Global Value Chains and Increasing Returns

Alonso de Gortari

alonso.degortari@gmail.com

January 2020

Abstract

This paper develops a simple, tractable, exact numerical solution method for multi-stage production models in the presence of internal economies of scale; a challenge facing the literature since these models are characterized by NP-hard optimization problems. The key insight is that firms increasingly trade-off fixed costs against marginal costs of production as output increases. The algorithm exploits this monotonicity property through iterative zero-one integer programs to trace a firm's optimal supply chain. The algorithm is fast, even in high dimensions. The model provides suggestive evidence that increasing returns amplify the effects of trade policy uncertainty and increase the costs of supply chain disruption.

JEL codes: C6, F1, F6.

1 Introduction

About 40% of the American companies surveyed by the American Chamber of Commerce were considering or had already relocated part of their production outside of China during the 2019 U.S.-China trade war. Indeed, top executives recently told the Wall Street Journal that "Once you move out, you don't go back," that "If we were to try to do a factory in the U.S., it would be enormously expensive," and that "It took us 20 years to build up the supply chain in China". These concerns are widespread: Blackrock's head, the world's largest asset manager, recently declared that "We're hearing from CEOs that more and more supply chains are moving out of China right now."[1]

But how do firms organize their supply chains? How are these decisions shaped by the potential impact of supply chain disruption? Conventional wisdom states that firms offshore to China because of its low wages. If that is the whole story, should firms not move back immediately once U.S.-China trade relations normalize? Why do firms care about the overhead associated with setting up a factory? Why does it take so long to build up a supply chain? Why does future trade policy uncertainty appear to have such a significant impact?

This paper incorporates internal economies of scale into an otherwise standard global value chain (GVC) model to argue that fixed costs of production amplify the effects of trade policy uncertainty and increase the costs of supply chain disruption. The mechanism is a simple convexity argument: in the absence of fixed costs, firms face flat average cost curves. With increasing returns, however, average cost curves are decreasing with quantities. Firms organize their supply chains by trading-off fixed costs against marginal costs of production and set up different supply chains depending on the quantity being produced. Since trade policy uncertainty typically affects marginal costs (ex: ad valorem tariffs), this has a muted effect in the constant returns to scale case. Firms only reorganize their supply chain following substantial trade cost changes since this requires that a new average cost curve drop below the initially minimal cost curve. With increasing returns, though, the impact is amplified since average cost curves cross each other and even small changes in trade costs bring about big supply chain reorganizations.

Increasing returns thus imply that firms invest more cautiously in their supply chains when future trade costs might shift. Further, increasing returns imply that firms have more reason to fear the costs of supply chain disruption. Indeed, this is consistent with the 33% of U.S. firms producing in China that delayed or canceled their investment decisions as a consequence of the trade war. The paper provides numerical evidence consistent with this intuition.

Internal economies of scale became a central ingredient of the trade literature following the seminal work of Krugman (1979) and Krugman (1980), and more recently with Melitz (2003). Generalizing these models to a world with GVCs in the form of multi-stage production is highly challenging since it involves solving NP-hard combinatorial optimization problems.[2] In this set-

[1] Wall Street Journal (2019), American Chamber of Commerce in People's Republic of China (2019), CNBC (2019).

[2] Multi-stage production combinatorial optimization problems are similar to traveling salesman type problems that require ordering a fixed number of locations optimally. These are different from models in the spirit of Jia (2008) (or Arkolakis and Eckert 2017) where making additional choices generates complementarities with previous locations.

ting, firms organize an N-stage supply chain by assigning each stage n to one of J countries; thus leading to an assignment problem of size J^N. Antràs and de Gortari (2017) showed that dynamic programming techniques can be used to solve the constant returns to scale case, but their procedure does not work with scale economies. A few studies such as Tintelnot (2017) and Antràs et al. (2017) have managed to incorporate some form of fixed costs in GVC models, but these typically focus on a specific stage of importing or exporting instead of fully-fledged supply chain models.

This paper develops a simple, tractable, exact numerical solution method for multi-stage production models in the presence of internal economies of scale. The algorithm builds on the key insight that as quantity increases, firms are more likely to seek low marginal costs even if this comes at the expense of higher fixed costs. This monotonicity property is leveraged through an iterative zero-one integer program to trace a firm's full set of optimal supply chains. The algorithm is very efficient even in high dimensions and takes relatively little time to converge even in cases with huge solution spaces of the order of 10^{26} (when $J = N = 20$). While the algorithm very much requires a unitary elasticity of substitution across stages of production, this is a first step towards modeling GVCs with increasing returns in more general settings. That said, most multi-stage GVC models without economies of scale also use Cobb-Douglas production (see Yi 2003, Johnson and Moxnes 2019) even though dynamic programming solution methods are available for more general production functions (see Antràs and de Gortari 2017).

The paper is organized as follows. Section 2 describes the model setup, numerical challenges, and relation to the existing GVC literature in economics and operations research. Section 3 develops the numerical solution method, discusses time complexity, and provides benchmark performance statistics. Section 4 illustrates the model's usefulness and solution method by studying the impact of trade policy uncertainty and the cost of supply chain disruption in the presence of increasing returns. Section 5 concludes.

2 Model Setup

There are J countries (contained in the set $\mathcal{J}$) and producing a final good requires N stages of sequential production. Production at the most upstream stage $n = 1$ is one-to-one with equipped labor, while production q^n at each subsequent stage n requires combining equipped labor L^n with the upstream input $\tilde{q}^{n-1}$

$$q^n = \left(\frac{L^n}{\alpha^n}\right)^{\alpha^n} \left(\frac{\tilde{q}^{n-1}}{1-\alpha^n}\right)^{1-\alpha^n}. \tag{1}$$

The key parameter is $\alpha^n \in (0,1)$ which captures labor's expenditure share at stage n in terms of stage n gross output. In terms of assumptions, stage $n = 1$ is a special case with $\alpha^1 = 1$ and, as will be discussed below, the Cobb-Douglas structure is crucial. With frictionless trade, $\tilde{q}^{n-1}$ equals the inputs required to produce q^n units of stage n output and so output at upstream stage $n-1$ equals stage n's input requirements: $q^{n-1} = \tilde{q}^{n-1}$. With iceberg trade costs $\tau > 1$, however, $q^{n-1} = \tilde{q}^{n-1}\tau$ since the upstream stage produces extra output that 'melts' when shipped.

The firm's problem is to decide where to locate each stage of production n. The firm trades-off three margins when making this decision: First, effective units of labor (a mix of labor, capital, local materials or inputs) command a unit price c_i^n in each country $i \in \mathcal{J}$ and may, in principle, vary across stages of production n. Second, producing two subsequent stages $n-1$ and n in two different countries i and j requires shipping $\tau_{ij} > 1$ units for one input to arrive at j. Third, setting up a production facility in country i to produce stage n requires paying a fixed cost F_i^n regardless of how much output is produced there. The firm would thus prefer to produce in locations with low labor costs, in locations that are close to each other, and in locations in which setting up plants is cheap. The rest of the paper shows how to find the optimal supply chain.

Let $\ell(n) \in \mathcal{J}$ be the location producing the nth stage and let $\boldsymbol{\ell} \in \mathcal{J}^N$ be a supply chain vector of locations summarizing where each stage is located: $\boldsymbol{\ell} = \{\ell(1), \ldots, \ell(N)\}$. The firm's goal is to minimize its average cost of production by choosing the optimal $\boldsymbol{\ell}$ among all possible supply chains.[3] Since average costs vary depending on the quantity being produced, the firm's problem can be summarized as computing the lower envelope of the cost functions across all possible supply chains for each level of final output. This can be done in two steps.

The first step is to obtain the minimum cost of production when producing q^F final goods through a given supply chain $\boldsymbol{\ell}$. This equals

$$C\left(q^F, \boldsymbol{\ell}\right) = q^F \times \prod_{n=1}^{N} \left(c_{\ell(n)}^n\right)^{\alpha^n \beta^n} \times \prod_{n=1}^{N-1} \left(\tau_{\ell(n)\ell(n+1)}\right)^{\beta^n} + \sum_{n=1}^{N} F_{\ell(n)}^n, \tag{2}$$

with $\beta^n = \prod_{m=n+1}^{N}(1-\alpha^m)$ capturing the ratio of gross output shipped at stage n relative to final good, or stage-N, gross output (since value is added at each stage note $\beta^1 \leq \beta^2 \leq \cdots \leq \beta^N = 1$). The cost function depends on two terms: variable costs and fixed costs. First, total variable costs depend on the amount of final goods being produced q^F (say iPhones) and the unit cost of each final good (the marginal cost of each iPhone). The latter depends on i) the costs of effective units of labor $c_{\ell(n)}^n$ at each stage weighted by $\alpha^n \beta^n$, their contribution to final good value-added, ii) the trade costs associated with shipping output across any two stages $\tau_{\ell(n)\ell(n+1)}$ weighted by β^n, the gross value shipped at stage n.[4] Second, total fixed costs equal the costs of setting up plants (any overhead incurred by Apple such as capital structures, logistics, rent, etc) along the supply chain.[5]

The second step is to find the supply chain minimizing total costs

$$\boldsymbol{\ell}\left(q^F\right) = \underset{\boldsymbol{\ell} \in \mathcal{J}^N}{\arg\min} \quad C\left(q^F, \boldsymbol{\ell}\right). \tag{3}$$

[3] Minimizing costs up to the assembly stage is the firm's goal in the absence of final good trade costs; an useful characterization since it implies a common final good demand function regardless of where assembly is located. Alternatively, shipping the stage $\ell(N)$ final good to consumers in j could require paying an additional $\tau_{\ell(N)j}$ trade cost in which case firms minimize the cost of serving consumers in market j. All subsequent results apply to this alternative case.

[4] Total variable costs can be found by by fixing a supply chain $\boldsymbol{\ell}$ and substituting recursively the upstream production function in (1) from stage $n-1$ into that at stage n across all stages. Then one can find the unit variable costs of supply chain $\boldsymbol{\ell}$ by solving a standard cost minimization problem. See Antràs and de Gortari (2017) for details.

[5] An extension with multi-stage plants is discussed in Section 5.

While finding the cost function in (2) for a given supply chain requires the standard toolkit, finding the optimal supply chain in (3) is extremely challenging because it requires solving a non-differentiable and non-linear assignment problem. The lack of differentiability prevents us from using the standard calculus tools from economics while the lack of linearity prevents us from using the standard linear programming tools from operations research. Moreover, the problem is hard to solve by brute force since it has J^N possible solutions and thus increases exponentially with N.

The lack of progress in solving problems in the spirit of (3) remains one of the main roadblocks impeding the evolution of the GVC literature into a world with increasing returns. In contrast, the traditional international trade literature has long incorporated these forces since the seminal work of Krugman (1979) and Krugman (1980) and, more recently, with Melitz (2003). Indeed, in the absence of GVCs or multi-stage production, the above model with $N = 1$ nests standard trade models with scale economies and (3) is easily solved by ranking countries according to c_i/F_i.

This paper provides a highly tractable−and easily implementable−numerical solution method for multi-stage production models in the spirit of (3). Namely, models incorporating both i) substitution of inputs across stages of production (i.e. the Cobb-Douglas functional form), and ii) fixed costs of production. In order to build up the intuition for the solution method, I now describe the features and limitations of the current frontier in terms of numerically solving GVC models.

GVC Models in Economics

Modeling GVCs through multi-stage production with constant returns to scale is the special case of (3) with no fixed costs

$$\boldsymbol{\ell}\left(q^F\right) = \underset{\boldsymbol{\ell}\in\mathcal{J}^N}{\arg\min}\ q^F \times \prod_{n=1}^{N}\left(c^n_{\ell(n)}\right)^{\alpha^n\beta^n} \times \prod_{n=1}^{N-1}\left(\tau_{\ell(n)\ell(n+1)}\right)^{\beta^n}. \tag{4}$$

A key property of these models is that marginal costs always equal average costs. Hence, there is a single global solution to this problem and the optimal supply chain $\boldsymbol{\ell}\left(q^F\right) = \boldsymbol{\ell}$ is the same regardless of how many final goods q^F are produced.

Modeling GVCs in the spirit of (4) was born with the seminal work of Yi (2003). While the absence of fixed costs greatly simplifies the problem relative to (3), more than a decade passed before the GVC literature figured out how to solve (4) tractably in high dimensions. The complexity is driven by the iceberg trade costs which generate interdependencies across production stages−much like traveling salesman problems. To see this, focus on the trivial case without trade costs. Finding the optimal chain is easy since the optimal location producing stage n is such that

$$\ell(n) \in \boldsymbol{\ell} \Leftrightarrow \ell(n) = \underset{\ell\in\mathcal{J}}{\arg\min}\ c^n_{\ell(n)}.$$

With no interdependencies across stages of production, each stage is optimally located in that location offering the lowest unit cost of effective units of labor. The optimal GVC is found by minimizing stage-by-stage and solving N minimization problems of size J each. This approach is akin to multi-sourcing production models in the spirit of Antràs et al. (2017).

Bilateral iceberg trade costs, in contrast, do generate interdependencies across stages of production: a country i featuring minimal c_i^n at stage n might not be included in the optimal supply chain $\boldsymbol{\ell}$ if it is far away from the upstream input supplier $\ell(n-1)$ and the downstream producer $\ell(n+1)$. Likewise, a country with high c_i^n might form part of the optimal supply chain if it is closely located to upstream and downstream locations with low unit costs. In other words, the firm can only organize its supply chain by comparing the full supply chain sequences in the $\mathcal{J}^N$ space. Solution by brute force is costly because it requires solving a single global minimization problem with a space of solutions increasing exponentially with N. Indeed, the seminal work of Yi (2003) focused on the simplest multi-stage model with $N = 2$ and used brute force to find optimal supply chains. Many subsequent GVC papers adopted similar solution techniques.

Recently, Antràs and de Gortari (2017) showed that (4) can easily be solved, even in high dimensions with large N, when using the tools of dynamic programming. Specifically, one can begin by focusing on a decentralized solution in which producers in all countries $j \in \mathcal{J}$ at stage 2 choose their optimal input supplier from stage 1; call it $\ell_j(1)$. Then, producers in all countries $j \in \mathcal{J}$ at stage 3 choose their optimal supply chain by comparing the input costs of stage 2 suppliers conditional on their stage 1 sourcing decision (i.e. the cost of supplier i at stage 2 is given by supply chain $\{\ell_i(1), i\}$). This approach can be repeated until stage N is reached and the optimal supply chain is obtained by invoking Bellman's principle of optimality. More generally, Antràs and de Gortari (2017) showed this numerical approach can be used with any constant-returns-to-scale production function (not only Cobb-Douglas) and that it is highly efficient since it turns the global optimization problem of exponential size J^N to one requiring only $(N-1) \times J^2$ computations. Similar recursive techniques have been used by Johnson and Moxnes (2019) and Tyazhelnikov (2019).

Why do scale economies as in (3) vastly complicate the firm's optimal supply chain decision? There are two main challenges relative to the case with no fixed costs. The first challenge is that the GVC literature has yet to find a tractable numerical method for solving (3) when conditioning on a given amount of final goods q^F. That is, even if Apple knew it wanted to produce q^F iPhones, it is not clear how to easily find the supply chain minimizing total costs. As discussed previously, since (3) is both non-differentiable and non-linear, standard calculus and linear programming tools do not apply. Further, dynamic programming tools do not apply either since the firm cannot simply compare average costs up to a given stage n and solve recursively. The reason is that production features substitution across stages, making it impossible to know how many parts q^n are required at stage n in order to produce q^F final goods without specifying the full supply chain.[6] In other words, a firm can only compute average costs and make the marginal cost vs fixed cost trade-off if

[6]More precisely, the relationship between prices is $p^n_{\ell(n)} = \left(c^n_{\ell(n)}\right)^{\alpha^n}\left(p^{n-1}_{\ell(n-1)}\tau_{\ell(n-1)\ell(n)}\right)^{1-\alpha^n}$ and the relationship between expenditures is $p^n_{\ell(n)}q^n_{\ell(n)} = \beta^n \times p^N_{\ell(N)}q^F$. Rearranging, this implies that if a firm wants to produce q^F units of final goods through supply chain $\boldsymbol{\ell}$ then it must produce the following amount of units at each stage

$$q^n_{\ell(n)} = \beta^n \frac{\prod_{m=1}^N \left(c^m_{\ell(m)}\right)^{\alpha^m\beta^m} \times \prod_{m=1}^{N-1}\left(\tau_{\ell(m)\ell(m+1)}\right)^{\beta^m}}{\prod_{m=1}^n \left(c^m_{\ell(m)}\right)^{\alpha^m\frac{\beta^m}{\beta^n}} \times \prod_{m=1}^{n-1}\left(\tau_{\ell(m)\ell(m+1)}\right)^{\frac{\beta^m}{\beta^n}}} \times q^F.$$

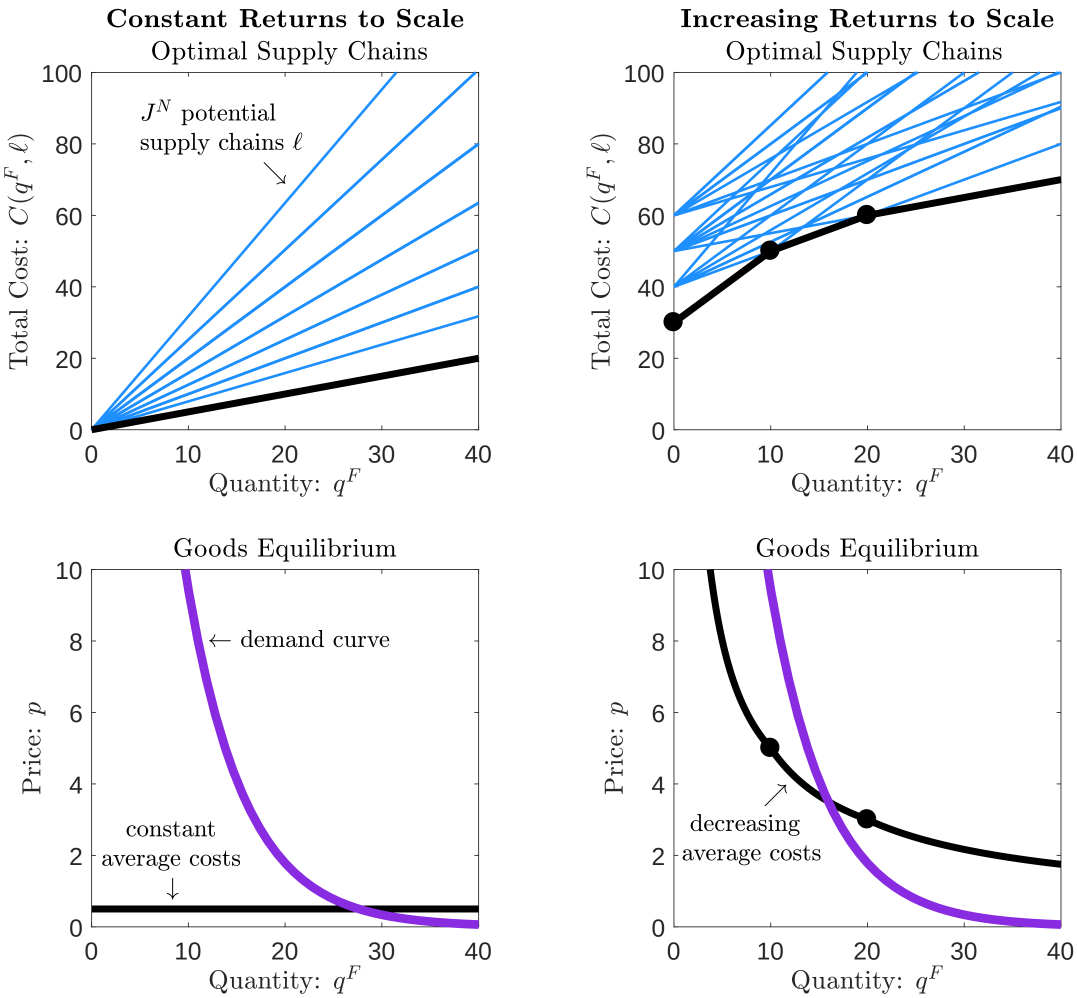

Figure 1: Supply Chain Optimization with Constant and Increasing Returns to Scale: Example with $J = N = 3$. The upper two panels plot the universe of possible cost functions with the thick lower envelope representing the minimum cost (or optimal supply chain) for each quantity. The lower two panels plot the minimum average cost function together with a representative demand curve.

it focuses on a fully specified sequence of locations $\boldsymbol{\ell}$. The firm can thus solve (3) by brute force, but this requires searching across a space of size J^N.

To better illustrate the intuition for this challenge, observe Figure 1 which simulates a simple $J = N = 3$ example. The northwest panel plots the J^N cost curves without fixed costs: each curve is a ray from the origin since marginal costs equal average costs and determine the curve's slope. The northeast panel plots the J^N cost curves with fixed costs: each cost curve starts at $\sum_n F^n_{\ell(n)}$ when $q^F = 0$, the slope is the marginal cost, and the average cost is decreasing because of increasing returns.[7] The firm's goal is to obtain the lower envelope of the total cost function across

Since there is substitution across stages, $q^n_{\ell(n)}$ depends on the costs incurred in all upstream and downstream locations. In the absence of fixed costs, this is a non-issue since there is a single optimal chain and tracking quantities is irrelevant.

[7]The left panel appears to have less than J^N cost curves since several chains share the same marginal costs in this example. With fixed costs, all curves appear on the right panel (with those sharing marginal costs parallel to each other).

all possible supply chains. This is easily done with dynamic programming in the absence of fixed costs, and the northwest panel shows that the optimal supply chain is that with minimal marginal costs (or a minimal slope). With fixed costs, however, this is challenging because the trade-off between marginal and fixed costs depends on q^F. In the northeast panel, when $q^F \in (10, 20)$ the optimal supply chain is neither the supply chain with minimal marginal costs nor the one with minimal fixed costs. Indeed, finding the optimal supply chain for each level of q^F requires solving (3)−which we do not yet know how to do.

The second challenge is that even if we solved (3) by brute force, that only gives us the minimal cost for a given q^F. Since the marginal vs fixed cost trade-off changes with q^F, finding an equilibrium in the goods market requires performing the brute force computation for many different levels of q^F. And, unfortunately, even then one cannot be certain that the cost function's lower envelope has been fully traced out.

To see this better again observe Figure 1 and, for the pure sake of exposition, assume that i) the firm faces an exogenous demand function and ii) the firm decides to sell its output at average cost. With no fixed costs, there is a single optimal supply chain. The southwest panel shows that the goods equilibrium is immediately found by finding the level of demand consistent with the firm's average (or marginal) cost. With fixed costs, however, we cannot know ex-ante what the equilibrium q^F will be since the average cost function is both i) decreasing and ii) features kinks when the optimal supply chain changes. In other words, one could begin by solving (3) by brute force with some arbitrary q^F. This would deliver the lowest average cost at q^F, but ex-ante it is not clear if this average cost curve will cross the demand function at a point in which that supply chain is still the optimal one. Indeed, the southeast panel shows that it will not if the initial $q^F < 10$ or $q^F > 20$. The optimal supply chain delivering the market clearing q^F requires solving (3) by brute force many times until the lowest average cost curve crossing the demand function is found.

Beyond average cost pricing, the intuition is exactly the same for firms in imperfectly competitive markets: a monopolist choosing quantity and price jointly needs to both know i) its demand function and ii) its cost function. But computing the cost function requires finding the lower envelope in the northeast panel which requires solving (3) by brute force many times.

GVC Models in Operations Research

Supply chain management is one of the classical topics in operations research. There are two main reasons why their approach to modeling supply chains is different from an economist's approach. The first is that production is modeled closer to an engineer's view of the world and thus the blueprint required to transform upstream inputs into downstream inputs is fixed in terms of physical quantities. In other words, the production function typically does not feature substitution across stages as in (1) but instead takes the Leontief form

$$q^n = \min \left\{ \frac{L^n}{\alpha^n}, \frac{q^{n-1}}{1-\alpha^n} \right\}.$$

Ignoring trade costs for now, this delivers the following cost function

$$C\left(q^F, \boldsymbol{\ell}\right) = q^F \times \sum_{n=1}^{N} \alpha^n \beta^n c^n_{\ell(n)} + \sum_{n=1}^{N} F^n_{\ell(n)}.$$

While this is still non-differentiable and the space of solutions is still of size J^N, this cost function is a standard linear assignment problem that can readily be solved using integer linear programming.

The second way operations research studies supply chains differently is in its handling of trade costs. Because the production function is Leontief, the amount of inputs required to produce any given level of final goods q^F is independent of the supply chain (which is not the case with substitution, see footnote 6). In other words, output at stage n is $q^n_{\ell(n)} = \beta^n \times q^F$ regardless of where the other stages of $\boldsymbol{\ell}$ are located. The handling of trade costs is consistent with this view of the world: instead of using iceberg trade costs, the costs associated with shipping inputs across any two countries $\ell(n)$ and $\ell(n+1)$ at stage n should also be independent of what happens throughout other parts of the supply chain. This is typically incorporated through a set of additive trade costs t^n_{ij} which are paid per unit of the final good. The cost function is thus

$$C\left(q^F, \boldsymbol{\ell}\right) = q^F \times \sum_{n=1}^{N} \left(\alpha^n \beta^n c^n_{\ell(n)} + t^n_{\ell(n)\ell(n+1)}\right) + \sum_{n=1}^{N} F^n_{\ell(n)},$$

which is also a simple assignment problem that can be solved with integer linear programming.[8]

3 Numerical Solution Method

I propose a simple, yet powerful, algorithm for finding optimal supply chains in the context of i) substitution across stages of production (with elasticity 1), ii) iceberg trade costs, and iii) internal economies of scale, as in (3). I first describe the algorithm, then show a graphical proof, and finally provide mathematical and numerical details.

The algorithm is based on the key insight that the trade-off between marginal and fixed costs is monotonic as the final output level q^F falls. Suppose we knew the optimal supply chain at an arbitrary q^F_0, how do we trace the cost curve's lower envelope? The algorithm's key insight is that as $q^F < q^F_0$ decreases, a necessary condition for the firm to reorganize its supply chain is that fixed costs fall and marginal costs rise with the new chain. To see why, imagine this did not occur and the firm reorganized its supply chain at q^F to one with higher fixed costs. This can only be optimal if marginal costs fall or else total costs are higher than at the original chain. But if this occurs at $q^F < q^F_0$, then total costs must also be lower at the original q^F_0 on this new chain thus contradicting

[8] Incorporating iceberg trade costs with Leontief production, however, is quite challenging. The cost function equals

$$C\left(q^F, \boldsymbol{\ell}\right) = q^F \times \sum_{n=1}^{N} \left(\alpha^n \beta^n c^n_{\ell(n)} \prod_{m=n}^{N-1} \tau_{\ell(m)\ell(m+1)}\right) + \sum_{n=1}^{N} F^n_{\ell(n)}.$$

the optimality of the original chain. Hence fixed costs must fall. In addition, marginal costs must increase or else the firm would have already been on this supply chain.

The algorithm thus builds on this monotonicity property as follows: when $q^F \to \infty$, fixed costs are irrelevant and the optimal supply chain is that with minimal marginal costs. As q^F falls, the firm reorganizes its supply chain only if fixed costs fall and, when fixed costs do fall, the firm moves to the curve with lowest marginal costs out of those with lower fixed costs. The left three panels of Figure 2 describe this algorithm graphically using Figure 1's example. The top panel shows the algorithm starts by finding the curve with minimum marginal costs; the firm produces here when $q^F \to \infty$. Once this curve is found, all curves with higher fixed costs can be discarded (since they also have higher marginal costs by construction). The middle panel shows the firm then proceeds to find the curve with lowest marginal costs out of all those with lower fixed costs than the previous curve. After this curve is found, all additional curves with higher fixed costs can be discarded. The algorithm iterates repeatedly until, as the bottom panel shows, all cost curves have been discarded and this delivers the cost function's lower envelope.

The monotonicity property is only a necessary condition because sometimes cost curves that are initially on the lower envelope are then subsequently dominated by a curve with even lower fixed costs (but higher marginal costs). In practice, the algorithm takes care of this by using the thresholds at which the cost curves intersect to determine which curves are actually optimal. Below I will show that this 'inefficiency' is irrelevant from a time complexity perspective in practice.

Mathematically, the algorithm exploits three properties of our supply chain model. First, while production features substitution, the Cobb-Douglas structure makes total marginal costs log-linear. Second, total fixed cost are linear. By themselves, these two properties are not useful for solving (3) since the total cost function is neither log-linear nor linear. But combining these two properties together with the monotonicity property provides a fully linear algorithm obtained by splitting the marginal and fixed costs trade-off into separate parts of the optimization problem.

The algorithm exploits the monotonicity property through iterations minimizing log marginal costs subject to searching across all supply chains with lower fixed costs than the previous iterations. This can be represented with the tools of zero-one integer programming:

$$\begin{aligned}
\min_{\zeta_{ij}^n, \zeta_i^N \in \{0,1\}} \quad & \sum_{n=1}^{N-1} \sum_{i,j} \zeta_{ij}^n \left(\alpha^n \beta^n \ln c_i^n + \beta^n \ln \tau_{ij}\right) + \sum_i \zeta_i^N \alpha^N \beta^N \ln c_i^N, \\
\text{[I]}: \quad & \sum_i \zeta_{ij}^n = \sum_i \zeta_{ji}^{n+1}, \ \forall j \in \mathcal{J}, n = 1, \ldots, N-2, \\
\text{[II]}: \quad & \sum_i \zeta_{ij}^{N-1} = \zeta_j^N, \ \forall j \in \mathcal{J}, \\
\text{[III]}: \quad & \sum_i \zeta_i^N = 1, \\
\text{[IV]}: \quad & \sum_{n=1}^{N-1} \sum_{i,j} \zeta_{ij}^n F_i^n + \sum_i \zeta_i^N F_i^N < FC_{max}.
\end{aligned} \tag{5}$$

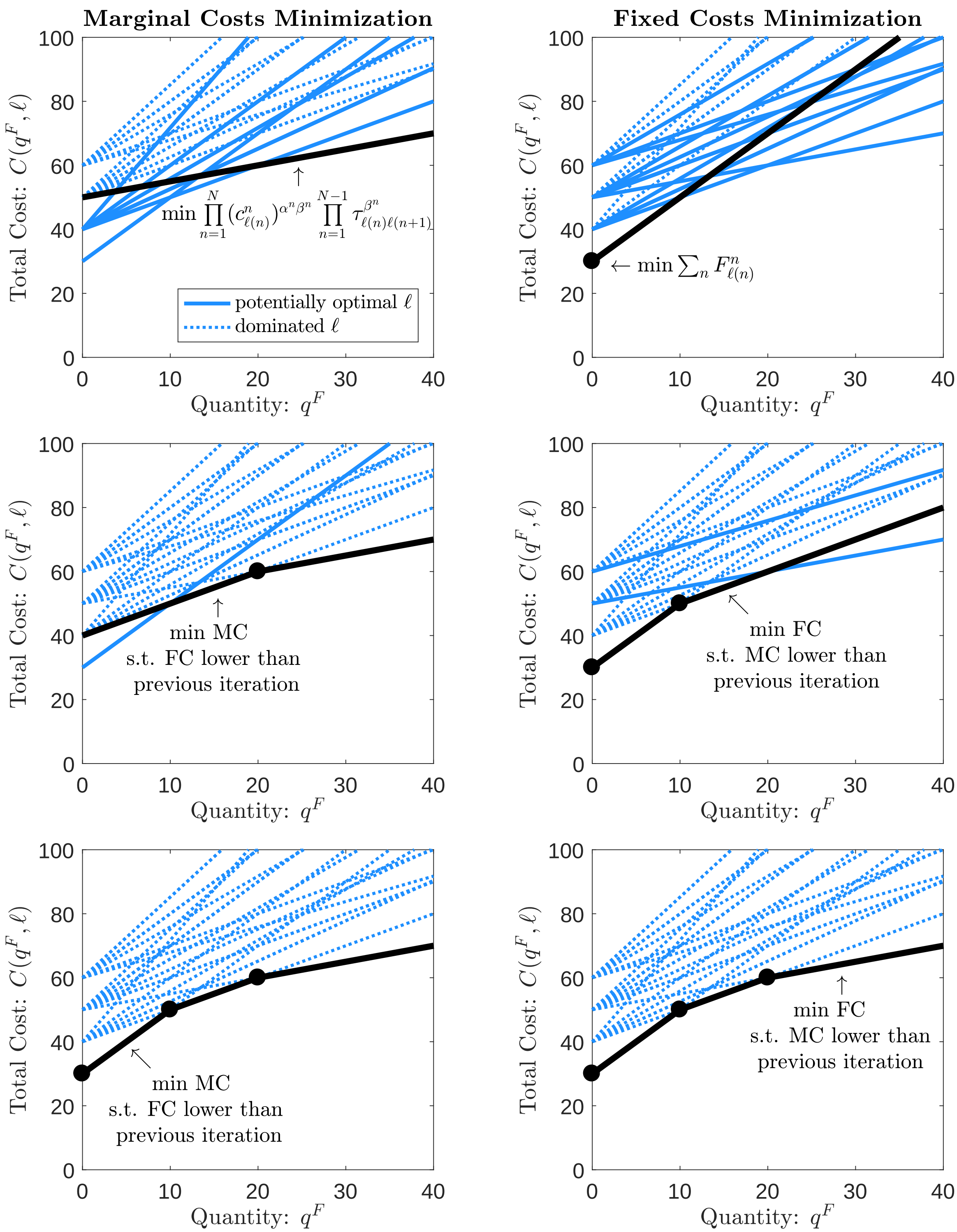

Figure 2: An Iterative Algorithm for Computing Optimal Supply Chains: Illustration of the iterative marginal cost minimization algorithm (5) and the iterative fixed cost minimization algorithm (6) using Figure 1's example. Both arrive to the same solution.

The endogenous variables are $\zeta_{ij}^n, \zeta_i^N \in \{0,1\}$, with $\zeta_{ij}^n = 1$ if and only if $\ell(n) = i$ and $\ell(n+1) = j$ and $\zeta_i^N = 1$ if and only if $\ell(N) = i$. The objective function minimizes marginal costs, constraints [I] and [II] ensure that if j receives the stage n input then j produces the stage $n+1$ input, [III] ensures that one and only one country assembles the final good, and [IV] restricts the solution space to supply chains with lower fixed costs than FC_{max} (with FC_{max} entering as a parameter).

The algorithm thus iterates as follows: start by solving (5) with $FC_{max} = \infty$ to obtain the cost curve for $q^F \to \infty$ and update FC_{max} to the fixed cost of this curve, then iterate by re-solving (5) with the new FC_{max} and update after each iteration. The algorithm is done when there are no more feasible solutions. The segment of q^F over which each cost is relevant is found by intersecting each pair of subsequent optimal cost curves. As mentioned above, this also helps discard a few cost curves that are not actually on the lower envelope. The top panel in Figure 3 shows the algorithm's output in practice in a $J = N = 5$ example where the problem's high dimensionality is evident.

A nice implication of the monotonicity property is that it can also be applied in reverse order. That is, the algorithm can alternatively be based on the insight that when $q^F \to 0$ the firm only cares about minimizing fixed costs. As q^F increases, the firm will reorganize its supply chains to those with higher fixed costs but only if they deliver lower marginal costs. This approach is described graphically in the right three panels of Figure 2. Importantly, it reaches the same solution as the previous algorithm. Mathematically, this version of the iterative algorithm is defined as:

$$
\begin{aligned}
\min_{\zeta_{ij}^n, \zeta_i^N \in \{0,1\}} \quad & \sum_{n=1}^{N-1} \sum_{i,j} \zeta_{ij}^n F_i^n + \sum_i \zeta_i^N F_i^N, \\
[\text{I}]: \ & \sum_i \zeta_{ij}^n = \sum_i \zeta_{ji}^{n+1}, \ \forall j \in \mathcal{J}, n = 1, \dots, N-2, \\
[\text{II}]: \ & \sum_i \zeta_{ij}^{N-1} = \zeta_j^N, \ \forall j \in \mathcal{J}, \\
[\text{III}]: \ & \sum_i \zeta_i^N = 1, \\
[\text{IV}]: \ & \sum_{n=1}^{N-1} \sum_{i,j} \zeta_{ij}^n \left(\alpha^n \beta^n \ln c_i^n + \beta^n \ln \tau_{ij}\right) + \sum_i \zeta_i^N \alpha^N \beta^N \ln c_i^N < MC_{max}.
\end{aligned}
\tag{6}
$$

The setup is similar to (5) with the difference being the initial iteration which starts with $MC_{max} = \infty$ and then iterates while updating MC_{max} until there are no more feasible solutions.[9]

These algorithms are highly tractable even in high dimensions since they require iterating a series of zero-one integer programs for which many lightning-speed algorithms exist. Furthermore, while solution by brute force is a heuristic since one has to decide along which set of quantities q^F to minimize costs, this algorithm is guaranteed to provide a full description of the minimum cost function. Because it depends on the monotonicity property, the algorithm necessarily traces out

[9]While both algorithms yield the same solution, one can be faster than the other in practice depending on problem's parameters. For example, (6) tends to be faster than (5) when the variance in fixed costs is low relative to the variance in marginal costs since the initial iterations discard a larger amount of cost curves.

J	N	J^N	mean # of optimal ℓ	MC minimization time			FC minimization time		
				mean	min	max	mean	min	max
2	2	4	1.9	0.002	0.001	0.027	0.002	0.001	0.013
4	4	256	4.3	0.013	0.002	0.055	0.014	0.002	0.082
6	6	4.7×10^4	7.4	0.127	0.017	0.449	0.111	0.008	0.439
8	8	1.7×10^7	10.6	0.588	0.103	2.038	0.526	0.087	1.740
10	10	1.0×10^{10}	13.8	1.801	0.392	5.295	1.545	0.342	4.338
15	15	$4.4{\times}10^{17}$	21.8	16.66	3.903	60.67	10.58	2.272	32.22
20	20	1.0×10^{26}	29.6	90.58	33.08	355.3	31.70	6.888	101.2
25	2	400	4.5	0.013	0.002	0.059	0.020	0.002	0.076
25	4	3.9×10^5	8.0	0.565	0.088	2.043	0.793	0.181	2.463
25	6	2.4×10^8	11.3	2.470	0.686	7.153	2.458	0.676	6.786
50	2	2500	5.4	0.057	0.008	0.199	0.139	0.013	0.430
50	4	6.3×10^6	9.2	2.303	0.449	6.622	3.297	0.967	8.471
50	6	1.6×10^{10}	13.1	12.07	3.601	29.94	11.54	3.740	28.08
100	2	1.0×10^4	6.0	0.214	0.030	0.736	0.645	0.063	1.983
100	4	1.0×10^8	10.1	11.94	2.222	37.04	16.29	5.496	35.43
100	6	1.0×10^{12}	14.7	95.63	22.84	420.9	60.88	20.56	139.8

Table 1: Algorithm Performance: Each row presents summary statistics across 1,000 simulations. Mean number of optimal ℓ refers to average number of distinct supply chain segments in cost function's lower envelope. Time units are seconds. Simulations implemented in Matlab using Gurobi to solve integer programs and on a 2019 Macbook Pro.

the complete set of optimal supply chains across all $q^F \in (0, \infty)$.

Performance

In terms of time complexity the algorithm improves upon brute force but still features exponential worst-case performance $O\left(J^N\right)$. To see why, imagine a degenerate case in which every single supply chain is on the lower envelope. In this case the algorithm will require J^N iterations of the zero-one integer programs. As is often the case, though, while worst-case performance is exponential, the algorithm is extremely fast in practice and it is unlikely that one would encounter a model parameterization implying worst-case performance.

Table 1 presents performance summary statistics for various degrees of model complexity. Specifically, I run 1,000 simulations of the GVC model for each pair J and N and solve each using both the marginal cost minimization program (5) and the fixed cost minimization program (6).

The algorithm runs very fast in practice. Moreover, the algorithm is efficient in that these statistics correspond to time needed to obtain the full cost function across all $q^F \in (0, \infty)$ (as opposed to brute force which minimizes at a single q^F). The algorithm is quick even in the highly dimensional cases where the solution space size is large. For reference, simply storing a vector of marginal or fixed costs of size 10^{10} in the brute force solution requires 75GB of memory; with $J = N = 20$ the solution space is on the order 10^{26} and requires 75 quadrillion GB! In contrast, the algorithm solves even the largest problems in less than a minute on average and often much faster. The fixed costs minimization algorithm, in particular, appears to do better in the very high dimensional cases.

Average time is increasing with J^N and is also closely related to the average number of distinct supply chains on the cost function lower envelope. This is expected since the latter proxies the number of iterations needed to fully trace out the cost function. Worst-case performance (as proxied by maximum time taken) is also very reasonable. Finally, note the algorithm can be parallelized in order to solve for the optimal cost curves of multiple goods at the same time. This is an useful property for embedding this setup in general equilibrium models with multiple goods.

4 Trade Policy Uncertainty and Supply Chain Disruption

What are the implications of trade policy uncertainty and the costs of supply chain disruption? I use the above model and the numerical machinery to shed light on how these issues might affect firms organizing their supply chains in the presence of increasing returns.

I conduct the following thought experiment: should a firm think twice before setting up its supply chain when there is uncertainty about how much it will cost to ship goods across borders in the future? And, should a firm fear the costs of supply chain disruption? To begin, the top panel of Figure 3 presents an example of the iterative algorithm's solution when $J = N = 5$, with the $J^N = 3125$ cost curves plotted for reference. Depending on how many final goods the firm wants to produce, there are 6 distinct optimal supply chains. Now imagine that the firm has to decide where to allocate its supply chain today, but that it constructs its factories and sells goods tomorrow. Will the firm think twice about committing to an optimal supply chain today if there is uncertainty about the future level of trade costs τ_{ij}?

To answer this question, I analyze how optimal supply chains $\boldsymbol{\ell}\left(q^F\right)$ and minimum total costs $C\left(\boldsymbol{\ell}\left(q^F\right), q^F\right)$ change at each q^F when trade costs change. Specifically, define shocked trade costs as $\tau'_{ij} = 1 + \zeta_{ij}\left(\tau_{ij} - 1\right)$ where τ_{ij} are the original trade costs underlying the solution in the top panel and ζ_{ij} is a mean one uniform random variable distributed over $\left[1 - \overline{\zeta}, 1 + \overline{\zeta}\right]$. For a given matrix of elements τ'_{ij}, I re-optimize and find the new set of optimal supply chains $\boldsymbol{\ell}'\left(q^F\right)$ and compare them to the original supply chains $\boldsymbol{\ell}\left(q^F\right)$ at each q^F. I then re-run this simulation 10,000 times and record the share of simulations in which the optimal supply chain changed at each q^F.

The bottom three panels in Figure 3 illustrate the impact of this form of trade cost uncertainty. The three panels depict the probability of the optimal supply chain changing with small ($\overline{\zeta} = 0.05$), medium ($\overline{\zeta} = 0.20$), and large shocks ($\overline{\zeta} = 0.40$). In other words, these probabilities proxy the firm's risk when committing to a given supply chain if it fears that trade costs might change.[10] Four regularities stand out. First, when $q^F \to 0$ the probability of the optimal supply chain changing is zero since only fixed costs matter. Second, the probability of switching at the original thresholds where the firm was indifferent between two supply chains is 100%. This is by construction as a small change in trade costs moves the thresholds and thus moves the optimal supply chain at the adjacent q^F. Third, the probability of switching is positive and substantial even between

[10] Note that $\mathbb{E}\left[\tau'_{ij}\right] = \tau_{ij}$ but $\mathbb{V}\left[\tau'_{ij}\right] = \frac{1}{3}\overline{\zeta}^2\left(\tau_{ij} - 1\right)^2$ so that higher $\overline{\zeta}$ implies more uncertainty.

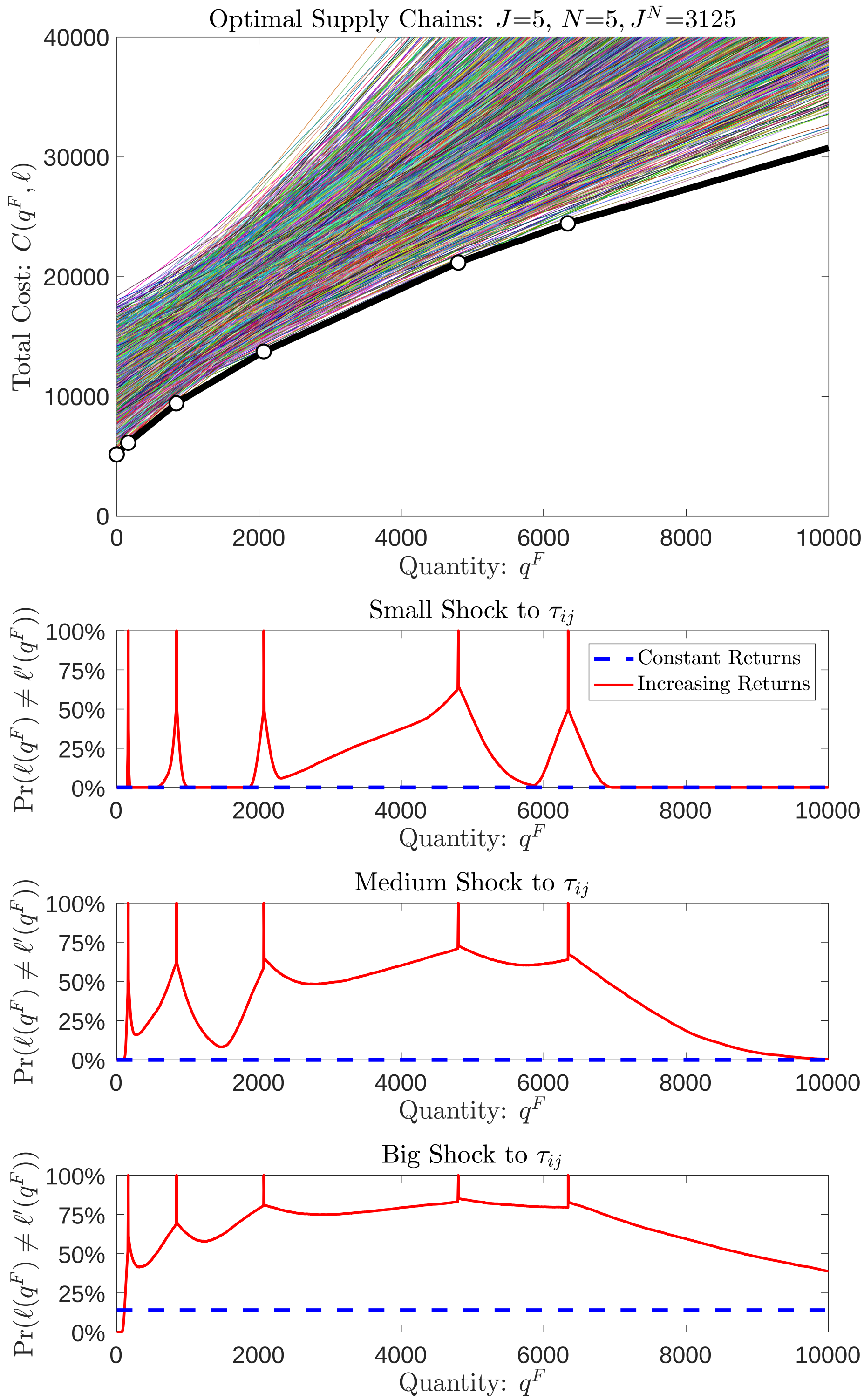

Figure 3: Optimal Supply Chains and Trade Policy Uncertainty: The top panel presents a numerical example with $J = N = 5$ with all $J^N = 3125$ cost curves and the lower envelope as computed by the iterative algorithm. The bottom three panels show the probability of the optimal supply chain $\boldsymbol{\ell}\left(q^F\right)$ changing at each q^F following a shock to τ_{ij}. Probabilities are constructed by shocking τ_{ij} across 10,000 simulations and computing the share of simulations in which $\boldsymbol{\ell}\left(q^F\right)$ shifts. The dotted line represents the probability of optimal $\boldsymbol{\ell}$ changing in the absence of fixed costs.

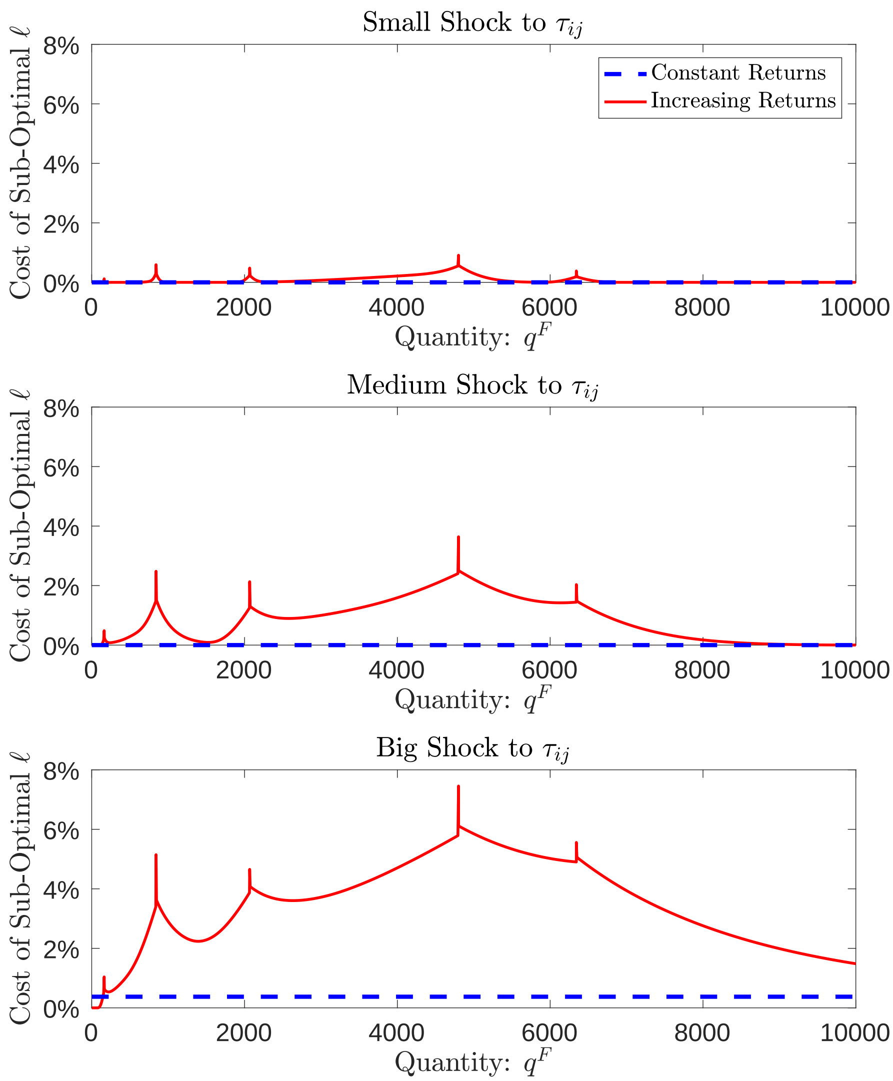

Figure 4: Costs of Supply Chain Disruption: Each panel plots the average ratio of $C\left(\boldsymbol{\ell}\left(q^F\right), q^F\right) / C\left(\boldsymbol{\ell}'\left(q^F\right), q^F\right) - 1$ across 10,000 simulations, with $\boldsymbol{\ell}\left(q^F\right)$ the optimal supply chain at the original τ_{ij}, $\boldsymbol{\ell}'\left(q^F\right)$ the optimal supply chain at the shocked trade costs τ'_{ij}, and with both cost functions evaluated at the shocked trade costs τ'_{ij}.

thresholds. Fourth, the probability of switching increases with risk.

Figure 3 also plots the probability of the optimal supply chain changing using the exact same model parameters but when shutting fixed costs down. In this case, there is a single global optimal supply chain and thus the probability of changing is constant across q^F. Two main patterns appear. First, the optimal supply chain only changes when there is a lot of uncertainty regarding trade costs. Second, the optimal supply chain is much more likely to shift in the presence of fixed costs. Both facts are intuitive: the optimal supply chain when $q^F \to \infty$ in the presence of fixed costs (which is equal to the optimal supply chain in the absence of fixed costs) dominates because it features the lowest marginal costs. In contrast, when q^F takes moderate levels, cost curves dominate partly because of low marginal costs and partly due to low fixed costs. Hence, when $q^F \to \infty$ there is little competition for being the optimal $\boldsymbol{\ell}\left(q^F\right)$ and trade cost uncertainty is of little consequence. But when q^F takes moderate levels there is much competition for being the optimal supply chain and trade cost uncertainty is highly costly. This is also easy to see in Figure 1: the optimal supply chain with $q^F \to \infty$ will only change if trade costs change by a lot, but the optimal supply chain for $q^F \in [10, 20]$ might easily switch even with small changes in trade costs.

Finally, Figure 4 estimates the cost of supply chain disruption using the same simulations as in Figure 3 and shows the cost of keeping the initially optimal supply chain when trade costs change. Specifically, each panel plots the cost of producing through $\boldsymbol{\ell}\left(q^F\right)$ relative to $\boldsymbol{\ell}'\left(q^F\right)$ when both cost functions are evaluated at the shocked trade costs τ'_{ij}. Four patterns are salient. First, the cost of supply chain disruption roughly mimics the probabilities of switching supply chains with the cost being highest at the original thresholds and lowest at the extremes. Second, while the switching probability is relatively high even with small trade cost changes, the cost of supply chain disruption is only substantial with medium and large shocks. Third, the losses can be quite substantial: cost increases in the order of 5-10% are highly detrimental to industries with razor-thin profits such as the auto industry. Fourth, and most importantly, constant returns to scale severely underestimate the costs of supply chain disruption. In the absence of fixed costs, big trade shocks imply supply chain disruption costs of less than 0.5%. With scale economies, however, disruption costs increase by an order of magnitude to around 5% on average (at intermediate levels of q^F).

These simulations thus provide suggestive evidence that shocks to trade barriers, such as the recent return to protectionism, i) have much larger effects in the presence of increasing returns and ii) have differential effects on firms of different size. While results are subject to various simplifications (all other parameters are constant, no general equilibrium effects, etc) they illustrate how the above model is useful for shedding light on key questions being debated in the GVC arena such as the impact of trade policy uncertainty on firms' supply chain decisions and the costs of supply chain disruption.

5 Conclusions

This paper developed a simple, tractable, and exact numerical solution method for multi-stage production models in the presence of internal economies of scale. The model can be extended in a number of ways. Ricardian comparative advantage can be embedded by defining $c_i^n = a_i^n w_i$, with w_i country i's wage and a_i^n a stage-specific productivity shifter. This can also be embedded in a general equilibrium model with multiple goods, as in Yi (2003) but with increasing returns, by modeling each firm z as having its own $a_i^n(z)$. Further, the same procedure can be used for alternative fixed costs setups such as multi-stage plants in which fixed costs are paid once regardless of how many stages a country produces. Total fixed costs thus take the form $\sum_{i\in\mathcal{J}} 1\left[i \in \boldsymbol{\ell}\right] \times F_i$.

Other extensions are more challenging but also more rewarding. Production featuring spiders, as in Tyazhelnikov (2019) but with increasing returns, is more realistic though much harder to model. Production with non-unitary elasticities of substitution across stages is also more realistic but requires even more powerful numerical methods since marginal costs are not log-separable. Finally, this paper has taken the stand that a single global firm organizes the full supply chain. Antràs and de Gortari (2017) showed that this solution exactly coincides with the decentralized equilibrium in which a single firm produces at each stage in the constant returns to scale case. But this does not hold with increasing returns. Much could be learned from studying GVCs with increasing returns in the context of a property rights model as in Antràs and Chor (2013).

Empirically, much remains to be done. Simulations illustrate that increasing returns amplify the effects of trade policy uncertainty and increase the costs of supply chain disruption. A stricter analysis could estimate the model's parameters using data and run specific counterfactual exercises. While beyond the scope of this short paper, the machinery for conducting such a study is now available. Other questions, such as the impact of industrial policy, the evolution of comparative advantage, and the dynamic importance of sunk costs and supply chain disruption, might also be studied.

References

American Chamber of Commerce in People's Republic of China (2019): "Second Joint Survey on the Impact of Tariffs," .

Antràs, P. and D. Chor (2013): "Organizing the Global Value Chain," *Econometrica*, 81, 2127–2204.

Antràs, P. and A. de Gortari (2017): "On the Geography of Global Value Chains," *NBER Working Paper 23456*.

Antràs, P., T. C. Fort, and F. Tintelnot (2017): "The Margins of Global Sourcing: Theory and Evidence from U.S. Firms," *American Economic Review*, 107, 2514–2564.

Arkolakis, C. and F. Eckert (2017): "Combinatorial Discrete Choice," .

CNBC (2019): "BlackRock's Fink: CEOs tell me they're pulling their supply chains out of China," .

Jia, P. (2008): "What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry," *Econometrica*, 76, 1263–1316.

Johnson, R. and A. Moxnes (2019): "GVCs and Trade Elasticities with Multistage Production," *NBER Working Paper 26018*, 1–42.

Krugman, P. (1980): "Scale Economies, Product Differentiation, and the Pattern of Trade," *The American Economic Review*, 70, 950–959.

Krugman, P. R. (1979): "Increasing Returns, Monopolistic Competition and International Trade," *Journal of International Economics*, 9, 469–479.

Melitz, M. J. (2003): "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71, 1695–1725.

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

Tyazhelnikov, V. (2019): "Production Clustering and Offshoring," .

Wall Street Journal (2019): "Manufacturers Move Supply Chains Out of China," .

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