Technology, Trade Costs, and the Pattern of Trade with Multi-Stage Production*

Robert C. Johnson† Andreas Moxnes‡

March 2013

Abstract

Comparative advantage and trade costs shape the geography of cross-border supply chains and trade flows. To quantify these forces, we build a model of trade with multi-stage production that features technology differences both across and within individual production stages. We estimate technology and trade costs in the model via simulated method of moments, matching bilateral shipments of final and intermediate goods for sixteen countries. Using the estimated model, we investigate the extent to which supply chains magnify trade elasticities and respond to changes in trade costs or productivity. We find no magnification effects for moderate changes in trade costs, but relatively large adjustments in supply chains.

*We thank Lorenzo Caliendo, Thibault Fally, Kjetil Storesletten, and Kei-Mu Yi for helpful comments, as well as seminar participants at Dartmouth, the Philadelphia Fed Trade Workshop, the University of Colorado, the 2013 ASSA/AEA Meetings, and the University of Oslo.

†Dartmouth College and NBER. E-mail: robert.c.johnson@dartmouth.edu.

‡Dartmouth College. E-mail: andreas.moxnes@dartmouth.edu.

In a global supply chain, sequential production stages are 'sliced up' and allocated across countries to minimize total production costs. Comparative advantage and trade costs govern the allocation of stages to countries. First, countries differ in the cost with which they perform individual production stages. Some countries have comparative advantage in downstream production stages (e.g., manufacturing assembly in China), while others have comparative advantage in upstream stages (e.g., production of disk drives in Japan). Second, as inputs are shipped from country to country through the chain, producers incur trade costs. Often these costs are paid multiple times as goods travel back and forth across borders. Further, the burden of these trade costs is large: ad valorem costs are paid on the gross value of goods shipped, but cost savings from moving marginal production stages only apply to a fraction of that gross value.

In this paper, we build a quantitative model of trade with cross-border supply chains to study the role of comparative advantage and trade costs in shaping production fragmentation and trade patterns. As in Yi (2003, 2010), the production of each good requires a discrete number of stages, which must be performed in sequence.[1] These stages are allocated across countries to minimize production costs, given both bilateral trade frictions and differences in technologies across countries and stages. In contrast to workhorse Ricardian models, such as Eaton and Kortum (2002), that emphasize comparative advantage across goods, the multi-stage model features comparative advantage across and within individual production stages.

We develop a new methodology to quantify the role of technology and trade costs in shaping world trade flows and fragmentation. Specifically, we estimate technology and trade costs via simulated method of moments by minimizing deviations between final and intermediate trade shares in the model and data. This estimation procedure is typically infeasible when the number of unknowns is high, but we overcome this problem by applying econometric techniques from the discrete choice literature.[2] Because we ask the model to match observed trade flows, we are able to estimate both trade costs and productivity as free parameters. This allows a tighter mapping between theory and data than previous calibration procedures.[3] Not only does this facilitate comparison between the multi-stage model and

[1]Related models with a discrete number of stages include Markusen and Venables (2007) and Baldwin and Venables (2010). Costinot, Vogel, and Wang (forthcoming) develop a model with a continuum of production stages, building on Dixit and Grossman (1982) and Sanyal (1983). Arkolakis and Ramanarayanan (2009) and Bridgman (2008, 2012) also work with models that feature vertical specialization.

[2]We discuss this issue further in Section 3.2.2 and Appendix C.

[3]E.g. Yi (2010) calibrates a related multi-stage model for the US and two Canadian regions using a mixture of data (on production, labor allocations, income, etc.) and parameter restrictions. Productivity is assumed equal across stages of production, which implies that comparative advantage *across* production stages is absent. We do not need to impose these restrictions on technology differences. Further, we are able to estimate trade costs, unlike Yi (2010) who measures trade costs based on auxiliary data.

competing alternatives, but it also paves the way for use of the multi-stage model in future applications.

We estimate the model using manufacturing trade data for 15 industrial and emerging market countries, plus a composite rest-of-the-world region, in 2004.[4] Our estimates show that there are large differences in manufacturing technology levels across countries and stages; for example, we find that China has a strong comparative advantage in downstream (stage 2) production, whereas the U.S. has a comparative advantage in upstream (stage 1) production. Differences in comparative advantage induce specialization in the product mix of a country's exports. Our estimates of downstream/upstream comparative advantage are therefore, not surprisingly, strongly (but not perfectly) correlated with the share of intermediates in a country's total exports.

We apply the estimated model in three quantitative exercises that advance our understanding of how trade costs and technology shape cross-border fragmentation and bilateral trade flows. First, we examine the role of multi-stage production in explaining the stylized fact that the elasticity of trade to trade barriers is high. Yi (2003) first pointed out that multi-stage production inflates the effect of tariffs on trade, while Yi (2010) argues that it also increases the influence of country borders on trade. Somewhat surprisingly, we conclude that fragmentation of production does not play an important role in inflating the trade elasticity.

Our main insight is that, in our model, the trade elasticity is a function of the *level* of trade costs. Because our estimates of trade costs are relatively high, breaking up the production process is costly, and therefore occurs relatively infrequently.[5] Hence, a given decline in trade costs does not produce amplified cost savings for most goods traded. As a consequence, the trade elasticity is not significantly inflated relative to a world without fragmentation. To demonstrate the role of the level of trade costs, we show that the trade elasticity is higher (in absolute value) in counterfactuals where the level of trade costs is significantly lower. An implication of our work is that structural estimates of trade costs obtained via models that do not take into account multi-stage production are, to a first approximation, unbiased.

Second, we examine the response of fragmentation and real wages to a uniform decline in geographic barriers to trade, perhaps due to improvements in transportation technology.

[4] It is straightforward to extend our procedure to allow for more sectors and/or stages, or to vary the number of countries. Because the estimation procedure is computationally costly, one needs to trade off these dimensions in practice. We have opted here for relatively few stages and a high level of aggregation to maximize country coverage, which allows richer cross-country analysis of counterfactual experiments.

[5] Our estimates of trade costs, which capture both observable and unobservable trade costs, are in line with the standard estimates in the literature.

Production becomes more fragmented as trade costs decline. For the median country, a 10 percent decline in trade frictions increases the share of foreign value added in domestic final goods production by 1.5 percentage points. The ratio of value added to gross trade for the world also falls by 2.5 percentage points. Changes in foreign sourcing and value-added to import ratios are particularly large for nearby countries, consistent with time-series evidence [Johnson and Noguera (2012b)]. For the median country, the real wage gain is 2 percent. Overall, the aggregate response of trade and real wages is very similar in our model compared to a model with only one stage of production.[6] This mainly reflects the finding that the trade elasticities are similar across models. Thus, our model yields macro-level responses comparable to more standard models, despite important differences in micro-level adjustment mechanisms.

Third, we examine the response of production chains and trade patterns to changes in productivity in one country. We present results for a scenario in which productivity doubles in China, with emphasis on studying how this change spills over and reorganizes production chains within Asia. Mirroring the finding above, production sharing is mainly concentrated among China's Asian neighbors.

The rest of the paper proceeds as follows. Section 1 lays out the many-country, multi-stage model and presents a solution procedure. Section 2 discusses some important features of the model. Section 3 describes how we estimate technology and trade costs, as well as how we calculate value added trade flows. Section 4 presents our estimates, while Section 5 presents our counter-factual analysis. Section 6 concludes.

1 Framework

We start this section by laying out the basic elements of the framework. We assume there are two sectors of the economy, which we refer to as manufacturing and non-manufacturing (including agriculture, natural resources, and services). The manufacturing sector features a discrete multi-stage production process, similar to the models by Yi (2003, 2010). The non-manufacturing sector features standard Ricardian production and trade, drawing on Eaton and Kortum (2003). After presenting the elements of the model, we discuss how we solve the model numerically, since the the model does not admit analytical solution.

[6] Specifically, we benchmark our results against a similar liberalization exercise in a two-sector version of the Eaton-Kortum model with input-output linkages developed by Caliendo and Parro (2012).

1.1 Production

Consider a world economy with many countries and two sectors. Countries are indexed by $i, j, k \in \{1, \dots, C\}$ and sectors are denoted by $s \in \{m, n\}$, standing for manfuacturing and non-manufacturing respectively. Within each sector, there is a unit continuum of goods, indexed by z and both sectors operate under perfect competition. By way of notation, we generally put country labels in the superscript. We put good and sector labels in parentheses, so that (z, s) denotes good z in sector s.

Manufacturing In the manufacturing sector, each good requires two stages to produce, and subscripts on each variable denote production stages. Production in stage 1 uses labor and a composite input, and we assume the production function for good z in sector m is:

$$q_1^i(z,m) = T_1^i(z,m)\Theta(m)X^i(z,m)^{\theta(m)}l_1^i(z,m)^{1-\theta(m)}, \tag{1}$$

where $T_1^i(z,m)$ is the good-specific productivity of country i in manufacturing stage 1, $l_1^i(z,m)$ and $X^i(z,m)$ are the quantities of labor and the composite input used in production, $\theta(m)$ is the share of the composite input in production, and $\Theta(m) = (1-\theta(m))^{1-\theta(m)}\,\theta(m)^{\theta}$ is a constant normalization.

The output of the first stage is an input that is used in stage 2 production of manufactured good z. Production in stage 2 combines the first stage input and labor, with the production function given by:

$$q_2^i(z,m) = T_2^i(z,m)\Theta(m)x_1^i(z,m)^{\theta(m)}l_2^i(z,m)^{1-\theta(m)}, \tag{2}$$

where $T_2^i(z,m)$ is productivity in stage 2, $x_1^i(z,m)$ is the quantity of the stage 1 input used, $l_2^i(z,m)$ is labor used, $\theta(m)$ is again the cost share attached to the stage 1 input, and $\Theta(m)$ is the same normalization as above.[7]

Output in each stage may be produced in any location, but every time output is shipped it incurs an bilateral sector-specific iceberg transportation cost $\tau^{ij}(m)$.[8]

[7]We do not explicitly include capital as a produced factor in the model. This implies that differences in capital are captured in the productivity term in our estimation. In computing counterfactuals, we implicitly hold all factors fixed. Including endogenous capital stocks, as in Yi (2003), would be a straightforward extension.

[8]Two points are worth noting here. First, we do not assume that the cost is stage-specific. Extensions in which trade costs for final and intermediate goods differ would allow one to consider the effects of input-tariff liberalization. Second, we assume that trade costs are ad valorem. Extensions with per unit trade costs, as in Irarrazabal, Moxnes, and Opromolla (2012), would give rise to differences in trade costs across stages because the gross value per unit shipped differs across stages.

Non-manufacturing In the non-manufacturing sector, each good is produced using one stage. Production requires labor and a composite input, and we assume for simplicity the that the composite input is the same composite input used by manufacturing. The production function for good z in sector n is then:

$$q^i(z,n) = T^i(z,n)\Theta(n)X^i(z,n)^{\theta(n)}l^i(z,n)^{1-\theta(n)}. \tag{3}$$

The notation here is the same as above, though here we drop the stage subscripts. Note here also that we allow the share of inputs in production to differ from that in the manufacturing sector.

As is standard, each non-manufacturing good can be produced in any location and shipping from source to destination incurs iceberg cost $\tau^{ij}(n)$, which may differ from shipping costs for manufactured goods.

1.2 Aggregation

Within each sector, goods are aggregated to form non-traded composites, which are sold to final consumers and used to form the composite input used in stage 1.[9] The composite goods, denoted $Q^i(s)$, are Cobb-Douglas combinations individual goods:

$$Q^i(s) = \exp\left(\int_0^1 \log(\tilde{q}^i(z,s))dz\right), \tag{4}$$

where $\tilde{q}^i(z,s)$ is the quantity of each good in sector s purchased (from low cost sources at home or abroad) by country i. For manufacturing, $\tilde{q}^i(z,m)$ represents purchases of stage 2 goods.

These sector-level composite goods are combined to form an aggregate final good and the composite input used by all producers in the non-manufacturing sector and stage 1 producers in manufacturing. We assume that the aggregate final good is given by: $F^i = AF^i(m)^\alpha F^i(n)^{1-\alpha}$, where $F^i(s)$ denotes the amount of the composite good in sector s that is sold to final consumers and $A = (1-\alpha)^{1-\alpha}\alpha^\alpha$. Similarly, the composite input is given by: $X^i = BX^i(m)^\beta X^i(n)^{1-\beta}$, with $X^i = \sum_s \left[\int_0^1 X^i(z,s)dz\right]$ in equilibrium and $B = (1-\beta)^{1-\beta}\beta^\beta$.[10] Finally, adding up requires that: $Q^i(s) = F^i(s) + X^i(s)$.

[9] One can think of this aggregation step as a third production stage, with zero value added. Because there is zero value added, one can alternatively write down the model with aggregation explicitly incorporated into preferences and production functions.

[10] Note that we assume the composite input is not sector-specific.

1.3 Households

Consumers supply labor inelastically to firms and consume the composite final good F_i. In effect, they therefore have Cobb-Douglas preferences over stage 2 goods. The consumer budget constraint is: $w^i L^i = P_F^i F^i + TB^i$, where w^i is the wage, L^i is the labor endowment, P_F^i is the price of the final composite, and TB^i is the nominal trade balance. The trade balance appears here in the budget constraint, since we treat it as an exogenous nominal transfer necessary to equate income and expenditure for each country.

1.4 Solving the Model

To estimate and simulate the model, we solve a discrete approximation of the continuum model described above. We assume that there are a large number (R) of goods within each sector, and let $r = \{1, \ldots, R\}$ index discrete products.

We describe the solution to the model by walking through a three step numerical procedure here. First, given wages w^i, we determine prices and sourcing decisions for each good. Second, given prices and this assignment, we find equilibrium quantities produced of each good. Third, given prices and quantities, we compute labor demand and check whether this matches labor supply in each country.

1.4.1 Prices and Sourcing Decisions

Given wages, we can describe the optimal sourcing decisions for each destination market. For non-manufacuting, this amounts to determining who the low cost suppliers are for each good to each destination. For manufacturing, we need to solve for the optimal assignment of stages to countries for production of all goods purchased by each destination.[11] We solve both problems by comparing prices across alternative sources, or alternative allocations of stages to countries for manufactured goods, for delivery of a given good to each destination, and picking the assignment that minimizes costs.

For manufacturing, this takes the form of a nested minimization problem:

$$\begin{aligned} &\tilde{p}_2^k(r,m) = \min_j \tau^{jk}(m) p_2^j(r,m), \quad \text{with} \quad p_2^j(r,m) = \frac{\left(w^j\right)^{1-\theta(m)} \left(\tilde{p}_1^j(r,m)\right)^{\theta(m)}}{T_2^j(r,m)} \\ \text{and} \quad &\tilde{p}_1^j(r,m) = \min_i \tau^{ij}(m) p_1^i(r,m), \quad \text{with} \quad p_1^i(r,m) = \frac{\left(w^i\right)^{1-\theta(m)} \left(P_X^i\right)^{\theta(m)}}{T_1^i(r,m)}, \end{aligned} \tag{5}$$

[11] To be clear, the assignment of stages to countries for each good depends on the destination at which that good is consumed.

where the price of the composite intermediate input is given by P_X^i (which we define below).

To be clear, $\tilde{p}_2^k\,(r,m)$ is the realized price of stage 2 output of good r in manufacturing (i.e., the price at which k actually purchases good (r,m)). It is equal the minimum over $\tau^{jk}(m)p_2^j(r,m)$, the possible prices at which each country j could deliver the stage 2 good if j chooses the minimum cost source for stage 1. These prices are in turn a function of the cost of stage 1 inputs in each source j, where the low cost supplier of stage 1 goods delivers inputs to country j at price $\tilde{p}_1^j\,(r,m)$. This input supply price is the minimum over delivered prices from alternative source countries (i): $\tau^{ij}(m)p_1^i(r,m)$. Finally, those input supply prices depend on the composite input price in country i, which itself is a function of the realized prices of stage 2 output in country i.

For non-manufacturing, the minimization problem is simpler due to the 'one-stage' structure:

$$\tilde{p}^k\,(r,n) = \min_j \tau^{jk}(n)p^j\,(r,n)\,, \quad \text{with} \quad p^j\,(r,n) = \frac{\left(w^j\right)^{1-\theta(n)}\left(P_X^i\right)^{\theta(n)}}{T^i\,(r,n)}. \tag{6}$$

Prices in both sectors depend on the price of the composite input. This composite input price is given by: $P_X^i = P^i(m)^\beta P^i(n)^{1-\beta}$, where $P^i(s)$ denotes the price of the composite good in each sector. These composite goods prices themselves are given by:

$$P^i(m) = \exp\left(\frac{1}{R}\sum_r \log(\tilde{p}_2^k\,(r,m))\right), \tag{7}$$

$$P^i(n) = \exp\left(\frac{1}{R}\sum_r \log(\tilde{p}^k\,(r,n))\right), \tag{8}$$

where these price indexes are discrete approximations to a price index defined for the continuum of goods in each sector.[12]

Starting with a guess for the composite input prices P_X^i, we can solve the minimization problem for prices $\{\tilde{p}_1^j\,(r,m)\,,\tilde{p}_2^k\,(r,m)\,,\tilde{p}^k\,(r,n)\}$. We then use these prices to update the value of P_X^i, and solve for an updated set of prices. We iterate on this fixed point problem to convergence. Having converged on a value for P_X^i, we can easily compute the solution for equilibrium prices, as well as the allocation of stages to countries for manufactured goods. For manufacturing, we denote the set of countries to which country i is the low cost supplier for a particular stage of each good as: $\{\Omega_1^i\,(r,m)\,,\Omega_2^i\,(r,m)\}$. And $\Omega^i\,(r,n)$ denotes the set of low cost suppiers to i for non-manufactured good r.

[12]The continuum price indexes take the form: $\log\left(P^i(m)\right) = \int_0^1 \log(\tilde{p}^k\,(z,m))dz$ and $\log\left(P^i(n)\right) = \int_0^1 \log(\tilde{p}^k\,(z,n))dz$.

1.4.2 Quantities Supplied and Demanded

Given the prices obtained in the previous step, we can compute production of goods at each stage in each country by working backwards from final demand. Total demand for the sector-level composite goods is given by:

$$P^k(m)Q^k(m) = \alpha P_F^k F^k + \beta P_X^k X^k \tag{9}$$

$$P^k(n)Q^k(n) = (1-\alpha)P_F^k F^k + (1-\beta)P_X^k X^k. \tag{10}$$

Since we have taken the wage as given and observe the trade balance, we know $P_F^k F^k = w^k L^k - TB^i$. However, we do not directly observe expenditure on the composite input $P_X^k X^k$. We therefore need to solve for this value.

Given $P_F^k F^k$ and a guess for X^k, we can compute $P^k(s)Q^k(s)$ for both sectors. This allows us to solve for demand for stage 2 goods in manufacturing and demand for individual goods in non-manufacturing in each destination.

Starting with manufacturing, demand for individual stage 2 goods in destination k is given by:

$$\tilde{q}^k(r,m) = \frac{\frac{1}{R}P^k(m)Q^k(m)}{\tilde{p}_2^k(r,m)}, \tag{11}$$

where again $\tilde{p}_2^k(r,m)$ is the delivered price from the actual source that supplies market k. Tracing these demands back to the countries that supply those goods, we can compute the quantity of stage 2 goods produced in each source j as:

$$q_2^j(r,m) = \sum_{k \in \Omega_2^j(r,m)} \tau^{jk}(m)\tilde{q}^k(r,m). \tag{12}$$

Then, given this stage 2 production in country j, demand for stage 1 inputs in country j is:

$$x_1^j(r,m) = \frac{\theta(s)p_2^j(r,m)q_2^j(r,m)}{\tilde{p}_1^j(r,m)}. \tag{13}$$

These input demands allow us to then solve for the quantity of each stage 1 good supplied by country i as:

$$q_1^i(r,m) = \sum_{j \in \Omega_1^i(r,m)} \tau^{ij}(m)x_1^j(r,m). \tag{14}$$

Turning to non-manufacturing, demand for individual goods in destination k is given by:

$$\tilde{q}^k(r,n) = \frac{\frac{1}{R}P^k(n)Q^k(n)}{\tilde{p}^k(r,n)}, \tag{15}$$

where $\tilde{p}^k(r,n)$ is the delivered price from the actual source that supplies market k in the non-manufacturing sector.

Finally, given stage 1 production in the manufacturing sector and production in the non-manufacturing sector, we can compute demand for the composite input:

$$X^i = \frac{1}{P_X^i}\left[\sum_r \theta(m)p_1^i(r,m)q_1^i(r,m) + \sum_r \theta(n)p^i(r,n)q^i(r,n)\right] \tag{16}$$

This gives us an updated value for purchases of the composite input X^i, and hence updated values for the total amount of the composite goods supplied in each sector $P^k(s)Q^k(s)$. We iterate on this fixed point problem to convergence.

1.4.3 Labor Market Clearing

The candidate solution above involved a guess for wages, so we need to check whether this guess clears the labor market. We can calculate total labor demand from both stages in manufacturing as:

$$l_1^i(r,m) = (1-\theta(m))\frac{p_1^i(r,m)\,q_1^i(r,m)}{w^i} \tag{17}$$

$$l_2^i(r,m) = (1-\theta(m))\frac{p_2^i(r,m)\,q_2^i(r,m)}{w^i}. \tag{18}$$

And labor demand in non-manufacturing is:

$$l^i(r,n) = (1-\theta(n))\frac{p_1^i(r,n)\,q_1^i(r,n)}{w^i} \tag{19}$$

Then total labor demand is: $L_D^i(\mathbf{w}) = \sum_r l^i(r,n) + \sum_r \left[l_1^i(r,m) + l_2^i(r,m)\right]$, where we have made total labor demand explicitly a function of the wage vector. The equilibrium wage vector then sets labor demand equal to labor supply: $L_D^i = L^i$ for $i = 2,..,N$ (where market 1 is dropped appealing to Walras' law). We choose w^1 as the numeraire.

2 Discussion

In this section, we comment on three aspects of the model. First, we comment on the mix between sequential multistage versus roundabout production in the model. Second, we briefly describe how this production structure can be represented in an input-output accounting framework, with details in the appendix. Third, we review intuition about the elasticity of trade to trade costs in models with multi-stage production.

2.1 Snakes and Spiders

Within the manufacturing sector, the model mixes sequential, multi-stage production with a roundabout input loop, so that input use consists of stage 1 inputs and the composite input. In contrast, for the non-manufacturing sector, there is only a roundabout input loop. We illustrate the basic set-up in a closed economy in Figure 1.

Borrowing terminology from Baldwin and Venables (2010), we can think of the manufacturing technology as characterized by recursive 'snakes' and 'spiders.' To isolate the 'snake' part of the model, suppose that we were to set the share of the composite input in stage 1 to zero in the manufacturing sector. In that case, we would re-write the first stage production function as $q_1^i(z,m) = T_1^i(z,m) l_1^i(z,m)$. And the stage 2 output would be used to satisfy final demand and as an input in the non-manufacturing sector. This would turn the manufacturing sector into a continuum of two-stage 'snakes' – similar to the model analyzed in Yi (2003).

In the full model, a 'spider' production process links the two ends of the 'snakes' in manufacturing.[13] Output from the second stage is aggregated and included in the composite input that is fed back into the first stage of the production process. Because the composite input links ends of the sequential production process, it converts the two-stage process into a multi-stage process with an effectively infinite number of production stages, where some fraction of output is drawn out at each stage to satisfy final demand. Not only does the spider link the ends of the sequential production process, it also links the production process across sectors. Manufactures uses non-manufactures in production (and vice versa) because the composite input is made from all goods. Put differently, inputs flow across sectors through spiders, while inputs flow within the manfuacturing sector both via snakes and spiders.

2.2 Input-Output Accounting in the Model

In the model, there are input-output linkages across production stages, sectors, and countries. These linkages can be represented in the form of a model-based global input-output (IO) table. Analogous to data-based IO tables, this model-based IO table records information on bilateral shipments of final and intermediate goods across sectors and countries.

The chief difference between the model-based IO table and data-based tables is that we can split input shipments that occur between stages 1 and 2 in manufacturing from input shipments for use in forming the composite input. In a sense, we observe input linkages at higher resolution in the model than we do in the data. Therefore, we adapt the model-based IO accounting framework to take this into account. The key modification is that we

[13] Within the non-manufacturing sector, inputs obviously flow through a spider only.

introduce two rows/columns in the IO table for manufacturing, corresponding to production stages. We discuss the resulting model-based IO table at length in Appendix A. Having constructed this IO table, we can use it to construct measures of value-added trade in the model. We describe two measures along these lines in Section 3.3.

2.3 Elasticity of Trade to Trade Costs

In the model with multi-stage production above, the elasticity of trade flows to trade costs depends on the level of trade costs. We discuss this relationship at length in simulations of the model below, but pause here to develop some intuition for this result. There are two observations that underpin our interpretation of the trade elasticity.

First, as trade costs rise in the multi-stage model, it becomes increasingly costly to split up manufacturing production stages across countries. Therefore, as trade costs increase, stages are more often co-located, which implies that the model behaves more like a standard 'one-stage' multi-sector Ricardian model. That is, we expect the model to behave like a multi-sector extension of the Eaton and Kortum (2002).[14] Since the single-stage model features a constant partial (i.e., holding prices fixed) elasticity of trade to changes in trade costs, the multi-stage model will also feature a near constant partial trade elasticity at high levels of trade costs. At high levels of trade costs, the multi-stage model will also generate changes in trade and welfare that are similar to the Ricardian benchmark, as we highlight below in discussion of our empirical results.

Second, as trade costs fall in the multi-stage model, it becomes increasingly attractive to split up discrete production stages across borders to take advantage of cost differences. The ability to substitute over the location of individual stages of the production process, rather than simply over entire goods themselves, tends to amplify the sensitivity of trade to changes in trade costs.

One force for amplification arises because trade costs are incurred multiple times when inputs are shipped abroad and then embodied in imported final goods. For example, if a good that is exported uses an imported input, then one pays ad valorem costs on the input twice – once when it is imported, and again when it is exported embodied in final good. Yi (2010) refers this as the 'multiple border crossing' force.

A second force for amplification arises because agents evaluate the burden of trade costs relative to the cost savings on shifting the location of a single stage of the production process. The benefits of moving the location of a single stage depend on that stage's share in total value added (equivalently, the value of the final good), and benefits are lower when the share

[14] Specifically, we use the model by Caliendo and Parro (2012) as a benchmark against which to evaluate our results.

in total value added is also low. The trade costs incurred in shifting a production stage are thus perceived to be more burdensome when total value added in the marginal stage is low. Yi (2010) refers to this as the 'effective rate of protection' force.

At intermediate levels of trade costs, the model economy features both standard Ricardian trade, where consumers substitute across entire goods, and trade through multi-stage production chains in which agents substitute over production locations for each stage. Therefore, the aggregate model elasticity of trade to trade costs depends on the mix of Ricardian vs. multi-stage multistage trade. As trade costs fall, the share of trade via multi-stage production chains rises, and the elasticity of trade to trade costs does as well.

3 Mapping the Model to Data

In this section, we discuss how we fit the model presented in Section 1 to data. We begin by presenting our data source. We then discuss how we calibrate a subset of the parameters of the model and estimate the remainder via simulated method of moments. We conclude the section by describing how we use the model-based and data-based input-output tables to compute measures of trade in value added.

3.1 Data

Our data source is the GTAP 7.1 Data Base assembled by the Global Trade Analysis Project at Purdue University, which includes trade, production, and input-output data for 2004.[15] While the underlying data includes more than 90 countries, we cannot use this fine country detail due to computational constraints. Therefore, we retain 15 major countries – United States, China, Japan, Germany, Italy, India, Great Britain, France, Canada, Spain, Brazil, Australia, Russia, Mexico, and South Korea – and aggregate the remaining countries to form a composite rest-of-the-world region.[16] Further, there are 57 sectors in the underlying data.

In the data, we have information on 6 objects for each country: y^i is a 57×1 vector of total gross production, f^{ii} is a 57×1 vector of domestic final expenditure (which includes consumption, investment, and government purchases), f^{Ii} is a 57×1 vector of domestic final import expenditure, A^{ii} is a 57×57 domestic input-output matrix, A^{Ii} is a 57×57 import input-output matrix, and $\{x^{ij}\}$ is a collection of 57×1 bilateral export vectors for exports from i to j.

[15] See the GTAP website at `http://www.gtap.agecon.purdue.edu` for documentation of the source data. This is the same dataset used in Johnson and Noguera (2012a).

[16] Because we have input-output data for the countries that comprise the rest-of-the-world region, we can aggregate them in a way that preserves basic input-output identities for the world as a whole.

Using these data, we compute bilateral final and intermediate input shipments using a proportionality assumption. To write this out, let us define $Inter^{Ij} = A^{Ij} diag(y^j)$ to be the matrix of imported intermediate use with (s, s') elements equal to the value of imported sector s inputs used by sector s' in country j. For pairs $i \neq j$, we compute disaggregate final bilateral and intermediate goods shipments as: $Inter^{ij}(s, s') = Inter^{Ij}(s, s') \left(\frac{x^{ij}(s)}{\sum_{k \neq j} x^{kj}(s)} \right)$ and $Final^{ij}(s) = f^{Ij}(s) \left(\frac{x^{ij}(s)}{\sum_{k \neq j} x^{kj}(s)} \right)$. For each country buying from itself, shipments are reported directly in the data, with $Final^{ii}(s) = f^{ii}(s)$ and $Inter^{ii}(s, s')$ equal to the (s, s') element of $A^{ii} diag(y^i)$.

We aggregate these final and intermediate shipments to the two-sector level – manufacturing and non-manufacturing – to estimate the model.[17] For reference, we will use bilateral final and intermediate goods shipments for the manufacturing sector to estimate the model below, which are defined as: $Final^{ij}(m) = \sum_{s \in m} Final^{ij}(s)$ and $Inter^{ij}(m) = \sum_{s \in m} \sum_{s'} Inter^{ij}(s, s')$. Rewriting these as shares of destination expenditure, we obtain: $Fshare^{ij}(m) = \frac{Final^{ij}(m)}{\sum_k Final^{kj}(m)}$ and $Ishare^{ij}(m) = \frac{Inter^{ij}(m)}{\sum_k Inter^{kj}(m)}$.

Due to the proportionality assumptions used above, aggregate final and intermediate import shares differ across partners due to differences in the composition of trade across partners.[18] For example, $Fshare^{ij}$ will be high (relative to $Fshare^{kj}$) when country j's imports from country i are high (relative to country k) in sectors that account for a large share of total final import demand. The fact that composition drives bilateral variation in trade shares raises the concern that our trade shares understate the true extent of variation in the data.[19] To check our trade shares, we can compare them to trade shares in an alternative database (the WIOD database) that does not use the proportionality assumption to estimate bilateral final and intermediate shipments.[20] For both $Fshare^{ij}$ and $Ishare^{ij}$, the correlation between trade shares in our data and this alternative data is above 0.99. Thus, at the aggregate level at which we use the data, the proportionality assumption appears innocuous.

[17] Manufacturing covers sectors 27-42 in the GTAP data. Non-manufacturing covers all other sectors, including agriculture, natural resources, food products, and services.

[18] There are large differences in the extent to which individual manufacturing sectors are used as intermediate versus final goods. These patterns are broadly sensible. For example, the textiles sector tends to be heavily used as an intermediate input, while the apparel sector tends to be dominated by final goods. More generally, machinery and transport equipment tend to be used relatively intensively as final goods. While chemicals, paper/wood products, metals and metal products, and mineral products all tend to be used intensively as intermediate inputs.

[19] To the extent that they do, they exert a conservative bias on our results, pushing the behavior of the multi-stage model closer to Ricardian models that typically assume identical bilateral sourcing patterns for final and intermediate goods within each sector.

[20] The WIOD database uses a modified Broad Economic Categories (BEC) classification to assign HS 6-digit categories to final or intermediate end use, corresponding to final vs. intermediate classifications in the national accounts. See `http://www.wiod.org/`.

3.2 Fitting the Model

There are a number of free parameters, including technology levels $\{T_1^i(r,s), T_1^i(r,s)\}$, trade costs $\tau^{ij}(s)$, and several share parameters in production functions and preferences $\{\theta(s), \alpha, \beta\}$. We mix calibration and estimation in pinning down these parameters.

3.2.1 Calibrated Parameters

We calibrate $\{\theta(s), \alpha, \beta\}$ to match ratios for 'typical' countries in the data. The parameter $\theta(s)$ governs the value added to output ratio in each sector.[21] We therefore set $\theta(s)$ to match the median value added to output ratio across countries in each sector. These median values are: $\theta(m) = 0.67$ and $\theta(n) = 0.43$.[22] These imply that the value added to output ratio is 0.24 lower for manufactures than non-manufactures.

The parameter α is also straightforward to calibrate, since it is the share of manufactures in final expenditure. We also set this value equal to the median across countries, given by $\alpha = 0.18$. Finally, β governs the extent to which manufactures versus non-manufactures are used in forming the composite input. As we describe in Appendix B, we choose β to match inter-sectoral flows for the world economy as a whole. This yields a value $\beta = 0.2$. Note that β and α are quite similar, so manufactures and non-manufactures receive similar weights in the composite input and final demand.

3.2.2 Estimation via Simulated Method of Moments

Parameters The remaining unknown parameters are technology levels and trade costs.

For technology levels, we assume that countries draw productivity from country, stage, and sector specific Fréchet distributions, where draws are assumed to be independent across countries/stages/sectors.[23] We parameterize these distributions with a common shape parameter κ, and location parameters $\{T_1^i(s), T_2^i(s)\}$ for sector s in country i. We set $\kappa = 4.12$, guided by Simonovska and Waugh (2011), and normalize $T_1^1(s) = T_2^1(s) = 1$, so that technology levels are measured relative to country 1.

We parameterize trade costs by assuming that bilateral trade costs are a power function in distance. Specifically, we estimate a function of the form: $\tau^{ij}(s) = \tau(s)\left(d^{ij}\right)^{\rho}(s)$, where

[21] To see this, note that value added in manufacturing is equal to the wage bill for each good r: $va^i(r,m) = w^i l_1^i(r,m) + w^i l_2^i(r,m) = (1-\theta(m))[p_1(r,m)q_1^i(r,m) + p_2(r,m)q_2^i(r,m)]$. Adding up across goods yields: $1-\theta(m) = \frac{va^i(m)}{p_1(r,m)q_1^i(r,m)+p_2(r,m)q_2^i(r,m)}$.

[22] The input share for manufactures varies from roughly 0.57 to 0.78 across countries, with most countries between 0.6 and 0.7. The input share for non-manufactures varies from 0.37 to 0.55.

[23] We have experimented with allowing draws to be correlated across stages for individual goods. The parameter governing this correlation is weakly identified by the data and introducing this correlation does not materially affect the results.

d^{ij} is the distance between country i and country j, $\tau(s)$ is a level parameter, and $\rho(s)$ is the elasticity of trade costs to distance. We set trade costs on domestic shipments to one in all countries ($\tau_{ii}(s) = 1$).

The unknown parameters are then $\Theta = \{T_1^i(m), T_2^i(m), T^i(n), \tau(m), \tau(n), \rho(m), \rho(n)\}$. As the computational requirements for estimating the full set of parameters is very high, we estimate a restricted version of the full model. Specifically, we restrict trade costs and techologies in the non-manufacturing sector, and this allows us to estimate the model using manufacturing trade only.[24] Two assumptions are required. First, we restrict $T^i(n)$ to be equal to the geometric mean of $T_1^i(m)$ and $T_2^i(m)$. Second, we assume that trade costs are equal in both sectors, i.e., $\tau^{ij}(m) = \tau^{ij}(n) = \tau\left(d^{ij}\right)^{\rho}$. This reduces the set of parameters to be estimated to be $\tilde{\Theta} = \{T_1^i(m), T_2^i(m), \tau, \rho\}$.

Moments We estimate $\tilde{\Theta}$ by matching the simulated data to measured shipments of manufactured final and intermediate goods across countries. We generate simulated data taking final expenditure in each market $P_F^i F^i$ and relative wages w^i as given and set equal to values in the data (with $w^1 = 1$ as a normalization).[25]

Given calibrated parameters $\{\kappa, \theta(s), \alpha, \beta\}$, relative wages, and final expenditure, we choose starting values of $\tilde{\Theta}$ and draw productivities for each good and sector. Given these parameters, we solve the model following the procedure in Section 1.4. Using the simulated data from the model, we then compute a vector of moments which we match to analogous moments in the data.

We form the first set of moments using trade shares for final manufactured goods. Shipments of final manufactured goods from i to j are:

$$Final^{ij}(m) = \frac{1}{R} \sum_r \mathbb{I}\left(j \in \Omega_2^i(r,m)\right) \alpha(m) P_F^j F^j,$$

where $\alpha(m) P_F^j F^j$ is total expenditure on manufactured goods in j and $\frac{1}{R}\sum_r \mathbb{I}\left(j \in \Omega_2^i(r,m)\right)$ is the share of stage 2 goods that i supplies to country j. The share of source i in country j final expenditure on manufactures is: $Fshare^{ij} = \frac{Final^{ij}(m)}{\sum_k Final^{kj}(m)}$.[26]

[24] In a previous draft, we estimated a version of the model above using data for goods (including agriculture, natural resources, and manufacturing) and services sectors. To estimate that alternative model, we assumed that goods were produced with a multi-stage production process and that the services sector was non-traded (i.e., had infinite trade costs). Similar to the approach in this draft, this allowed us to use goods trade data only in estimation. The core results we report below are very similar using this alternative strategy.

[25] We compute total expenditure in each market by dividing observed final expenditure on goods value by α. We also observe the trade surplus for goods TB^i in the data. Since the consumer budget constraint is given by $w^i L^i = P_F^i F^i + TB^i$, then we back out wages as: $w^i = \left(P_F^i F^i + S^i\right)/L^i$. We use 2004 population data from the Penn World Table 7.1 to proxy for labor endowment L^i.

[26] Note that in the model, these trade shares are identical to the trade shares for all stage 2 output,

The second set of moments consists of trade shares for manufactured inputs. Input shipments from country i to j include both stage 1 goods and stage 2 goods destined for the composite input. Input shipments from i to j of sector m goods are:

$$Inter^{ij}(m) = \sum_{r} \mathbb{I}\left(j \in \Omega_1^i(r,m)\right)[\theta(m)p_2^j(r,m)q_2^j(r,m)] + \frac{1}{R}\sum_{r} \mathbb{I}\left(j \in \Omega_2^i(r,m)\right)[\beta P_M^j M^j].$$

Then the share of inputs from source i in country j's total purchases of manufactured inputs in destination j is: $Ishare^{ij} = \frac{Inter^{ij}(m)}{\sum_k Inter^{kj}(m)}$.

These final and intermediate trade shares correspond to data shares defined in Section 3.1. Since the final and intermediate trade shares sum to one, we only use off diagonal trade shares ($i \neq j$), in total $2\left(N^2 - N\right)$ moments. We denote the log difference between actual and simulated moments $\mu^{ij}(\Theta) = \ln m^{ij} - \ln \hat{m}^{ij}(\Theta)$, and stack all μ^{ij}'s in a column vector $M(\Theta)$.

Procedure Our estimation procedure is based on the moment condition $E\left[M\left(\tilde{\Theta}_0\right)\right] = 0$, where $\tilde{\Theta}_0$ is the true value of $\tilde{\Theta}$. Hence, we estimate a $\hat{\tilde{\Theta}}$ that satisfies:

$$\arg\min_{\tilde{\Theta}} \left\{ M\left(\tilde{\Theta}\right)' M\left(\tilde{\Theta}\right) \right\}$$

Since we have $2N$ variables and $2\left(N^2 - N\right)$ moments, the model is over identified.

This minimization problem is not straightforward to solve numerically, since the simulated moments are not continuous in the underlying parameters. In Appendix C, we show that the expressions for the trade shares can be re-written in such a way that they resemble choice probabilities from the discrete choice literature. We can therefore employ a technique to smooth the objective function from McFadden (1989). With this smoothing, we can turn to standard numerical routines to solve the minimization problem.

3.3 Value Added in Trade

In evaluation of the model, we report responses of trade measured in both gross and value-added terms. Value-added measures serve as convenient summary statistics regarding the nature of changes in supply chains in the counter-factual episodes that we examine. Further, measures of trade in value added have received attention in a number of recent contribu-

including output dedicated for intermediate use. So they can alternatively be computed using total stage 2 trade flows.

tions.[27] We are therefore interested in the extent to which the model we estimate is capable of helping us interpret stylized facts documented in this literature. We focus on two alterantive metrics of value-added trade.

First, we compute the amount of value added from each source country embodied in final goods produced by a given country, which we refer to here as 'value added inputs.' This type of metric has been used by Erumban, Los, Stehrer, Timmer, and de Vries (2012a, 2012b) to describe how 'slices' of manufacturing production chains are allocated across countries. Increases in the share of foreign value added in final goods production imply that larger slices of the value chain are located abroad.

Second, we compute the amount of value added from each source country consumed in each destination, which we refer to as 'value-added exports' as in Johnson and Noguera (2012a). In the aggregate, the ratio of value-added exports to gross exports is a convenient summary statistic for the extent of double-counting in trade data. This double-counting is a marker for production fragmentation. For example, increased use of imported inputs in production of exports is associated with declines in the ratio of value-added exports to gross exports. At the bilateral level, value-added exports can be thought of as capturing country j's demand for value added from country i. Changes in bilateral value-added relative to gross trade are therefore associated with changes in reduced form demand linkages between countries.

The common element in both calculations is the observation that multiplying the Leontief inverse of the global input-output matrix by a vector of final goods returns the amount of gross output (by country and sector) needed to produce those final goods. These gross output requirements can then easily be converted to value added requirements, by multiplying by value added to output ratios. We describe the details regarding how we compute these measures in both the model and data in Appendix D.

4 Estimation Results

In this section, we present our estimates for technology and trade cost parameters, and discuss model fit. We also evaluate the magnitude of the partial elasticity of bilateral trade to trade costs in our model.

[27] Among others, see work by Trefler and Zhu (2010), Daudin, Rifflart, and Schweisguth (2011), Erumban, Los, Stehrer, Timmer, and de Vries (2012a, 2012b), Johnson and Noguera (2012a, 2012b, 2012c), and Koopman, Wang, and Wei (2012).

4.1 Technology and Trade Costs

Technology We present estimates for technology levels by stage for the 15 countries and the composite region in Table 1. The first two columns present the geometric means $T_1^i(z,m)$ and $T_2^i(z,m)$ in each country, expressed relative to relative to the United States.[28] The final column computes the ratio of mean technololgy in stage 2 relative to stage 1 in each country.

The estimates indicate that most countries have technology levels lower than the U.S. level. We plot the estimated aggregate technology level for each country – computed as the geometric mean of the stage technology levels in Table 1 – against income per capita in Figure 2. As expected, average technology levels are highly correlated with income per capita.

Based on examination of Table 1, it is evident that productivity levels are correlated across stages. Countries with high absolute productivity in stage 1 tend to also have high absolute productivity in stage 2. Despite this correlation, there are sizable differences in relative stage productivities across countries. To be clear about interpretation, the final column Table 1 measures relative productivity in each country compared to relative productivity in the U.S. Numbers greater than one indicate that a country has a comparative advantage in stage 2 (downstream) production relative to the U.S. Scanning the table, the U.S. and Australia have a comparative advantage in upstream production, while China, India, and South Korea all have a strong comparative advantage in downstream production. Other countries lie between the these extremes.

Trade Costs Turning to estimated trade costs, we assumed that the trade cost function took the form: $\tau^{ij} = \tau\,(d^{ij})^{\rho}$. Our estimate of the elasticity of trade costs to distance is $\rho = 0.29$, and the level parameter is $\tau = .26$.[29] For the country pair separated by the median distance in our data (8400km), these estimates imply that international trade costs are 3.57 times (257% higher than) domestic trade costs: $\tau^{median} = \exp\left(\ln\tau + \rho\ln(8400)\right) = 3.57$. These costs are large, but in line with standard estimates in the literature.[30]

[28]Since $T_n^i(z,m)$ is drawn from a Fréchet, the unweighted geometric mean is given by $\exp(\gamma/\kappa)T_n^i(m)^{1/\kappa}$, where γ here is the Euler-Mascheroni constant. Since we normalize each country relative to the U.S., the numbers reported in the table are effectively $\left(\frac{T_n^i(m)}{T_n^{US}(m)}\right)^{1/\kappa}$.

[29]This magnitude of this level parameter is not directly interpretable since it depends on the units in which we measure distance. Therefore, we focus on total implied trade costs, which are interpretable as costs of international relative to domestic trade.

[30]For example, Eaton and Kortum (2002) return estimated distance costs of roughly 300% for country pairs in the 3000 to 6000 mile distance range. Anderson and van Wincoop argue trade costs are equivalent to an ad-valorem tax of 170% for a representative rich country.

Model Fit Before turning to detailed analysis of the model, we quickly summarize how the model fits various moments in the data. We start by examining how the model fits the moments that we have targeted in estimation – intermediate and final goods trade shares – in Figure 3, the true trade shares are on the x-axis and simulated trade shares are on the y-axis (both log scales).[31] The model generally fits these trade shares well. In Figure 4, we plot the share of final goods exports in total exports for each country, in the data as well as in the model. For most countries, the model is able to match their export composition, e.g. the relatively high share of final goods in Mexican and Chinese exports. The final goods share is also highly correlated with the estimated $T_2^i(m)/T_1^i(m)$, showing that identification of comparative advantage partly comes from variation in the final goods share across countries.

Turning to untargeted moments, we are able to reproduce variation in bilateral measurs of value added trade. In the top panel of Figure 5, we plot value added input shares – the share of value added from each bilateral source country in production of final goods for a given country. In the bottom panel of Figure 5, we plot value added to export ratios.[32]

The model does an excellent job at replicating bilateral sourcing of value added. The model also generates a positive correlation between value added to export ratios in actual and simulated data, though the overall fit here is not as tight. The dimension on which the model misses is in generating value added to export ratios near/above one, which are observed in the actual data but not the simulated data.

Finally, we report several reduced form correlations between bilateral trade and distance that are helpful for interpreting counterfactuals below. To do this, we estimate a simple gravity regression of the form:

$$\log y^{ij} = \chi^i + \chi^j + \delta \ln Distance^{ij} + e^{ij}, \tag{20}$$

where where y^{ij} is either actual or simulated total bilateral exports ($Exports^{ij}$) or VAX ratios ($VAExports^{ij}/Exports^{ij}$), and χ^i and χ^j are exporter and importer fixed effects.[33]

We present the results in Table 2. Not surprisingly, the model is able to reproduce the well known dampening effect of distance on trade, producing a distance coefficient of -1.12, slightly larger than the -0.99 in the actual data. Further, the model reproduces the positive correlation between value added to export ratios and distance, though the magnitude is

[31]The cluster of points in the upper right corner is the share of each country's purchases from itself. Not surprisingly, these own shares are uniformly large.

[32]In both figures, we include domestic as well as cross-border transactions. The cluster of points in the upper right of the top panel is the share of domestic value added in final goods production for each country. We include these domestic transactions because these are important moments for the model to replicate in order to generate the correct degree of aggregate openness for each country.

[33]In estimating this regression, we include exports to/from the rest of the world. Results are virtually identical if we exclude these flows.

somewhat smaller in the simulated than actual data.[34]

4.2 The Trade Elasticity

Yi (2003, 2010) argues that the elasticity of trade to changes in trade costs is magnified in multi-stage models, and suggests that this magnification is important for understanding the response of world trade to multilateral tariff liberalization and the large effect of borders on trade. We seek to quantify these magnification effects (if any) in our estimated model.

In standard models that generate CES import demand, the elasticity of bilateral trade to trade costs, $\partial \ln Exports^{ij} / \partial \ln \tau^{ij}$, is constant and equal to a structural parameter – e.g. equal to the Fréchet parameter in Eaton and Kortum (2002) or the CES demand elasticity in Anderson and van Wincoop (2003). In our model, we can calculate $\partial \ln Exports^{ij} / \partial \ln \tau^{ij}$ using the chain rule:

$$\frac{\partial \ln Exports^{ij}}{\partial \ln \tau^{ij}} = \frac{\partial \ln Exports^{ij}}{\partial \ln Distance^{ij}} \Bigg/ \frac{\partial \ln \tau^{ij}}{\partial \ln Distance^{ij}} \ . \tag{21}$$

Using this equation and our simulated data, we can compute the partial elasticity of exports to trade costs implied by our model. In our simulated data, the first term – the elasticity of exports to distance – is -1.12 (see Table 2). Further, we have structurally estimated the the second term – the elasticity of trade barriers to distance (ρ) – to be 0.29. This implies that the the partial elasticity of exports to trade costs is -3.86 in our model. Recall that the single-stage elasticity is $\kappa = 4.12$. Therefore, we conclude that the trade elasticity is not magnified in our estimated model. In other words, supply chains seem to have a limited role in explaining the trade elasticity, compared to standard models.[35]

Looking at this result in another way, this finding suggests that standard gravity coefficients measuring the elasticity of exports with respect to distance are not inflated in the presence of multi-stage production. Rearranging (21), the distance elasticity of trade is $\frac{\partial \ln Exports^{ij}}{\partial \ln Distance^{ij}} = \left(\frac{\partial \ln Exports^{ij}}{\partial \ln \tau^{ij}}\right)\left(\frac{\partial \ln \tau^{ij}}{\partial \ln Distance^{ij}}\right)$. If the trade elasticity is not inflated, then the estimated distance coefficient will not be inflated either. Rather, it will reflect the true fundamental elasticity of trade costs with respect to distance.

[34] Johnson and Noguera (2012b) present and discuss this correlation of VAX ratios and distance at length.

[35] Note that even with zero fragmentation, our model would not produce a trade elasticity precisely equal to $\kappa = 4.12$. If we were to restrict our model to disallow fragmentation and force both stages for each good to be located in the same country, then the productivity with which output would be produced would be a combination of productivities in the first and second stage, given by $T_2^i(z,m)T_1^i(z,m)^{\theta(m)}$ for good z. Despite the fact that $T_2^i(z,m)$ and $T_1^i(z,m)$ are drawn from Fréchet distributions, this composite productivity is not distributed Fréchet. Therefore, even in this special case, the model would not yield an aggregate trade elasticity equal to the Fréchet parameter.

A potential concern is that, as the Frechet parameter κ is not identified by our data, we must rely on an estimate from external sources. Therefore, we would like to know whether we get magnification of the trade elasticity for other values of κ. In a simple robustness check, we estimate the model based on $\kappa = 8$, and then calculate the two elasticities in equation (21), as before. This gives us $\rho = 0.14$ and a distance elasticity of -1.09, implying an elasticity of exports to trade costs of -7.8. Our results are therefore robust to alternative choices of κ.[36]

We pick these results up again in Section 5.2.1, where we explore why our findings differ from Yi (2003, 2010) in greater detail.

5 Technology and Trade Cost Counterfactuals

We now turn to analysis of several counterfactuals in the model. We start with a scenario in which we lower trade costs across all partners. We then examine a large positive productivity shock in China. Both counterfactuals will help us understand how fragmentation responds to various shocks, as well as why we do not find evidence of magnification of the trade elasticity.

5.1 A Single-Stage Benchmark

To evaluate the extent to which the multi-stage model delivers responses (e.g., changes in trade or welfare) that differ from single-stage models, we need to specify an appropriate single-stage benchmark. Because the multi-stage model does not strictly nest a single-stage Ricardian model, the construction of such a benchmark requires some care. Our approach is to use a version of the multi-sector Eaton-Kortum style model developed by Caliendo and Parro (2012) that is calibrated to match our simulated data.

Following Dekle, Eaton, and Kortum (2008) and Caliendo and Parro (2012), we only need a few key parameters to run counterfactuals.[37] These parameters include bilateral trade shares, income, sector-level production and expenditure, sector-level input cost shares, and the trade surplus in an initial equilibrium, plus a value for the Fréchet shape parameter. We obtain these values using our simulated data. This means that we run counterfactuals in this single-stage model starting from an equilibrium that is observationally equivalent in key

[36] A higher κ implies that the estimated level of trade costs is lower. This does not, however, mean that supply chains are more common (relative to when $\kappa = 4.12$). Lower trade costs means that fragmentation is more profitable. On the other hand, a higher κ means that there is less heterogeneity in efficiency, so that that fragmentation is less profitable. The two forces cancel out, so that the magnitude of fragmentation is invariant to the choice of κ.

[37] Importantly, we do not need information about the trade cost function and technology parameters to solve for counterfactuals in the model.

dimensions to our estimated multi-stage model.[38] For each counterfactual, we then feed the same change in parameter values through our multi-stage model and this alternative single stage model. For reference, we describe the exact specification, calibration, and solution procedure for this single-stage model in Appendix E.

5.2 Reductions in Trade Costs

Starting from the baseline estimated equilibrium, we first analyze the case where every bilateral trade barrier in manufacturing declines, e.g. due to due to improvements in transportation technology.[39] We open the analysis by studying how the level of trade costs influences the extent to which the trade elasticity is magnified in a multi-stage model. We then turn to examining the response of gross trade, fragmenatation and value added trade, and real wages to a 10% change in trade costs. In both sets of analysis, we highlight similarities and differences between results for the multi-stage and single-stage benchmark models.

5.2.1 Magnification Effects and the Level of Trade Costs

In Section 4.2, we argued that our estimated model does not produce a significantly magnified elasticity of trade to changes in trade costs, which seems to contradict findings in Yi (2003, 2010). This result bears further examination.

In this section, we argue that a key reason we do not find large magnification effects is that the degree of magnification depends on the level of trade costs. Our structural estimates of international trade costs are relatively high, which implies that breaking up the production process is relatively costly, and therefore occurs relatively infrequently. This limits the extent of magnification in our baseline equilibrium.

Distance Elasticity of Trade Recalling Section 4.2, the trade elasticity can be decomposed into $\partial \ln Exports^{ij}/\partial \ln Distance^{ij}$ and $\partial \ln \tau^{ij}/\partial \ln Distance^{ij}$, which equals ρ in our structural model. Holding ρ constant, magnification in the trade elasticity is equivalent to magnification in the distance elasticity of trade. Therefore, we would like to show that $|\partial \ln Exports^{ij}/\partial \ln Distance^{ij}|$ increases as the level of trade costs declines in the multi-stage case, but is constant in the single-stage case.

[38] Our simulated data naturally includes information about shipments between production stages. We discard this information in calibrating the Caliendo-Parro model, since the meaning of a production stage is undefined in that model.

[39] Prior to running counterfacutals, we exogenously close aggregate trade imbalances in the multi-stage model. We use this balanced trade equilibrium data to parameterize the single-stage model.

To do this, we estimate $\partial \ln Exports^{ij} / \partial \ln Distance^{ij}$ in simulated data for different levels of trade costs in the multi-stage model. This amounts to re-estimating Equation 20 for multiple cross-sections of simulated data, where each cross-section is obtained by re-simulating the model for a different level of trade costs. Note that as we change the level of trade costs, we hold relative trade costs across bilateral trade partners constant.[40] Starting from our baseline equilibrium, we reduce the level of trade costs by around 200 percentage points, taking the mean bilateral trade cost in the model from near 250% to 50%. To illustrate the non-linear response in the model, we also increase trade costs by around 150 percentage points, relative to the baseline.

We plot the resulting distance coefficients against the mean bilateral trade cost in Figure 6. The solid line indicates distance elasticities for the multi-stage multistage model. The dotted line represents the implied trade elasticity for the single-stage benchmark model, which features a constant trade elasticity and hence distance elasticity. The vertical dotted line indices the baseline equilibrum, where average iceberg costs are near 250%.

As trade costs fall, the absolute value of the distance elasticity increases in the multistage case. This is the same thing as saying that the absolute value of the trade elasticity rises. As we lower trade costs, this increase is initially gradual, but accelerates sharply as average trade costs fall to around 150%. With ad valorem trade costs at 50%, the multistage elasticity is inflated by roughly 13% relative to the baseline elasticity. In other words, a lower *level* of trade costs make trade more sensitive to distance in a world with multistage production.

To shed light on the mechanism driving this result, we plot the share of goods for which production stages are collocated – meaning that both stages 1 and 2 are produced in the same location – in Figure 7. In the baseline equilibrium (vertical dotted line), 70% of goods are collocated, indicating that the magnitude of fragmentation is not large enough to significantly inflate the trade elasticity. As trade costs fall, stages are more often fragmented across countries, and so the effective trade elasticity rises.

One final point to note in Figures 6 and 7 is that both the distance elasticity and the share of goods with collocated stages are non-linear in the level of trade costs. These non-linearities are obviously linked. Fragmentation becomes more sensitive to changes in trade

[40] In lowering the level of trade costs by large amounts, we run into a subtle technical constraint in analyzing the resulting data. As the level of trade costs falls, country pairs with initially low trade costs may hit the lower bound of $\tau^{ij} = 1$. Once at this bound, we cannot lower trade costs further for these pairs. As a result, we start to distort relative distances in the data when we proceed to lower trade costs among all other pairs (i.e., pairs not at the lower bound). To estimate gravity coefficients using a constant set of relative distances and trade costs at different absolute levels of trade costs, we drop these country pairs for which the lower bound is attained from all regressions (both the high and low trade cost regressions). Because we drop some countries here, the reduced form gravity coefficient is somewhat higher than in our full data set, which is reflected in a slightly larger initial absolute value at the highest level of trade costs in Figure 6.

costs as the level of trade costs falls, which implies larger extensive margin changes in trade in response to further changes in trade costs and larger trade elasticities. This also explains why the trade elasticity is roughly constant for higher levels of trade costs (i.e. to the right of the vertical dotted line)

Bilateral Trade Elasticities Thus far, we have focused primarily on the *average* partial equilibrium elasticity of trade with respect to trade costs. That is, the elasticity that one would obtain from running Armington or Eaton-Kortum gravity-style regressions. Importantly, this elasticity is estimated under the assumption that the trade elasticity is constant and common to all countries.

This constant elasticity assumption does not hold in our model: trade elasticities are heterogeneous across countries and country pairs in the multi-stage model. This follows directly from the observation above that trade elasticities depend on the level of trade costs. For particular countries or country pairs that have relatively low trade costs (e.g., nearby countries like the U.S. and Canada), we may expect to find larger trade elasticities.

To illustrate this heterogeneity, we calculate the elasticity of trade for individual country pairs for the same set of trade costs changes that we examine above. To be precise, for each level of trade costs, we compute the elasticity of bilateral trade for a country pair for an additional 10% reduction in global trade costs.

In Figure 8, we plot the resulting trade elasticities for U.S. exports to Canada, Mexico, Japan, and Germany. The horizontal axis shows the level of trade costs for the country-pair. The vertical dotted line in each figure then depicts the elasticity of trade to trade costs for a 10% change in trade cost starting from the baseline equilibrium. Note that the level of trade costs corresponding to the dotted vertical line differs across countries, since the level of trade costs is different across bilateral pairs in the baseline equilibrium. The vertical axis depicts the elasticity of bilateral exports.

In the figure, we plot bilateral elasticities for both the multi-stage and benchmark single-stage Ricardian model. An important point to note in interpreting these elasticities is that these are general equilibrium, not partial equilibrium, elasticities. As we lower the level of trade costs in this experiment, wages and price levels are changing, and these changes influence trade. Therefore, despite the fact that the partial equilibrium elasticity in the single-stage Ricardian model is constant, the general equilibrium elasticity actually declines as the level of trade costs falls. Therefore, the key information in the graphs is measured by the divergence between the size of the elasticities between the multi-stage and single-stage models.

Two points stand out in the figures. First, in the baseline equilibrium, the multi-stage elasticities are virtually identical to the single-stage elasticities in all cases, even for exports to Canada. This implies that the response of bilateral trade to changes in trade costs is quite similar in the multi-stage and single-stage models near the baseline equilibrium, a point we return to below.

Second, the elasticities slowly diverge as the level of trade costs falls, with the level of $\tau^{ij} \approx 2$ marking the point of divergence in elasticities across all the country pairs. At this point, the multi-stage elasticity stops falling and starts increasing for all country pairs. As is evident, the elasticity in the single-stage Ricardian model continues to fall beyond this point. The divergence between these two elasticities marks the point where multi-stage elasticity magnification effects start to overwhelm the tendency for elasticities to fall in Ricardian models, leading the elasticity to start rising. An important point to note about this inflection point is that magnification occurs earlier for countries with relatively low initial trade costs (i.e., Canada and Mexico), as compared to Japan and Germany.

Comparison to Yi (2003, 2010) In sum, the finding that the level of trade costs matters for magnification helps explain why our results differ from Yi (2003, 2010). Our methodology captures both observable and unobservable trade costs, and therefore points to relatively high trade cost estimates, consistent with evidence surveyed by Anderson and van Wincoop (2004). In contrast, Yi (2003, 2010) uses data on observable trade costs, which are much smaller.[41] For example, Yi (2010) estimates that measurable tariff and distribution costs are in the range of 10-40% for US-Canada trade. Our estimates instead suggest that total observable plus unobservable trade costs are near 138% for US-Canada trade. Yi (2003) analyzes tariff liberalization starting with initial tariffs near 15%. Our estimates instead put average trade costs (multilateral barriers, roughly speaking) near 250%.

At low levels of trade costs in the range used by Yi (2003, 2010), we do find that our model is capable of generating substantial magnification. However, this low level of trade costs is far too low to rationalize the observed home bias and concentration in trade flows observed in the data. Therefore, we believe our results pointing to the lack of magnification provide a more reasonable guide for interpreting estimated trade elasticities for representative country samples. That said, it is possible to detect amplification effects at current levels of trade costs for particular sectors or country-pairs with low trade costs.

[41] It is natural to ask how the low level of trade costs assumed in Yi's analysis yield realistic levels of trade. In Yi (2003), one reason is that the model is calibrated to two symmetric countries. For example, suppose that we examine a symmetric two-country Ricardian model, and let us assume for simplicity that relative wages are equal to one and constant (e.g., set in an outside, freely traded sector). With Frchet productivity distributions, then the share of imports in GDP is given by: $\frac{\tau^{-n}}{1+\tau^{-n}}$. If $n \in (3, 6)$, then one needs trade costs between 20-44% to achieve an import share of 0.25. In our model and data, even very small countries exhibit extreme home bias, which requires much higher trade costs.

5.2.2 Gross Trade, Value Added Trade, and Real Wages

We now turn to exploring the response of gross trade, value added trade, and real wages to a 10% decrease in global trade costs from the baseline equilibrium.

Figure 9 depicts changes in gross trade in both the multi-stage and benchmark single-stage model. Given the findings in the previous section regarding the similarity of trade elasticities across models near the baseline equilibrium, it is not surprising that trade responses are very similar across the two models. Overall, the median increase in exports (imports) is 27% (28%) percent.

Turning to trade measured in value added terms, we plot responses for two measures – the share of foreign value added in final goods production and the value-added to export ratio – in Figure 10. Lower trade costs lead to higher foreign shares in final goods production and lower value added to export ratios. For the median country, the foreign value added input share increases by 1.5 percentage points, while the VAX ratio decreases by 2.5 percentage points. The fall in the global VAX ratio, i.e. the change in world value added trade relative to gross trade, is also 2.5 percentage points. To set this magnitude in context, Johnson and Noguera (2012b) find that the global VAX ratio declined by 6 percentage points during the period 1995-2005. Thus, modest declines in global trade costs could account for nearly half of this change.

Looking at bilateral changes in value added to export ratios, we find that declines tend to be larger among nearby partners. To illustrate this, we plot changes in aggregate bilateral VAX ratios in Figure 11. While VAX ratios fall for most (though not all) bilateral pairs, declines tend to be largest for countries separated by short distances, in the lower left part of the figure. The dependence of these changes on distance is also consistent with patterns documented in Johnson and Noguera (2012b). The most surprising aspect of this result is that these differential changes arise following a *global* reductions in trade cost. They do not require trade costs reductions to be larger for nearby partners.

Finally, we plot changes in real wages, w^i/P^i_F, in Figure 12. Real wages increase by 2% for the median country, and the response is very similar across the single- and multi-stage case. Again, this reflects the finding that the trade elasticities are similar across models near the baseline equilibrium. There is, however, considerable heterogeneity across countries. Overall, real wages increase more in markets with higher initial import shares, since imported goods have a larger share of the price index in these countries.

In sum, the model generates changes in value-added trade following modest reductions in global trade costs that are qualitatively and quantitatively consistent with time-series evidence. Moreover, responses for gross trade and real wages are quite similar in single-stage and multi-stage models. This is a consequence of the fact that estimated trade costs are relatively high, which implies that the amplification effects of vertical cross-border production are weak. That said, the mechanics of adjustment and shuffling of the allocation of stages to countries in the multi-stage model are of course different than substitution across goods in the single-stage model. In other words, the same macro-behavior disguise different underlying micro-behavior.

5.3 Changes in Technology

We now turn to investigating how improvements in local technology in one country induce changes in global production chains. We focus on one experiment here: productivity growth in China. Starting with our estimated model, we increase both T_1^{CHN} and T_2^{CHN} by two log points, which brings Chinese technology to a level higher than Mexico but somewhat lower than South Korea.

Figure 13 illustrates the changes in value added to export ratios and the share of foreign value added in final goods production following the change in Chinese technology. The largest adjustment occurs within China itself, where the VAX ratio rises and the foreign value added sourcing share falls. Both adjustments reflect the fact that China sources a larger fraction of inputs from itself.

Nearby countries – such as Japan, Australia and South Korea – experience the opposite adjustment. For those countries, VAX ratios fall and foreign value-added input shares rise, as China supplies more intermediate inputs into production in the Asian region. For European countries and the U.S., the adjustments are generally small. In Figure 14, we plot changes in foreign value-added input shares by country on the vertical axis against distance from China on the horizontal axis. The takeaway from both figures is that adjustments in the production chain are mostly confined to proximate trading partners. As in the case of the decline in trade costs above, this again reflects the predominantly local scope of production chains.

6 Conclusion

Despite substantial academic and policy interest in the rise of global supply chains, few quantitative models incorporate discrete, multi-stage production processes. In contrast, this paper puts the decision to collocate or fragment production stages at center stage. This allows us to quantify the role of technology and trade costs in driving fragmentation, as well as the role of fragmentation in magnifying trade elasticities. We find that while elasticity magnification effects are small in our estimated model, these effects depend on the initial level of trade costs. As a result, magnification effects can be substantial for countries or country pairs with initially low levels of trade costs. We also find fairly large changes in supply chains following global reductions in trade costs, and show that the model can help us interpret changes in both gross and value added trade.

In addition to the questions we address in this paper, our model and estimation strategy is well suited to address many other interesting questions. Maintaining our focus on understanding trade elasticities, we think it would be worthwhile to extend the empirical exercises in several directions.

First, anecdotal evidence suggests that sourcing decisions are 'sticky', in the sense that it is difficult for firms to adjust their production chains in response to shocks. The resulting sluggish adjustment in the allocation of stages to countries, wherein sourcing patterns adjust more in the long run than in the short run, implies that short run and long run elasticities diverge from one another. Our model can be used to quantify these differences.

Second, while we presented results for a 'small' (10%) change in global trade costs, we think it would be useful to revisit the model's predictions for 'large' changes – e.g., movements toward free trade and/or autarky. Near our estimated baseline equilibrium, the model behaves much like a two-sector, single-stage Ricardian model. As we raise trade costs, this similarity continues to hold, since fragmentation becomes even less common. However, as we lower trade costs, fragmentation becomes more common, making our model look less like the single-stage Ricardian benchmark. Therefore, further analysis of large liberalization exercises – whether among subsets of countries or for the world as a whole – would shed new light on dimensions where the multi-stage view of production leads to conclusions regarding trade elasticities and welfare that differ from the alternative single-stage view of production.

Third, while we found no aggregate magnification effects in our baseline equilibrium, there is substantial heterogeneity in how quickly magnification effects kick in across countries and country pairs, due to the dependence of trade elasticities on the level of trade costs. This type of heterogeneity may have important implications for quantifying the effects of trade policy changes and regional integration initiatives. For example, this logic suggests that the effects of trade agreements on trade would differ across trading partners, with trade agreements having larger effects for country pairs with initially low levels of trade costs. Further, it suggests that regional integration initiatives may have larger effects on trade than non-regional ones, since by definition geographic trade frictions are lower for intra-regional trade than for extra-regional trade. In both cases, multi-stage production could play a key role in explaining growth in bilateral or regional trade.

References

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

Anderson, James E., and Eric van Wincoop. 2004. "Trade Costs." *Journal of Economic Literature*, XLII: 691–751.

Arkolakis, Costas, and Anath Ramanarayanan. 2009. "Vertical Specialization and International Business Cycle Synchronization." *Scandinavian Journal of Economics*, 111(4): 655–680.

Baldwin, Richard, and Anthony Venables. 2010. "Relocating the Value Chain: Offshoring and Agglomeration in the Global Economy." NBER Working Paper 16611.

Baldwin, Richard, and Daria Taglioni. 2011. "Gravity Chains: Estimating Bilateral Trade Flows When Parts and Components Trade Is Important." NBER Working Paper 16672.

Bridgman, Benjamin. 2008. "Energy Prices and the Expansion of World Trade." *Review of Economic Dynamics*, 11: 904–916.

Bridgman, Benjamin. 2012. "The Rise of Vertical Specialization." *Journal of International Economics*, 86: 133–140.

Caliendo, Lorenzo, and Fernando Parro. 2012. "Estimates of the Trade and Welfare Effects of NAFTA." NBER Working Paper 18508.

Costinot, Arnaud, Jonathan Vogel, and Sue Wang. forthcoming. "An Elementary Theory of Global Supply Chains." *Review of Economic Studies*.

Daudin, Guillaume, Christine Rifflart, and Daniele Schweisguth. 2011. "Who Produces for Whom in the World Economy?" *Canadian Journal of Economics*, 44: 1403–1437.

Dekle, Robert, Jonathan Eaton, and Samuel Kortum. 2008. "Global Rebalancing with Gravity: Measuring the Burden of Adjustment." *IMF Staff Papers*, 55(3): 511–540.

Disdier, Anne-Celia, and Keith Head. 2008. "The Puzzling Persistence of the Distance Effect on Bilateral Trade." *The Review of Economics and Statistics*, 90(1): 37–48.

Dixit, Avinash K., and Gene M. Grossman. 1982. "Trade and Protection with Multistage Production." *Review of Economic Studies*, 49(2): 583–594.

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

Erumban, Abdul Azeez, Bart Los, Robert Stehrer, Marcel Timmer, and Gaaitzen de Vries. 2012*a*. "Fragmentation, Incomes and Jobs. An Analysis of European Competitiveness." WIOD Working Paper Number 9.

Erumban, Abdul Azeez, Bart Los, Robert Stehrer, Marcel Timmer, and Gaaitzen de Vries. 2012*b*. "Slicing Up Global Value Chains: The Role of China." WIOD Working Paper Number 12.

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

Irarrazabal, Alfonso, Andreas Moxnes, and Luca David Opromolla. 2012. "The Tip of the Iceberg: A Quantitative Framework for Estimating Trade Costs." Unpublished Manuscript, Dartmouth College.

Johnson, Robert C., and Guillermo Noguera. 2012*a*. "Accounting for Intermediates: Production Sharing and Trade in Value Added." *Journal of International Economics*, 82(2): 224–236.

Johnson, Robert C., and Guillermo Noguera. 2012*b*. "Fragmentation and Trade in Value Added over Four Decades." NBER Working Paper 18186.

Johnson, Robert C., and Guillermo Noguera. 2012*c*. "Proximity and Production Fragmentation." *The American Economic Review: Papers and Proceedings*, 102(3): 407–411.

Koopman, Robert, Zhi Wang, and Shang-Jin Wei. 2012. "Tracing Value-added and Double Counting in Gross Exports." NBER Working Paper No. 18579.

Markusen, James R. forthcoming. "Expansion of trade at the extensive margin: A general gains-from-trade result and illustrative examples." *Journal of International Economics.*

Markusen, James R., and Anthony J. Venables. 2007. "Interacting factor endowments and trade costs: A multi-country, multi-good approach to trade theory." *Journal of International Economics*, 73(2): 333–354.

McFadden, Daniel. 1989. "A method of simulated moments for estimation of discrete response models without numerical integration." *Econometrica*, 57: 995–1026.

Sanyal, Kalyan K. 1983. "Vertical Specialization in a Ricardian Model with a Continuum of Stages of Production." *Economica*, 50(197): 71–78.

Simonovska, Ina, and Michael Waugh. 2011. "The Elasticity of Trade: Estimates and Evidence." NBER Working Paper 16796.

Train, Kenneth. 2009. *Discrete Choice Methods with Simulation.* . Second ed., New York:Cambridge University Press.

Trefler, Daniel, and Susan Chun Zhu. 2010. "The Structure of Factor Content Predictions." *Journal of International Economics*, 82: 195–207.

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

Yi, Kei-Mu. 2010. "Can Multi-Stage Production Explain the Home Bias in Trade?" *The American Economic Review*, 100(1): 364–393.

Table 1: Estimated Technology

	Stage 1	Stage 2	$\frac{Stage2}{Stage1}$
United States	1	1	1
China	0.20	0.52	2.61
Japan	0.67	1.19	1.77
Germany	0.69	1.37	1.98
Italy	0.65	1.07	1.64
India	0.10	0.31	3.01
United Kingdom	0.63	1.01	1.60
France	0.59	0.97	1.63
Canada	0.58	0.87	1.51
Spain	0.53	0.83	1.56
Brazil	0.26	0.43	1.66
Australia	0.64	0.69	1.06
Russia	0.25	0.36	1.43
Mexico	0.35	0.61	1.75
South Korea	0.39	0.92	2.39
Rest of World	0.16	0.93	5.75

Note: Columns labeled Stage 1 and Stage 2 report the geometric mean of the Fréchet distribution for each manufacturing stage: $\exp(\gamma/\kappa)T_n^{1/\kappa}$, where γ is the Euler-Mascheroni constant and $\kappa = 4.12$ as in our simulated model. Average technology levels are normalized to one in the U.S. in both stages. Relative productivities $\frac{Stage2}{Stage1}$ therefore measure comparative advantage across stages relative to U.S. comparative advantage.

Table 2: Distance and Trade in Data and Model

	$\log(Exports^{ij})$		$\log(VAX^{ij})$	
	Data	Model	Data	Model
Log Distance	-0.99***	-1.12***	0.20***	0.14***
	(0.05)	(0.02)	(0.02)	(0.01)
R^2	0.90	0.99	0.60	0.94
N	240	240	240	240

All regressions include exporter and importer fixed effects. Robust standard errors in parentheses.

Figure 1: Diagram of Production Process in a Closed Economy

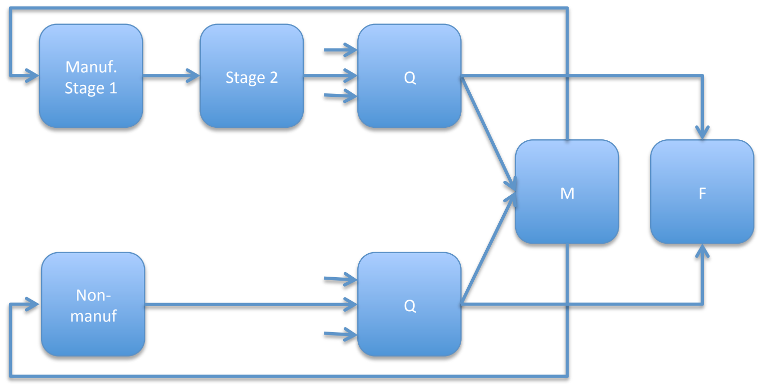
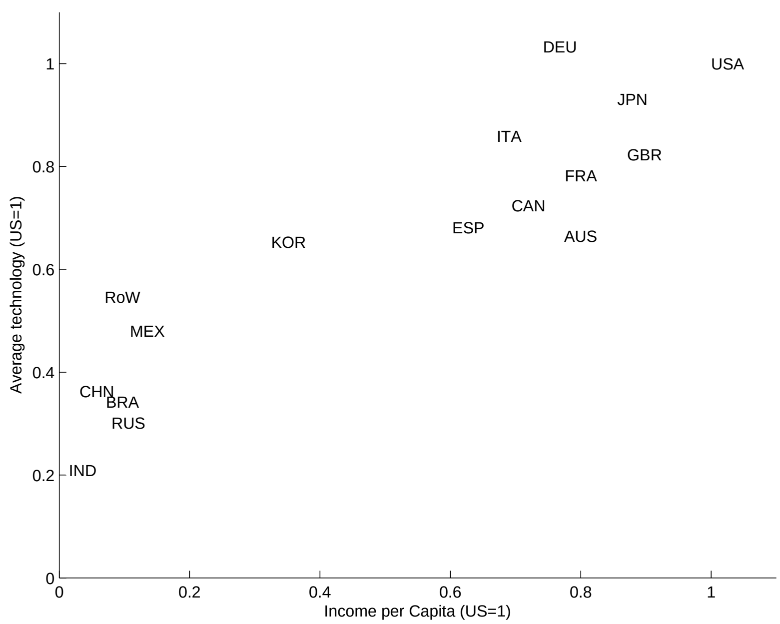
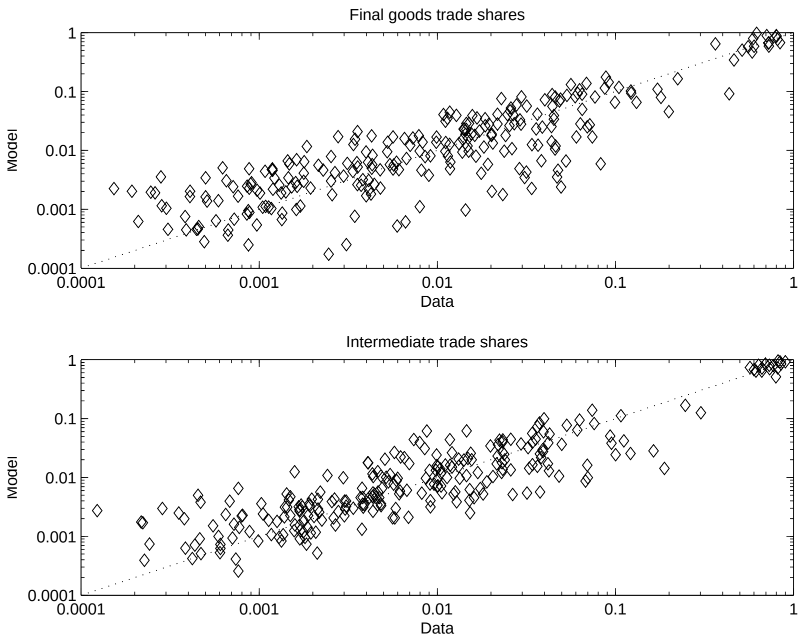


Figure 2: Aggregate Technology and Income per Capita



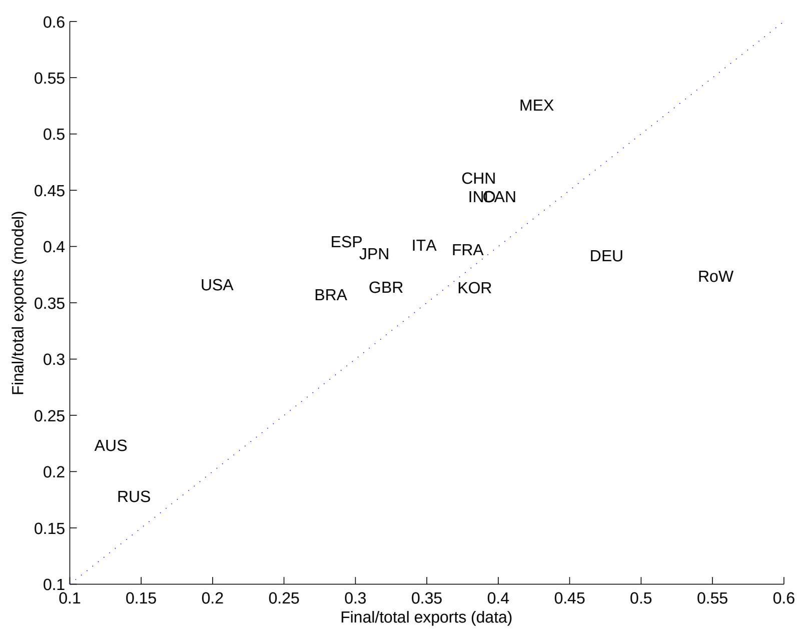
Average technology is an unweighted mean of the Stage 1 and Stage 2 technology estimates reported in Table 1. Income per capita is nominal expenditure at market exchange rates divided by population.

Figure 3: Trade Shares in Data and Model



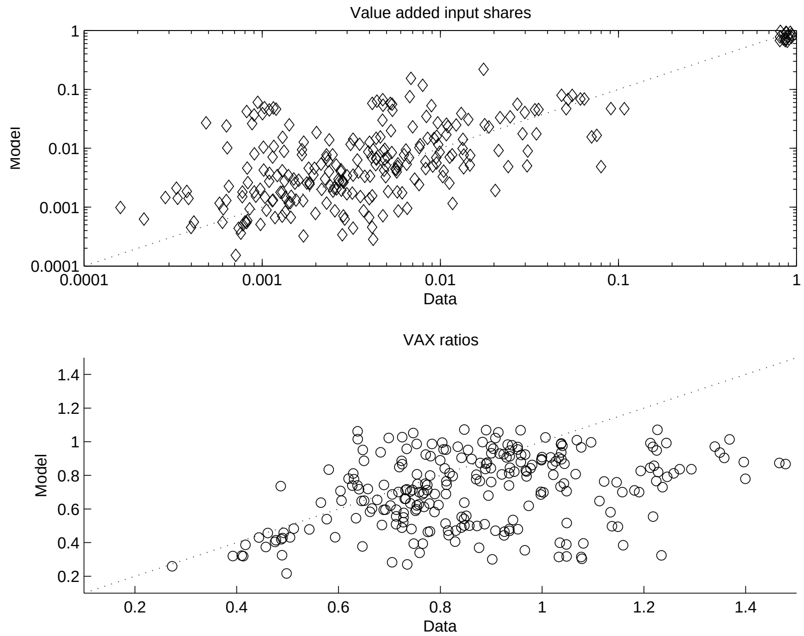
The dotted line in both figures is the 45-degree line. Axes in logs.

Figure 4: Final Relative to Total Exports in Data and Model



The dotted line in both figures is the 45-degree line.

Figure 5: Bilateral Value Added Trade in Data and Model



The dotted line in both figures is the 45-degree line. Top panel axes in logs.

Figure 6: The Gravity Distance Coefficient and the Level of Trade Costs

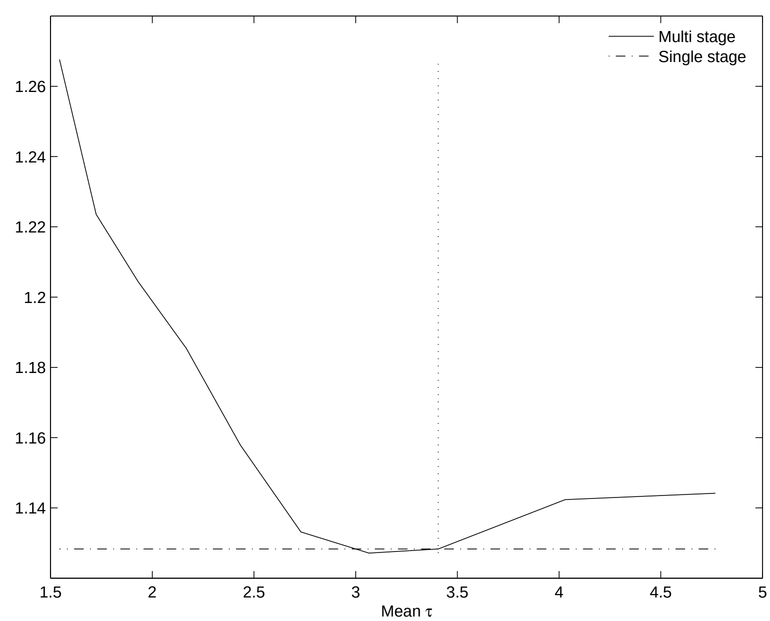


Figure 7: The Share Goods with Collocated Production Stages

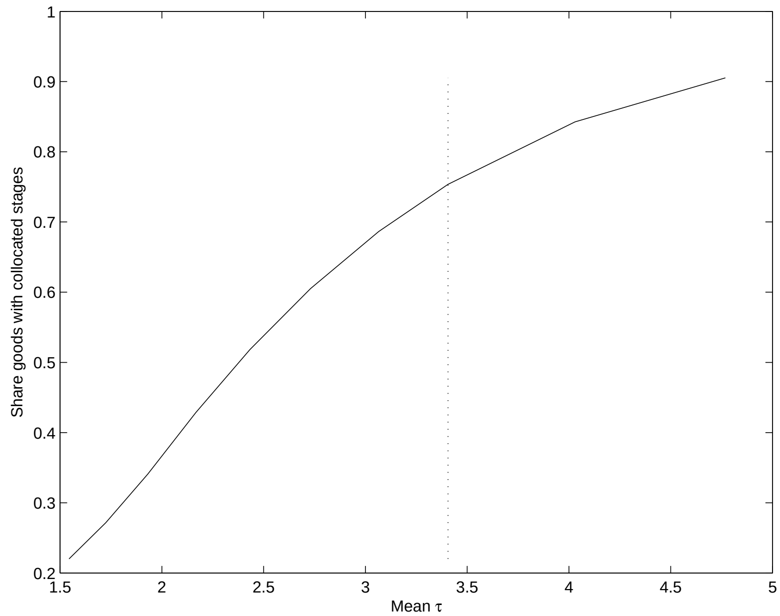


Figure 8: United States Bilateral Export Elasticities

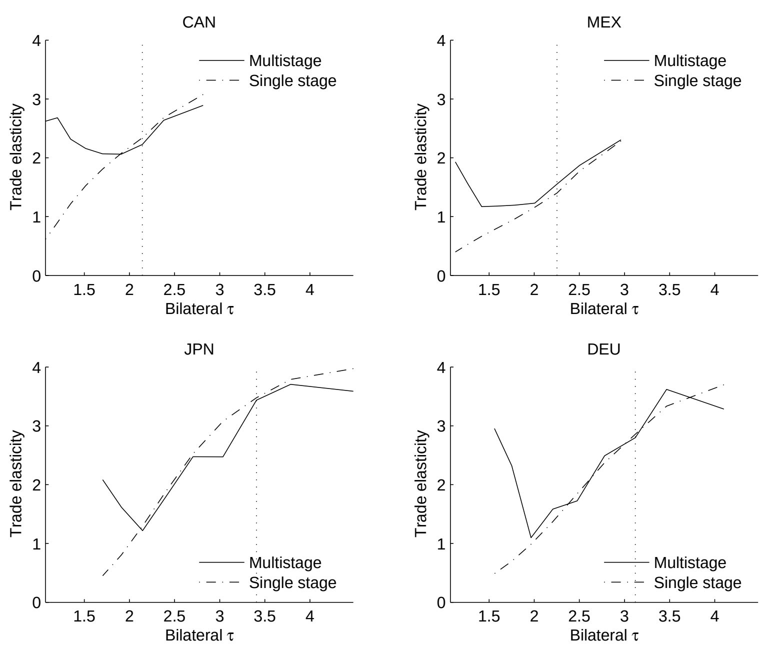


Figure 9: Gross Trade Responses to 10% Reduction in Trade Costs

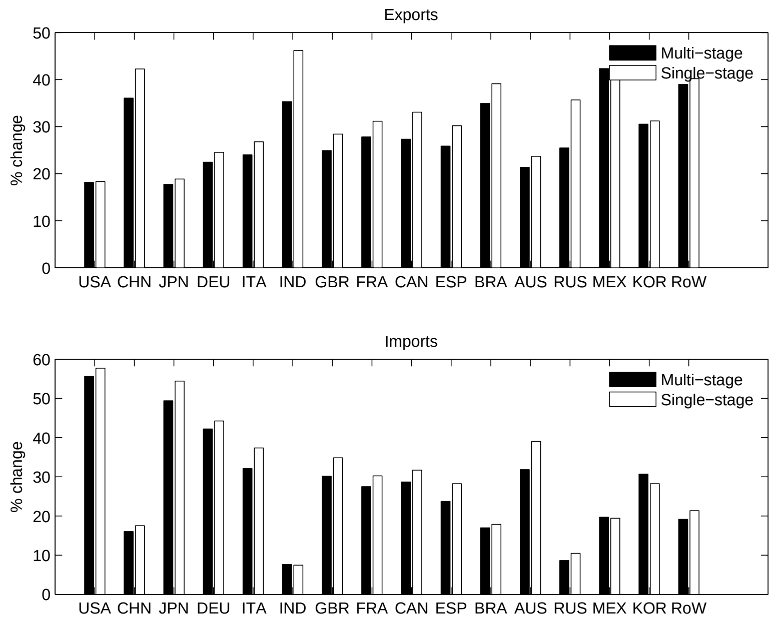


Figure 10: Value Added Trade Responses to 10% Reduction in Trade Costs in Multi-Stage Model

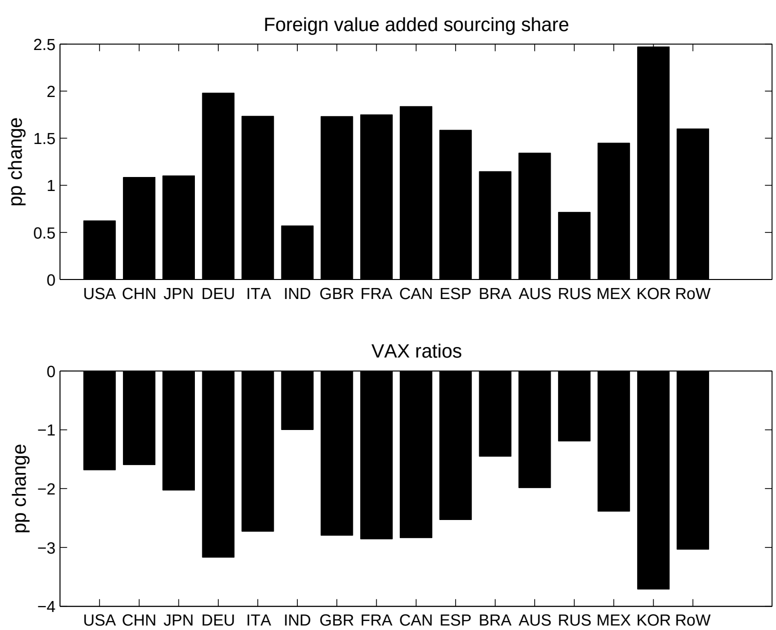


Figure 11: Value Added to Export Ratio Changes by Distance in Response to 10% Reduction in Trade Costs in Multi-Stage Model

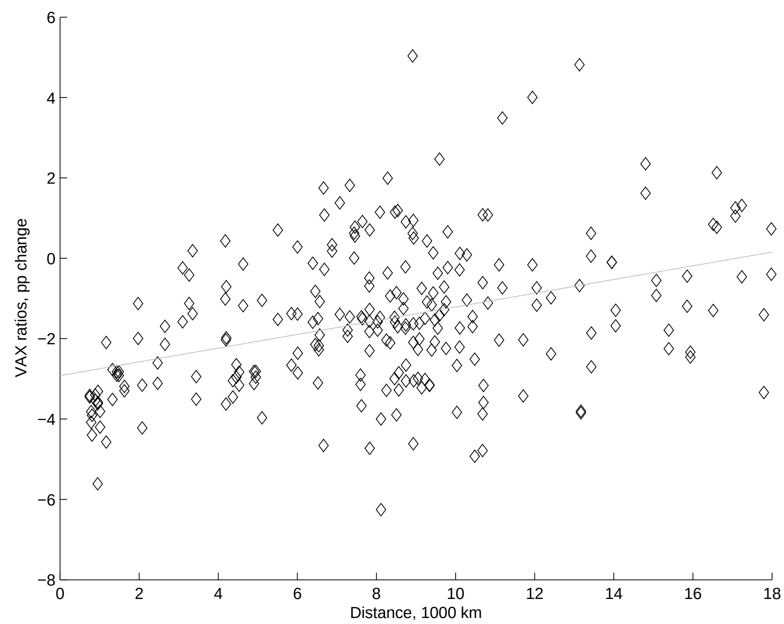


Figure 12: Real Wage Response to 10% Reduction in Trade Costs

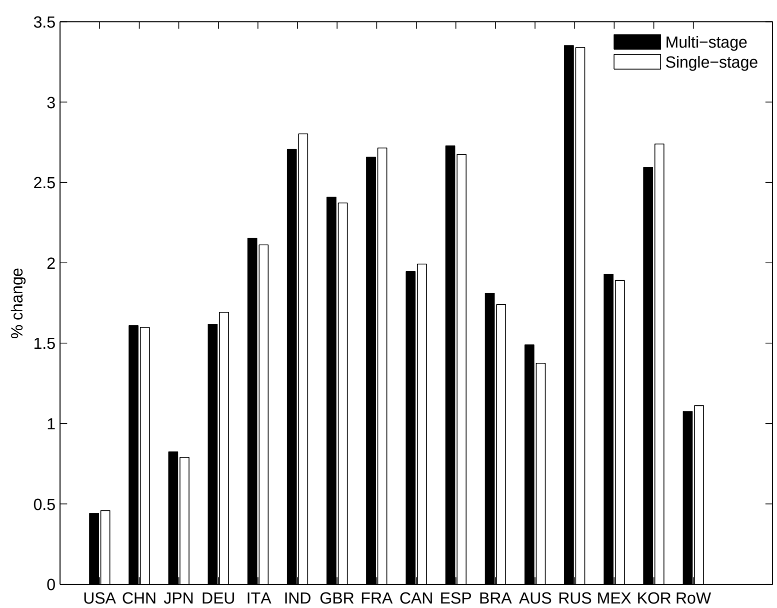


Figure 13: Value Added Trade Responses to Increase in Chinese Technology in Multi-Stage Model

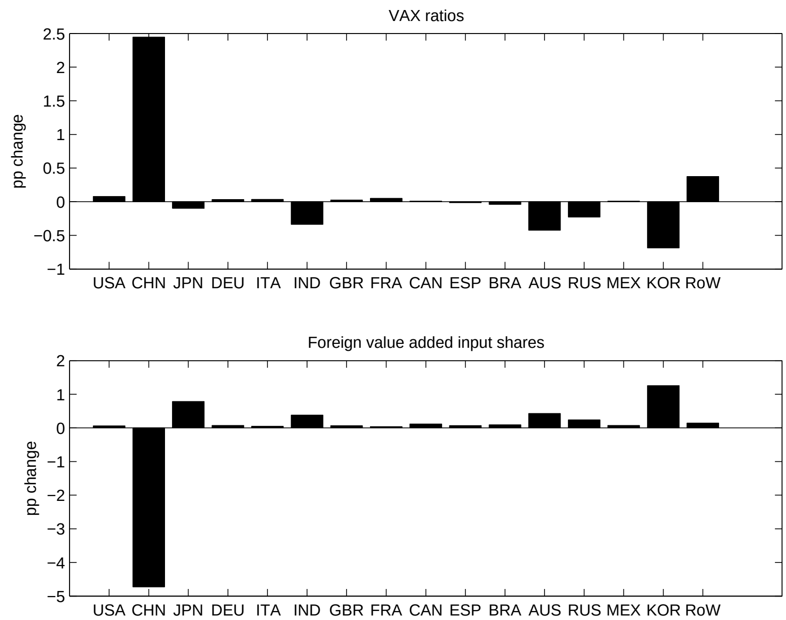
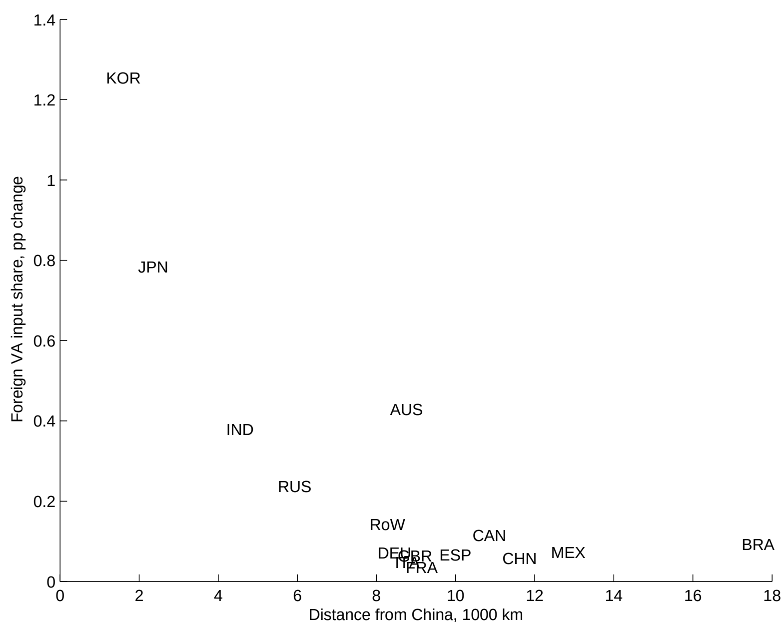


Figure 14: Changes in Foreign Value Added Inputs in Response to Increase in Chinese Technology in Multi-Stage Model



A Multi-Stage Input-Output Framework

In this appendix, we describe how we set up the model-based input-output accounting framework. To start, it is helpful to define notation here for values of bilateral shipments and gross output. Using the letter y to denote prices times delivered quantities, then we can write bilateral shipments as:

$$y_1^{ij}(r,m) \equiv \mathbb{I}\left(j \in \Omega_1^i(r,m)\right) \tilde{p}_1^j(r,m) x_1^j(r,m), \tag{22}$$

$$y_2^{ij}(r,m) \equiv \mathbb{I}\left(j \in \Omega_2^i(r,m)\right) \tilde{p}_2^j(r,m) \tilde{q}^j(r,m), \tag{23}$$

$$y^{ij}(r,n) \equiv \mathbb{I}\left(j \in \Omega_2^i(r,n)\right) \tilde{p}^j(r,n) \tilde{q}^j(r,n), \tag{24}$$

where the indicator functions $\mathbb{I}(\cdot)$ take the value 1 when country i is the low cost supplier to country j for a particular good/stage. Using this notation, we can then aggregate across goods and destinations as necessary. Denoting manufacturing stages by k, we can define total production by good $y_k^i(r,m)$, total bilateral shipments by sector $y_k^{ij}(m)$, and total production by sector $y_k^i(m)$ as:

$$y_k^i(r,m) \equiv \sum_j y_k^{ij}(r,m) \tag{25}$$

$$y_k^{ij}(m) \equiv \sum_r y_k^{ij}(r,m) \tag{26}$$

$$y_k^i(m) \equiv \sum_j \sum_r y_k^{ij}(r,m). \tag{27}$$

We define $y^i(r,n)$, $y^{ij}(s)$, and $y^i(n)$ analogously for non-manufacturing, dropping the stage subscripts.

Manufacturing We start with market clearing condition for stage 1 output in manufacturing. Note that bilateral shipments of stage 1 inputs can be written as:

$$\begin{aligned} y_1^{ij}(m) &= \sum_r \mathbb{I}\left(j \in \Omega_1^i(r,m)\right) \tilde{p}_1^j(r,m) x_1^j(r,m) \\ &= \sum_r \mathbb{I}\left(j \in \Omega_1^i(r,m)\right) \theta(m) y_2^j(r,m), \end{aligned} \tag{28}$$

where the second line uses the fact that $\tilde{p}_1^j(r,m)x_1^j(r,m) = \theta(m)y_2^j(r,m)$. Then the market clearing condition for stage 1 output can be written as:

$$\begin{aligned} y_1^i(m) &= \sum_j y_1^{ij}(m) \\ &= \sum_j \left[\frac{y_1^{ij}(m)}{y_2^j(m)}\right] y_2^j(m) \\ &= \sum_j \left[\theta(m)\sum_r \mathbb{I}\left(j \in \Omega_1^i(r,m)\right)\left(\frac{y_2^j(r,m)}{y_2^j(m)}\right)\right] y_2^j(m). \end{aligned} \tag{29}$$

In the second line, the ratio $\frac{y_1^{ij}(r,m)}{y_2^j(m)}$ records the share of stage 1 inputs from country i used by country j as a share of stage 2 output. The third line says that this ratio is equal to the Cobb-Douglas input cost share times the weighted count of goods in which country i is the low cost supplier of stage 1 inputs to country j, where the weights equal the share of country j's stage 2 production of each good in total stage 2 production in j.

Turning to stage 2 output, we need to divide this output across uses since it is both absorbed as a final good and used to form the composite input. We can break down output as follows:

$$\begin{aligned} y_2^{ij}(r,m) &= \mathbb{I}\left(j \in \Omega_2^i(r,m)\right)\tilde{p}_2^j(r,m)\tilde{q}^j(r,m) \\ &= \mathbb{I}\left(j \in \Omega_2^i(r,m)\right)\frac{P^j(m)Q^j(m)}{R} \\ &= \frac{\mathbb{I}\left(j \in \Omega_2^i(r,m)\right)}{R}\left[\alpha P_F^j F^j + \beta P_X^j X^j\right] \\ &= \frac{\mathbb{I}\left(j \in \Omega_2^i(r,m)\right)}{R}\left[\alpha P_F^j F^j + \beta\theta(m)y_1^j(m) + \beta\theta(n)y^j(n)\right], \end{aligned} \tag{30}$$

where the last line uses the fact that $P_X^j X^j = \theta(m)y_1^j(m) + \theta(n)y^j(n)$. The last line breaks down stage 2 shipments into final use and intermediate use by sector.

Then the full sector-level market clearing conditions for stage 2 output are given by:

$$\begin{aligned} y_2^i(m) &= \sum_j y_2^{ij}(m) \\ &= \sum_j\sum_r y_2^{ij}(r,m) \\ &= \sum_j \left[\frac{\sum_r \mathbb{I}\left(j \in \Omega_2^i(r,m)\right)}{R}\right]\left[\alpha P_F^j F^j + \beta\theta(m)y_1^j(m) + \beta\theta(n)y^j(n)\right]. \end{aligned} \tag{31}$$

The ratio $\left[\frac{\sum_r \mathbb{I}\left(j\in\Omega_2^i(r,m)\right)}{R}\right]$ is the fraction of stage 2 goods for which i is the low cost supplier to country j in sector m. For shorthand, we define $R^{ij}(m) \equiv \sum_r \mathbb{I}\left(j \in \Omega_2^i(r,m)\right)$, so then this fraction is given by: $\frac{R^{ij}(m)}{R}$.

Non-manufacturing For the non-manufacturing sector, bilateral shipments can be broken down as:

$$\begin{aligned} y^{ij}(r,n) &= \mathbb{I}\left(j \in \Omega^i(r,n)\right) \tilde{p}^j(r,n)\tilde{q}^j(r,n) \\ &= \mathbb{I}\left(j \in \Omega_2^i(r,n)\right) \frac{P^j(n)Q^j(n)}{R} \\ &= \frac{\mathbb{I}\left(j \in \Omega_2^i(r,n)\right)}{R} \left[(1-\alpha)P_F^j F^j + (1-\beta)P_X^j X^j\right] \\ &= \frac{\mathbb{I}\left(j \in \Omega_2^i(r,n)\right)}{R} \left[(1-\alpha)P_F^j F^j + (1-\beta)\theta(m)y_1^j(m) + (1-\beta)\theta(n)y^j(n)\right]. \end{aligned} \tag{32}$$

So the full market clearing condition for non-manufacturing is:

$$\begin{aligned} y^i(n) &= \sum_j y^{ij}(n) \\ &= \sum_j \sum_r y^{ij}(r,n) \\ &= \sum_j \left[\frac{\sum_r \mathbb{I}\left(j \in \Omega^i(r,n)\right)}{R}\right] \left[(1-\alpha)P_F^j F^j + (1-\beta)\theta(m)y_1^j(m) + (1-\beta)\theta(n)y^j(n)\right]. \end{aligned} \tag{33}$$

Using the same shorthand as above, we re-write the ratio $\left[\frac{\sum_r \mathbb{I}\left(j \in \Omega^i(r,n)\right)}{R}\right]$ as $\frac{R^{ij}(n)}{R}$.

Input-Output Table With these market clearing conditions, we can set up the input-output table. The component pieces are bilateral input use matrices A^{ij} and bilateral final goods shipments, which we will denote f^{ij}. The input use matrices have four rows/columns, corresponding to stages and sectors, and take the form:

$$\begin{aligned} A^{ij} &= \begin{bmatrix} 0 & \frac{y_1^{ij}(m)}{y_2^j(m)} & 0 \\ \frac{R_2^{ij}(m)}{R}\beta\theta(m) & 0 & \frac{R_2^{ij}(m)}{R}\beta\theta(n) \\ \frac{R_2^{ij}(n)}{R}(1-\beta)\theta(m) & 0 & \frac{R_2^{ij}(n)}{R}(1-\beta)\theta(n) \end{bmatrix}, \\ \text{with} \quad \frac{y_1^{ij}(m)}{y_2^j(m)} &= \theta(m) \sum_r \mathbb{I}\left(j \in \Omega_1^i(r,m)\right) \left(\frac{y_2^j(r,m)}{y_2^j(m)}\right). \end{aligned} \tag{34}$$

The ordering of rows/columns is (sector m, stage 1), (sector m, stage 2), and sector n.

These bilateral matrices can be arrayed to form the $4N \times 4N$ dimensional global input-output matrix:

$$A \equiv \begin{pmatrix} A^{11} & A^{12} & \dots & A^{1N} \\ A^{21} & A^{22} & \cdots & A^{2N} \\ \vdots & \vdots & \ddots & \vdots \\ A^{N1} & A^{N2} & \dots & A^{NN} \end{pmatrix} \tag{35}$$

Then we can organize the use of stage 2 goods as final goods in vector form as:

$$f^{ij} = \begin{bmatrix} 0 \\ \frac{R_2^{ij}(m)}{R} \alpha P_F^j F^j \\ \frac{R_2^{ij}(n)}{R} (1-\alpha) P_F^j F^j \end{bmatrix}. \tag{36}$$

And let F be the $4N \times N$ matrix of all f^{ij} vectors, where destinations j are arrayed along columns and source countries i are stacked vertically. And defining ι as a $N \times 1$ column of ones, note that $F\iota$ is the vector of final goods produced by source i.

Finally, let us assemble output by stage and sector into vectors:

$$y^i = \begin{bmatrix} y_1^i(m) \\ y_2^i(m) \\ y^i(n) \end{bmatrix}. \tag{37}$$

And let us stack these vertically to form an $4N \times 1$ dimensional vector Y.

Given this set-up, the standard input-output accounting identity holds: $Y = AY + F\iota$. We will use this input-output system to compute different model-based measures of trade in value added.

B Calibration of the Composite Input Aggregator

This appendix discusses our basis for calibrating the weight of manufacturing and non-manufacturing goods in the composite intermediate input, i.e., the parameter β. To pick a value for this parameter, we lay out an approach to choosing a value for β that is appropriate in a closed economy. We then use data for the world economy as a whole to calibrate β, which is by definition closed. Because we focus on a closed economy here, we suppress the country superscript on variables below.

Using notation similar to Section A, we can write total gross output in manufacturing as: $y(m) = y_1(m) + y_2(m)$. In the closed economy, all stage 2 goods are produced and use domestic stage 1 goods as inputs. Therefore, $y_1(m) = \theta(m)y_2(m)$. Using this fact, we re-write gross output as: $y(m) = (1+\theta(m))y_2(m)$. This links stage two output to observable sector level output $y(m)$ and a parameter $\theta(m)$ that can be measured from data. This is the first useful accounting identity.

The second useful accounting identity is the market clearing for stage 2 goods from sector 1: $y_2(m) = \alpha P_F F + \beta P_X X$ in the closed economy. We then recall that total purchases of the composite intermediate inputs are given by: $P_X X = \theta(m)y_1(m) + \theta(n)y(n)$. Combining these yields:

$$y_2(m) = \alpha P_F F + \beta\theta(m)y_1(m) + \beta\theta(n)y(n). \tag{38}$$

Recall that in our data, final purchases are observed at the sector level, so $\alpha P_F F$ is data. We can also link $y_2(m)$ and $y_1(m)$ to data on gross output at the sector level, as in the previous paragraph. Finally, gross output in the non-manufacturing sector $y(n)$ is also observable.

This leaves us with one equation in one unknown β:

$$(1+\theta(1))^{-1}y(m) = \alpha P_F F + \beta\theta(m)\left[\frac{\theta(m)}{1+\theta(m)}y(m) + \beta\theta(n)\right]y(n). \tag{39}$$

We pick β guided by this equation. To implement the calibration for the world, we aggregate all countries in our data to form the composite input-output table for the world, which records sector-to-sector sales of inputs, gross output, and final demand by sector. Using this data, we compute the sector-level input shares (i.e., $\{\theta(m), \theta(n)\}$) that are consistent with this world-level data, which happen to be nearly identical to cross-country median input shares. Plugging in these values along with values for final demand and gross output into the equation and solving yields a value of $\beta \approx 0.2$.

C Smoothing the Objective Function

In this section, we show how to calculate approximated simulated moments that are smooth in the parameter vector. This allows us to use standard gradient based optimization methods when minimizing the objective function.[42]

Our approach to simulating the trade shares borrows from the discrete choice literature, building on the observation that the trade shares are mathematically equivalent to choice probabilities. We use the logit-smoothed AR simulator to compute the trade shares (see McFadden (1989) and Train (2009)).

The first step in performing this transformation is to note that the indicator functions in the expressions for $Final^{ij}(m)$ and $Inter^{ij}(m)$ (see Section 3.2.2) can be re-written as statements about supply prices from alternative sources. Country j buys output for a particular stage from country i if i is the low cost supplier, which means that: $\mathbb{I}\left(j \in \Omega_1^i(r,m)\right) = \mathbb{I}\left(p_1^{ij}(r,m) < p_1^{kj}(r,m), \forall k \neq i\right)$ and $\mathbb{I}\left(j \in \Omega_2^i(r,m)\right) = \mathbb{I}\left(p_2^{ij}(r,m) < p_2^{kj}(r,m), \forall k \neq i\right)$. For example, the final trade shares can be written as: $Fshare^{ij} = \frac{1}{R}\sum_r \mathbb{I}\left(p_2^{ij}(r,m) < p_2^{kj}(r,m), \forall k \neq i\right)$. The second step then approximates the indicator function with the logit function, as in:

$$Fshare^{ij} = \frac{1}{R}\sum_r \frac{e^{-p_2^{ij}(r,m)/\lambda}}{\sum_k e^{-p_2^{kj}(r,m)/\lambda}}, \tag{40}$$

where λ is a scale factor.

Similarly, we can approximate the input trade shares as:

$$\begin{aligned} Inshare^{ij} &= \sum_r \mathbb{I}\left(j \in \Omega_1^i(r,m)\right)\left[\frac{\theta(s)y_2^j(r,m)}{In^j(m)}\right] + \frac{1}{R}\sum_r \mathbb{I}\left(j \in \Omega_2^i(r,m)\right)\left[\frac{\beta(m)P_M^j M^j}{In^j(m)}\right] \\ &= \sum_r \frac{e^{-p_1^{ij}(r,m)/\lambda}}{\sum_k e^{-p_1^{kj}(r,m)/\lambda}}\left[\frac{\theta(s)y_2^j(r,m)}{In^j(m)}\right] + \frac{e^{-p_2^{ij}(r,m)/\lambda}}{\sum_k e^{-p_2^{kj}(r,m)/\lambda}}\left[\frac{\beta(1)P_M^j M^j}{In^j(m)}\right]. \end{aligned} \tag{41}$$

[42] Gradient techniques are helpful to us, since the parameter space has relatively high dimensionality. We attempted to use non-gradient methods initially, but they generally performed poorly (i.e., were both slow and had difficulty finding the minimum).

The scale factor λ determines the degree of smoothing. As $\lambda \to 0$, the logit function converges to the indicator function and the smoothed trade shares approach the exact trade shares (in the discrete model). There is little guidance on the appropriate level of λ in general. By trial and error, we find that $\lambda = 0.02$ yields a very good approximation to the exact trade shares. Finally, we also need to choose R which yields an acceptable trade-off between simulation accuracy and computing time. In Monte Carlo simulations, we have found that the empirical model is able to recover the true parameters of the model when $R = 20,000$, so we use this value.

D Measuring Value Added in Trade

This appendix discusses how we compute value added trade in the data and model. As a preliminary step, we discuss how we assemble the data-based global input-output table.

Data-Based Global Input-Outut Table Using the data set described in Section 3.1, we define $Mshare^{ji}$ to be a 57×57 matrix with elements $\frac{x^{ji}(s)}{\sum_j x^{ji}(s)}$ along the diagonal. Then, for $i \neq j$, bilateral input-output matrices bilateral input-output matrices $\bar{A}^{ji}$ and final expenditure vectors $\bar{f}^{ji}$ are given by: $\bar{A}^{ji} = Mshare^{ji} A^{Ii}$ and $\bar{f}^{ji}(s) = Mshare^{ji} f^{Ii}$. Together with the domestic input-output matrices A^{ii} and domestic final goods sourcing f^{ii}, we then assemble the data as follows.

We stack the production data for all countries to form an $57N \times 1$ vector of gross output $\bar{Y}$, we combine the input output matrices to form a $57N \times 57N$ global input-output matrix $\bar{A}$, and we arrange the final expenditure vectors $\bar{f}^{ij}$ into $57N \times N$ matrix $\bar{F}$. For all these variables, we use the bar notation to make clear that these are data, since similar objects (without the bars) are defined for the model in Appendix A. With these definitions, we can write the input-output identity in the data as: $\bar{Y} = \bar{A}\bar{Y} + \bar{F}\iota$.

Value Added Inputs Starting with the data, we compute value-added inputs for final output from the goods sector as follows. We construct total final goods shipped from each country as $\bar{F}\iota$, and then reshape the resulting vector into corresponding 57×1 vectors of final goods shipped from each country, which we write $\bar{f}^i$. Zeroing out elements of these vectors corresponding to services sectors, we get modified vectors $\bar{f}^{i,goods}$. Then we arrange the collection of $\bar{f}^{i,goods}$ for all countries to form a $57N \times N$ block diagonal matrix $\bar{F}^{VAI}$, and compute foreign value added in final output in the goods sector as:

$$\overline{VAInputs} \equiv \bar{R}(I - \bar{A})^{-1} \bar{F}^{VAI}, \tag{42}$$

where $\bar{R}$ is a $N \times 57N$ block diagonal matrix with row vectors of value added to output ratios for each country along the diagonal.

To explain this calculation, note that $(I - \bar{A})^{-1} \bar{F}_{VAI}$ returns a $57N \times N$ matrix where column j is the vector of output needed to produce final goods shipped from j to all destinations. To compute value added embodied in those goods, we multiply by sector-level value added to output ratios and sum across sectors, where both operations are accomplished simultaneously via pre-multiplication by $\bar{R}$. The ij elements of the resulting matrix are the amount of value added from country i embodied in final goods produced in country j. For example, it measures the amount of Mexican value added in final goods produced in the United States. We construct value added inputs in the model in an identical way, but slide in model-based definitions for the input-output matrix, final goods production, and value added to output ratios.[43] We denote the resulting values $VAInputs$.

Value Added Exports The procedure for computing value added exports is similar to that above, except that final goods are distinguished according to the destination in which they are consumed. We express value added exports for the goods sector in matrix form as:

$$\overline{VAExports} = \bar{R}^{VAX}(I - \bar{A})^{-1}\bar{F}, \tag{43}$$

where here $\bar{R}^{VAX}$ takes the same form as $\bar{R}$, but replaces all value added to output ratios for services with zeros. The ij elements of $\overline{VAExports}$ record the amount of value added from the goods sector in country i that is absorbed in destination j, embodied in the final goods that j consumes. As above, we can construct the model-equivalent measures using an identical formula with values from the model-based input-output framework substituted for values from data, denoting resulting values $VAExports$.

E Benchmark Ricardian Trade Model with Input-Output Linkages

In this appendix, we describe the model that we use in Section 4 to evaluate how the multi-stage model differs from a benchmark two-sector Ricardian model. The benchmark model we use is a special case of the model in Caliendo and Parro (2012), which itself is based on Eaton and Kortum (2002).[44] Therefore, we refer to it at the EK/CP model. For brevity, we describe the key equilibrium conditions here, and refer the reader to Caliendo and Parro (2012) for a complete description of the underlying model.

To define the equilibrium concisely, we need to introduce some new notation. First, we define gross expenditure $E^i(s)$ to be total spending on final goods plus intermediates goods from sector s. Second, we define sector-level trade balances as $TB^i(s)$. Otherwise, the notation used here matches that use in the main text, with modification in the meaning of variables as necessary. For example, $c^i(s)$ denotes unit costs and $P^i(s)$ denotes an aggregate price level of an aggregate of sector s goods, but the functional forms are different here than in the main text reflecting differences between this model and the multi-stage model. For a

[43] One point to note is that value added to output ratios in the model are pinned down by parameters, equal to $1 - \theta(s)$ in each sector and common to all countries by assumption.

[44] It is a special case both in that we consider two sectors only, and in that we restrict the value of a number of parameters in ways that are consistent with our multi-stage model.

two sector economy, the equilibrium of the EK/CP model can be written concisely as:

$$c^i(m) = (w^i)^{1-\theta(m)} \left[P^i(m)^{\gamma^i(m)} P^i(n)^{1-\gamma^i(m)}\right]^{\theta(m)} \tag{44}$$

$$c^i(n) = (w^i)^{1-\theta(m)} \left[P^i(m)^{\gamma^i(n)} P^i(n)^{1-\gamma^i(n)}\right]^{\theta(n)} \tag{45}$$

$$P^i(m) = \left[\sum_j T^j(m) \left(c^j(m)\tau^{ji}(m)\right)^{-\bar{\kappa}}\right]^{-1/\bar{\kappa}(m)} \tag{46}$$

$$P^i(n) = \left[\sum_j T^j(n) \left(c^j(n)\tau^{ji}(n)\right)^{-\bar{\kappa}}\right]^{-1/\bar{\kappa}(n)} \tag{47}$$

$$\pi^{ij}(m) = T^i(m) \left[\frac{c^i(m)\tau^{ij}(m)}{P^j(m)}\right]^{-\bar{\kappa}(m)} \tag{48}$$

$$\pi^{ij}(n) = T^i(n) \left[\frac{c^i(n)\tau^{ij}(n)}{P^j(n)}\right]^{-\bar{\kappa}(n)} \tag{49}$$

$$E^i(m) = \sum_s \gamma^i(s)\theta(s) \left[E^i(s) + TB^i(s)\right] + \alpha P_F^i F^i \tag{50}$$

$$E^i(n) = \sum_s (1-\gamma^i(s))\theta(s) \left[E^i(s) + TB^i(s)\right] + (1-\alpha) P_F^i F^i \tag{51}$$

$$TB^i(m) = \sum_j \pi^{ij}(m)E^j(m) - E^i(m) \tag{52}$$

$$TB^i(n) = \sum_j \pi^{ij}(n)E^j(n) - E^i(n) \tag{53}$$

$$TB^i = TB^i(m) + TB^i(n) \tag{54}$$

$$P_F^i F^i = w^i L^i - TB^i, \tag{55}$$

There are several new parameters here. The parameter $\gamma^i(s)$ is a Cobb-Douglas input share, where $\gamma^i(s)$ is the share of total spending on intermediate inputs that sector s in country i dedicates to inputs from sector m. The parameter $\bar{\kappa}(s)$ is a sector-specific trade elasticity. We describe how we assign values to these parameters below. The remaining parameters $\theta(s)$ and α are defined as in the main text, and they they not vary across countries by assumption.

Following Dekle, Eaton, and Kortum (2008) and Caliendo and Parro (2012), the system of equations can be re-written in terms of changes relative to an initial equilibrium. To do this, we define variables $\hat{x} = \frac{x'}{x}$, where x' is the value of a variable in the new equilibrium and x is the value in the initial equilibrium. This yields equilibrium conditions:

$$\hat{c}^i(m) = (\hat{w}^i)^{1-\theta(m)} \left[\hat{P}^i(m)^{\gamma^i(m)} \hat{P}^i(n)^{1-\gamma^i(m)}\right]^{\theta(m)} \tag{56}$$

$$\hat{c}^i(n) = (\hat{w}^i)^{1-\theta(n)} \left[\hat{P}^i(m)^{\gamma(n)} \hat{P}^i(n)^{1-\gamma(n)}\right]^{\theta(n)} \tag{57}$$

$$\hat{P}^i(m) = \left[\sum_j \pi^{ji}(m)\hat{T}^j(m)\left(\hat{c}^j(m)\hat{\tau}^{ji}(m)\right)^{-\bar{\kappa}}\right]^{-1/\bar{\kappa}(m)} \tag{58}$$

$$\hat{P}^i(n) = \left[\sum_j \pi^{ji}(n)\hat{T}^j(n)\left(\hat{c}^j(n)\hat{\tau}^{ji}(n)\right)^{-\bar{\kappa}}\right]^{-1/\bar{\kappa}(n)} \tag{59}$$

$$\hat{\pi}^{ij}(m) = \hat{T}^i(m)\left[\frac{\hat{c}^i(m)\hat{\tau}^{ij}(m)}{\hat{P}^j(m)}\right]^{-\bar{\kappa}(m)} \tag{60}$$

$$\hat{\pi}^{ij}(n) = \hat{T}^i(n)\left[\frac{\hat{c}^i(n)\hat{\tau}^{ij}(n)}{\hat{P}^j(n)}\right]^{-\bar{\kappa}(n)} \tag{61}$$

$$E^i(m)\hat{E}^i(m) = \sum_s \gamma^i(s)\theta(s)\left[E^i(s)\hat{E}^i(s) + TB^i(s)\hat{TB}^i(s)\right] + \alpha P_F^i F^i \widehat{P_F^i F^i} \tag{62}$$

$$E^i(n)\hat{E}^i(n) = \sum_s (1-\gamma^i(s))\theta(s)\left[E^i(s)\hat{E}^i(s) + TB^i(s)\hat{TB}^i(s)\right] + (1-\alpha) P_F^i F^i \widehat{P_F^i F^i} \tag{63}$$

$$TB^i(m)\hat{TB}^i(m) = \sum_j \pi^{ij}(m)E^j(m)\hat{\pi}^{ij}(m)\hat{E}^j(m) - E^i(m)\hat{E}^i(m) \tag{64}$$

$$TB^i(n)\hat{TB}^i(n) = \sum_j \pi^{ij}(n)E^j(n)\hat{\pi}^{ij}(n)\hat{E}^j(n) - E^i(n)\hat{E}^i(n) \tag{65}$$

$$TB^i\hat{TB}^i = TB^i(m)\hat{TB}^i(m) + TB^i(n)\hat{TB}^i(n) \tag{66}$$

$$P_F^i F^i \widehat{P_F^i F^i} = w^i L^i \hat{w}^i \hat{L}^i - TB^i\hat{TB}^i. \tag{67}$$

In all simulations, we assume that labor input is fixed in all countries $\hat{L}^i = 1$. As in Caliendo and Parro, we treat changes in trade costs ($\hat{\tau}^{ij}(s)$), technology ($\hat{T}^j(s)$), and aggregate trade balances ($\hat{TB}^i$) as exogenous forcing variables. This leaves $10 + 2N^2$ endogenous variables $\{\hat{w}^i, \widehat{P_F^i F^i}, \{\hat{c}^i(s)\hat{P}^i(s), \hat{E}^i(s), \hat{TB}^i(s), \{\hat{\pi}^{ij}(s)\}_j\}_s\}_i$ and $10 + 2N^2$ equations, before choosing a normalization.[45]

To solve for these endogenous variables, we need parameters $\{\alpha, \{\bar{\kappa}(s), \gamma^i(s), \theta(s)\}_s\}$ and values for $\{w^i L^i, P_F^i F^i, TB^i, \{E^i(s), TB^i(s), \{\pi^{ij}(s)\}_j\}_s\}_i$ in an initial equilibrium. We set these parameters based on our the simulated data generated by the multi-stage model – i.e., we treat predicted equilibrium values from our estimated model as 'data', which implies that we start simulations from the same equilibrium in our multi-stage model and this

[45]Note that we treat nominal final expenditure $P_F^i F^i$ as one variable, hence the wide-hat notation on $\widehat{P_F^i F^i}$. We do not need to separate the final price level and real final expenditure to compute the counterfactuals that interest us.

alternative benchmark model. Recall that we estimated the benchmark model allowing for unbalanced trade. To perform counterfactuals, we exogenously close trade balances in the multi-stage model, and then evaluate counterfactuals relative to this balanced trade equilibrium. This implies that $P_F^i F^i = w^i L^i$ and $TB^i = 0$ in our initial equilibium. So we need only need $\{P_F^i F^i, \{E^i(s), TB^i(s), \{\pi^{ij}(s)\}_j\}_s\}_i$ in the balanced trade equilibirum of the multi-stage model, plus the structural parameters $\{\alpha, \{\bar{\kappa}(s), \gamma^i(s), \theta(s)\}_s\}$, to compute changes in equilibrium variables in the EK/CP model. We now describe the details regarding how we obtain values for these parameters.

Reading values for $\{P_F^i F^i, \{E^i(s), TB^i(s), \{\pi^{ij}(s)\}_j\}_s\}_i$ from our simulated data is completely straightforward. Further, $\theta(s), \alpha\}$ are set to the same values as in the multi-stage model. The parameter $\gamma^i(m)$ for sector m (the multi-stage sector) is equal to the value of inputs from sector m used by sector m as a share of total input use by sector m. Total input use by sector m is equal to use of stage 1 inputs by stage 2, which are equal to $\theta(m)y_2^i(m)$, plus use of the composite input, which is equal to $\theta(m)y_1^i(m)$.[46] Then all stage 1 inputs used by stage 2 in sector m originate from sector m, but only a fraction (β) of the composite input originates from sector m. This implies that:

$$\gamma^i(m) = \frac{\theta(m)y_2^i(m) + \beta\theta(m)y_1^i(m)}{\theta(m)y_2^i(m) + \theta(m)y_1^i(m)}$$

The value of this parameter varies across countries to the extent that the mix of stage 1 versus stage 2 output varies across countries. Turning to $\gamma^i(n)$, sector n uses sector m inputs only embodied in the composite input, and the composite input itself is the only input in production. Therefore, $\gamma^i(n) = \frac{\beta\theta(n)y^2(n)}{\theta(n)y^2(n)} = \beta$, where $\theta(n)y^2(n)$ is equal to the value of the composite input used by sector n. Note that this parameter then does not vary across countries in our model.

Finally, turning to the values for $\bar{\kappa}(s)$, we note that these values correspond to the elasticity of log bilateral trade to log bilateral trade costs in this EK/CP model. Therefore, we obtain them by regressing simulated bilateral trade in each sector from the multi-stage model on estimated log bilateral trade costs. For reference, this returns estimates $\bar{\kappa}(m) = 3.85$ and $\bar{\kappa}(s) = 4.08$.

[46] The notation here is defined in Appendix A.