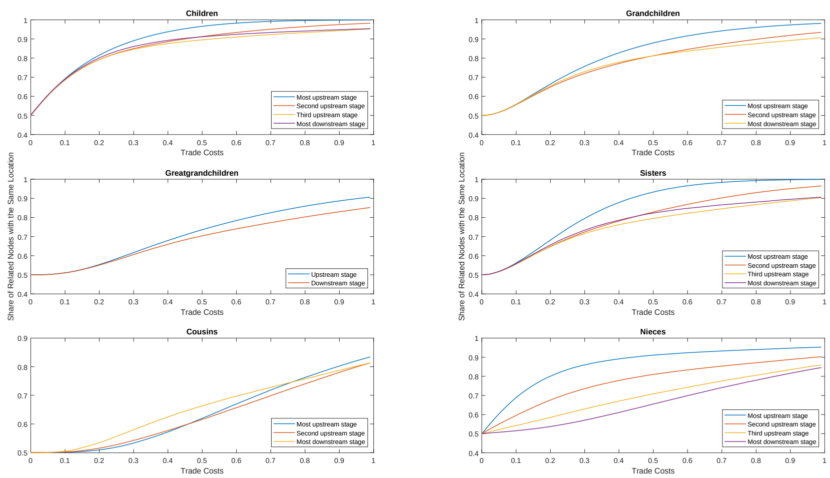

FIGURE 11. DIRECT AND INDIRECT CLUSTERING

E. Reshoring

When a node $\{i, b\}$ is produced in country i for some values of trade costs τ_0 and $\tau_1 > \tau_0$ and produced in country $j \neq i$ for $\tau_1 > \tau > \tau_0$, I say that reshoring occurs. To quantify the amount of offshoring, I change trade costs from 0 to ∞, and if reshoring occurs in a given node at least once, I consider this node to be engaged in reshoring. As a measure of reshoring, I then use the share of nodes at each stage engaged in reshoring.

Figure 12 reflects the effect of tree characteristics on reshoring intensity. A baseline specification is a standard tree with $N = 5$, $M = 2$, $K = 2$, which I simulate $S = 100,000$ times. Panel (a) represents the effect of the length of a tree on the reshoring intensity. As discussed in Section III.B, reshoring does not occur for the case of a spider ($N = 2$), but a further increase in tree length leads to a significant increase in reshoring intensity. One can see that this effect has a diminishing return for a node with the given upstream index as, consistent with Figure 9, an increase in the tree's length has a small effect on the strength of clustering of upstream nodes. The reshoring intensity at the most downstream stage is represented by the beginning of each graph, marked with red crosses; one can see that each additional stage of production significantly increases the reshoring intensity at the most downstream level.

Panel (b) shows the effect of the tree order on reshoring. For $M = 1$, consistent with Section III.A, reshoring is more common for the most up- and downstream nodes and rarely occurs for nonterminal nodes. As a tree's order increases, the intensity of reshoring increases for all stages but, does so disproportionately for downstream stages. This result is not surprising as the tree order has a higher impact on clustering in downstream stages, which means that downstream production decisions are driven less by cost considerations and more by clustering forces, which leads to a higher intensity of reshoring.

In the case of a snake, reshoring is not a common phenomenon, occurring in 5% of the cases for the terminal nodes and barely occuring for nonterminal nodes. However, even for a tree of order 2, reshoring occurs in approximately 10-15% of cases at all stages and can occur in up to 40% of the cases for the most downstream nodes when M is high.

Finally, panel (c) represents the effect of the number of countries on the reshoring intensity. A higher number of countries negatively affects the clustering intensity, but despite this fact, the intensity of reshoring increases. This occurs because, with the higher number of countries, the probability that the combination of draws that leads to reshoring will be drawn becomes greater.

IV. Conclusion

This paper introduces a novel framework, that enables analysis of global sourcing decisions of a firm with a complex production structure. This production structure, which I call a tree, nests two common archetypes of offshoring models – snakes and spiders, and exhibits features of both sequential and simultaneous production. To regularize trees, I introduce a concept of complete trees – they have fractal nature and an important property – an arbitrary tree can be expressed as a complete tree.

The key feature of tree technology is that it generates a clustering mechanism – firms have incentives to colocate intermediate goods produced on adjacent stages with this effect being stronger for higher trade costs. This mechanism generates several interesting outcomes; in particular, it is consistent with the phenomenon of reshoring —- the same part can be produced in one country for high and low values of trade costs and offshored for intermediate values of trade costs.

Offshoring literature on sequential production has mostly focused on models that have closed-form solutions; these models require stronger assumptions on production and transportation costs, which shut down this clustering mechanism. I employ an alternative approach, recently gaining popularity in International

Trade, and offer an optimal control algorithm which can solve the firm's problem for tree technology with arbitrary values of trade costs.

Armed with this algorithm, I perform Monte Carlo simulations and establish the link between properties of complete trees and clustering patterns. In particular, I show that the intensity of clustering is generally lower in more complex production structures with a higher number of stages, number of branches at each stage, and number of countries. This effect, however, is uneven, with clustering being stronger and more robust to changes for upstream nodes. Reshoring, on the other hand, is more prevalent in complex production structures and tends to take place more often in the downstream stages.

REFERENCES

Antràs, Pol, and Alonso De Gortari. 2020. "On the geography of global value chains." *Econometrica*, 88(4): 1553–1598.

Antràs, Pol, and Davin Chor. 2013. "Organizing the global value chain." *Econometrica*, 81(6): 2127–2204.

Antras, Pol, and Elhanan Helpman. 2004. "Global sourcing." *Journal of political Economy*, 112(3): 552–580.

Antras, Pol, Teresa C Fort, and Felix Tintelnot. 2017. "The margins of global sourcing: Theory and evidence from us firms." *American Economic Review*, 107(9): 2514–64.

Arkolakis, Costas, and Fabian Eckert. 2017. "Combinatorial discrete choice." *Available at SSRN 3455353*.

Baldwin, Carliss Young, and Kim B Clark. 2000. *Design rules: The power of modularity.* Vol. 1, MIT press.

Baldwin, Richard, and Anthony J Venables. 2013. "Spiders and snakes: offshoring and agglomeration in the global economy." *Journal of International Economics*, 90(2): 245–254.

Baqaee, David, and Emmanuel Farhi. 2019. "Networks, barriers, and trade." National Bureau of Economic Research.

Costinot, Arnaud, Jonathan Vogel, and Su Wang. 2013. "An elementary theory of global supply chains." *The Review of Economic Studies*, 80(1): 109–144.

Eaton, Jonathan, and Samuel Kortum. 2002. "Technology, geography, and trade." *Econometrica*, 70(5): 1741–1779.

Fally, Thibault, and Russell Hillberry. 2015. *A coasian model of international production chains.* The World Bank.

Feenstra, Robert C, and Gordon H Hanson. 1996. "Globalization, outsourcing, and wage inequality." *The American Economic Review*, 86(2): 240.

Frigant, Vincent, and Yannick Lung. 2002. "Geographical proximity and supplying relationships in modular production." *International Journal of Urban and Regional Research*, 26(4): 742–755.

Galvin, Peter, and Andre Morkel. 2001. "The effect of product modularity on industry structure: the case of the world bicycle industry." *Industry and Innovation*, 8(1): 31.

Grossman, Gene M, and Esteban Rossi-Hansberg. 2008. "Trading Tasks: A Simple Theory of Offshoring." *The American Economic Review*, 98(5): 1978–1997.

Grossman, Gene M, and Esteban Rossi-Hansberg. 2012. "Task trade between similar countries." *Econometrica*, 80(2): 593–629.

Harms, Philipp, Oliver Lorz, and Dieter Urban. 2012. "Offshoring along the production chain." *Canadian Journal of Economics/Revue canadienne d'économique*, 45(1): 93–106.

Jiang, Edwin, and Vladimir Tyazhelnikov. 2020. "Shadow Offshoring and Complementarity in Inputs."

Jia, Panle. 2008. "What happens when Wal-Mart comes to town: An empirical analysis of the discount retailing industry." *Econometrica*, 76(6): 1263–1316.

Johnson, Robert C, and Andreas Moxnes. 2019. "GVCs and trade elasticities with multistage production." National Bureau of Economic Research.

Johnson, Robert C, and Guillermo Noguera. 2012. "Accounting for intermediates: Production sharing and trade in value added." *Journal of International Economics*, 86(2): 224–236.

Kikuchi, Tomoo, Kazuo Nishimura, John Stachurski, and Junnan Zhang. 2019. "Coase Meets Bellman: Dynamic Programming and Production Chains." *arXiv preprint arXiv:1908.10557.*

Lim, Kevin. 2018. "Endogenous production networks and the business cycle." *Work. Pap.*

Oberfield, Ezra. 2018. "A theory of input–output architecture." *Econometrica*, 86(2): 559–589.

Ramondo, Natalia, and Andre Rodrıguez-Clare. 2013. "Trade, Multinational Production, and the Gains from Openness." *Journal of Political Economy*, 121(2).

Razhev, Mark. 2015. "Input-Output Structure and Trade Elasticity." In *Working Paper.*

Simonovska, Ina, and Michael E Waugh. 2014. "The elasticity of trade: Estimates and evidence." *Journal of international Economics*, 92(1): 34–50.

Sirkin, Harold L, Michael Zinser, Douglas Hohner, et al. 2011. *Made in America, again: Why manufacturing will return to the US.* Boston Consulting Group Boston.

Sraffa, Piero. 1975. *Production of commodities by means of commodities: Prelude to a critique of economic theory.* CUP Archive.

Tintelnot, Felix. 2017. "Global production with export platforms." *The Quarterly Journal of Economics*, 132(1): 157–209.

Tintelnot, Felix, Ayumu Ken Kikkawa, Magne Mogstad, and Emmanuel Dhyne. 2018. "Trade and domestic production networks." National Bureau of Economic Research.

Wu, Xiaole, and Fuqiang Zhang. 2011. "Efficient supplier or responsive supplier? an analysis of sourcing strategies under competition."

Yi, Kei-Mu. 2010. "Can multistage production explain the home bias in trade?" *The American Economic Review*, 100(1): 364–393.

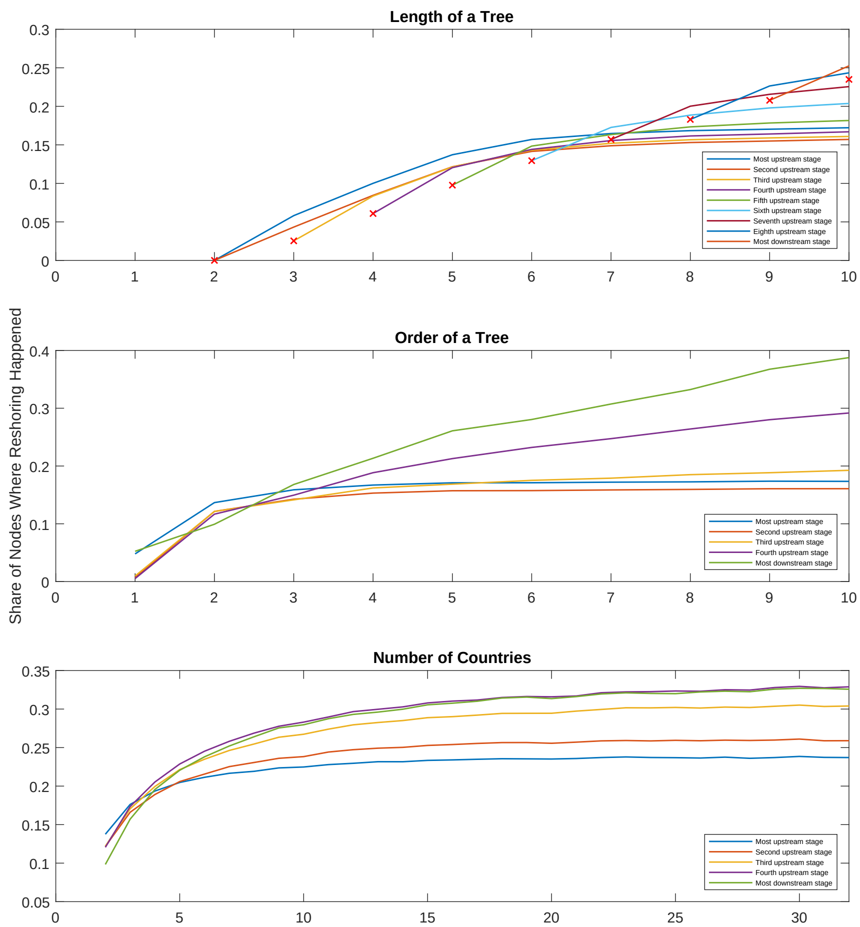

Figure 12. Reshoring