Australian School of Business Research Paper No. 2014 ECON 36

The Composition of Trade Flows and the Aggregate Effects of Trade Barriers

Scott French

The Composition of Trade Flows and the Aggregate Effects of Trade Barriers

Scott French*

August, 2014

Abstract

A widely used class of quantitative trade models implicitly assumes that patterns of comparative advantage take a specific form such that they have no influence over the effect of trade barriers on aggregate trade flows and welfare. In this paper, I show that this assumption is inconsistent with patterns present in the product-level trade data and develop a framework in which to analyze the role of interactions among countries' patterns of comparative advantage in determining the aggregate effects of trade barriers. The model preserves much of the tractability of standard aggregate quantitative trade models while allowing for the effects of any pattern of comparative advantage, across many products and countries, to be taken into account. After fitting the model to product-level trade data, I find that the composition of trade flows is quantitatively important in determining the welfare gains from trade and the aggregate effects of changes in trade barriers. A key finding is that the welfare gains from trade tend to be larger and more skewed toward low-income countries than an aggregate model would suggest.

JEL Classification: F11, F14, F17, F62, O19

Keywords: International Trade, Composition, Comparative Advantage, Trade Barriers, Welfare, Gravity, Income Differences, Elasticity of Substitution

---

*School of Economics, University of New South Wales. scott.french@unsw.edu.au. I am thankful to Dean Corbae for his guidance and support. I am also thankful for advice and comments from James Anderson, Arghya Ghosh, Richard Holden, Russell Hillberry, James Markusen, James Morley, Kim Ruhl, and Alan Woodland as well as seminar participants at the Australian National University, George Washington University, Monash University, the University of Melbourne, the University of New South Wales, the University of Oregon, the University of Sydney, and the University of Wyoming. All errors are my own. This paper was previously circulated under the title "The Composition of Exports and Gravity"

1 Introduction

When seeking quantitative predictions of the effects of barriers to international trade, researchers often turn to models in which trade barriers and their effects can be inferred from data on aggregate bilateral trade flows. Some of these models feature rich micro-level market structures, and all of them have the advantageous feature that the amount of data required to make predictions regarding aggregate variables – such as income, welfare, and trade flows – is quite low. In fact, Arkolakis et al. (2012) have recently shown that, for a large class of such models, the welfare gains from trade are a function of only aggregate trade flows and the elasticity of trade flows with respect to trade costs, regardless of the underlying micro-level structure of the model.[1] However, the restrictions of these models which make them so analytically tractable and conducive to quantitative analysis also implicitly assume either that there is no trade arising from comparative advantage across products or that countries' patterns of comparative advantage take a very special form, both of which imply that the effect of trade barriers on aggregate trade flows is independent of the composition of those trade flows.

In this paper, I relax such restrictions, developing a framework with the flexibility to allow for arbitrary patterns of comparative advantage across products for every country. These patterns can interact in non-trivial ways to influence the effects of trade barriers on aggregate bilateral trade flows and welfare. My model nests the aggregate quantitative trade models that fall into the class described by Arkolakis et al. (2012) and makes clear that only in very particular cases can countries' patterns of comparative advantage be ignored when predicting the effects of trade barriers, even when one is only interested in the effects on aggregate variables.[2] Using data on bilateral trade flows at the product level, I show that product-level trade flow patterns are inconsistent with these cases, that interactions among countries' patterns of comparative advantage are quantitatively important in determining aggregate bilateral trade flows, and that taking these patterns of comparative advantage into account is important for predicting the welfare effects of trade barriers.[3] For example, I find that 25% of the variation in bilateral trade flows arises from interactions among countries' patterns of comparative advantage. I also find that these patterns of comparative advantage account for about two-thirds of the gains from trade relative to autarky and that a model which ignores these patterns predicts only half of the gains from trade for the average country and a much smaller percentage for developing countries.

In Section 2, I outline the basic theoretical framework for this analysis. The point of departure is the model of Alvarez and Lucas (2007), which is based on the Ricardian trade model of Eaton and Kortum (2002) – henceforth, the EK model – in which international trade occurs due to idiosyncratic

[1] These models include the model of monopolistic competition and increasing returns to scale of Krugman (1980), the Armington model of Anderson and van Wincoop (2003), the Ricardian trade model of Eaton and Kortum (2002), and models of heterogeneous firms á la Melitz (2003), such as Chaney (2008).

[2] I refer to the models described by Arkolakis et al. (2012) as "aggregate quantitative trade models" because they imply that only aggregate data is required to make quantitative predictions regarding the aggregate effects of trade barriers, regardless of their micro-level structure.

[3] In particular, I use data from the UN Comtrade database at 6-digit level of Harmonized System, which consists of over 6,000 products, including more than 4,500 manufactured goods.

differences in productivity across product varieties. I assume that the ex-ante expected productivity of varieties in a given country varies across product-categories into which varieties are grouped, in contrast with the EK model, in which every variety is ex-ante identical. This setup allows much of the analytical tractability of the EK model to be maintained, while also allowing for any arbitrary pattern of comparative advantage across products to be incorporated into the model.

The key theoretical result of the paper is that aggregate bilateral trade flows follow an aggregate gravity-type equation that is nearly identical to that of an aggregate quantitative trade model. The only difference is that there is an additional bilateral term which summarizes the strength of the exporter's comparative advantage across products relative to the importer's ability to obtain each product from anywhere in the world. In the global economy, the latter is a function of every country's patterns of relative productivity across products and the barriers to trade between those countries and the importer. Thus, the model provides a succinct way to summarize and quantify the strength of a basic Ricardian force that is absent from a typical aggregate quantitative trade model – i.e., country i will export relatively more to country n if country i is relatively productive for goods that country n cannot purchase cheaply from other sources (including domestic producers in n).

The second important insight from the model is that this bilateral term summarizing the strength of comparative advantage for an exporter in a given market, when the destination market is the domestic market, also serves as a measure of the exporter's welfare gains from trade that arise from comparative advantage across products. This is because, if the products that a country can purchase relatively cheaply from abroad are those for which it is relatively unproductive, then for a given level of international trade flows, this country benefits relatively more from specialization according to comparative advantage.

To understand the relationship between this model and those that belong to the class of aggregate models often used for quantitative analysis, I detail the special cases in which this model reduces to one of these aggregate models. When relative productivity across products is equal for all countries, then there is no across-product comparative advantage, and trade arises solely due to idiosyncratic productivity differences across producers of given varieties. In this case, the model is consistent with the EK model and a version of the heterogeneous-firm model of Chaney (2008), in which the effects of trade barriers depend only on the level of dispersion in the distribution of idiosyncratic productivity.[4] In the opposite limiting case, in which relative productivity differences are so extreme that each country produces a unique set of products, the model is consistent the homogeneous-firm monopolistic competition model of Krugman (1980) and the Armington model of Anderson and van Wincoop (2003), in which the effects of trade barriers depend only on the elasticity of substitution across products. In general, however, the model lies between these two extremes, and the effects of trade barriers vary across each country pair, depending on the particular patterns of comparative that exist in the world economy.

[4]In the case of Chaney (2008), this is precisely true only if firm entry is exogenous. With endogenous entry, fixed trade costs affect aggregate trade flows by changing the equilibrium number of firms, and the magnitude of this effect is inversely related to the elasticity of substitution across product varieties.

In Section 3, because the strength of across-product comparative advantage cannot be inferred from aggregate trade flow data alone, I use product-level trade flow data to assess whether any of the special cases implicitly assumed by aggregate quantitative trade models accurately describes the world. To this end, I develop a bilateral index based on the expression for the elasticity of imports with respect to trade costs. Under the assumptions of an aggregate quantitative trade model, this index is constant across country pairs. However, I show that, in the data, there is a great deal of heterogeneity, with the exports of developing countries typically facing a higher trade cost elasticity than those of developed countries. This suggests that the patterns of across-product comparative advantage that exist in the data can play an important role in determining the effects of trade barriers, especially in their effects on the cross-country income distribution.

In Section 4, I use product-level trade data to estimate trade barriers and infer countries' patterns of productivity across products. Then, I compare the model's predictions with those of a version restricted to be consistent with an EK model, whose predictions depend only on aggregate data. I find that both models perform equally well in matching the cross-country income distribution, while the product-level model does a better job of predicting the cross-country relationship between the price of tradeable goods and real GDP per worker, and the product-level model goes quite far, compared to the aggregate model, in accounting for patterns of product-level trade flows.

Given the model's ability to predict salient features of the aggregate and product-level data, I go on to quantitatively assess the welfare effects of trade barriers using the product-level model and compare these predictions to those of the aggregate model. I do this by comparing the models' predictions of the effects on welfare of moving from the baseline case to complete autarky and of eliminating border-related trade costs. These counterfactual predictions are presented in Section 5. For both scenarios, the gains from falling trade barriers are larger and more skewed toward low-income countries than implied by the aggregate model. In the first case, this is because the strength of comparative advantage is found to be relative strong in the developing countries' domestic markets, which implies that these countries benefit relatively more from specialization across products in the baseline case. In the second case, this is also because the estimates of border costs based on the product-level model are higher on average and relatively higher for developing countries than those based on the aggregate model. This is the case because the product-level model infers from the data that the forces of comparative advantage are generally stronger, especially for developing countries, relative the aggregate model. Thus, higher trade costs are required to match the level of trade flows observed in the data. Such results indicate the importance of accounting for the patterns of comparative advantage that exist in the data when seeking to make quantitative predictions regarding the effects of trade barriers.

This paper is primarily related to the very large literature which uses quantitative trade models to determine the effects of trade barriers on aggregate variables, especially aggregate income and welfare, including Eaton and Kortum (2002); Anderson and van Wincoop (2003); Alvarez and Lucas (2007); and Helpman et al. (2008) and recent papers, such as Waugh (2010) and Fieler (2011), that resolve discrepancies between more traditional quantitative trade models and the data.[5] Costinot and Rodriguez-Clare (2013) review recent advances in this literature in measuring the welfare gains from international trade relative to autarky. The main contribution of my paper is that it demonstrates how non-trivial patterns of comparative advantage across products can be incorporated into the class of trade models most commonly used to address quantitative questions related to international trade. It does so in a way that maintains, to a large extent, the tractability and parsimony of the standard models while utilizing the wealth of information contained in product-level trade data that is available for most of the world's countries. It also provides a succinct and intuitive way to express the gains from trade due to countries' patterns of comparative advantage across products and to decompose the gains from trade into across-product and within-product components.

This is not the first paper to evaluate the effects of trade barriers using disaggregated trade data. For instance, Anderson and Yotov (2011), Chen and Novy (2011), Costinot et al. (2012), Caliendo and Parro (2012) and Levchenko and Zhang (2013) take into account sectoral heterogeneity at the industry level in predicting the effects of trade barriers. My paper differs from such studies in two ways. First, it derives succinct and intuitive expressions that demonstrate how patterns of comparative advantage influence the effect of trade barriers on aggregate trade flows and welfare. Second, it develops and implements a framework which takes advantage of the product-level trade data to fully take into account the role of potentially complex patterns of comparative advantage across thousands of products and more than 150 countries. By contrast, industry-level analyses that allow for heterogeneity across, at the very most, a few dozen industries can only take into account broad structural differences across countries. Thus, they ignore the effects of patterns of comparative across products, within each industry, on aggregate trade flows and welfare.

In Section 6, to evaluate the way in which the predictions based on this framework differ from those of industry-level analyses, I also consider a multi-sector extension to the baseline model in which trade costs are allowed to vary by sector. The basic result is that, for the counterfactual experiments considered in this paper, accounting for across-product comparative advantage has much the same implications in the multi-sector version of the model as in the single-sector version, and if anything the effects are quantitatively larger in the multi-sector specification. Thus, even in models that feature sectoral heterogeneity, it is important to account for product-level patterns of comparative advantage. In addition, because the data requirements associated with quantitatively modelling sectoral heterogeneity can significantly limit the sample size of a study, and because product-level trade data is available for a much wider range of countries, the framework of this paper can be applied to take into account at least one important form of heterogeneity even for sets of countries for which other forms of disaggregated data are unavailable. The analysis of Section 6 also demonstrates how the effects of comparative advantage across products, within sectors, can be incorporated into such multi-sector analyses without requiring any additional restrictions on the modelling framework or restricting the sample of countries and sectors that can be studied.

[5] Anderson and van Wincoop (2004) provide a survey of older papers that have extended theoretically-founded gravity models, such as Anderson (1979) and Krugman (1980), in a number of dimensions.

This paper is also related to Arkolakis et al. (2012) in that both papers address important features shared among the literature's workhorse class of quantitative trade models, but the papers make very distinct points. Arkolakis et al. (2012) demonstrate that, in this class of models, the welfare gains from trade depend only on two aggregate variables. This paper demonstrates that welfare in these models depends only on aggregate variables because of particular assumptions that imply no role for patterns of comparative advantage across products in influencing the welfare effects of trade barriers. It then goes on to show that, when the patterns that exist in the data are taken in to account using a more general framework, the role of such patterns is quantitatively important.

2 Model

The world economy is made up of N countries, indexed by n and i. Each country consists of a continuum of households with mass L_i which is each endowed with k_i units of capital and one unit of labor, which are supplied inelastically. Each household obtains utility from the consumption of a non-tradeable final good.

2.1 Production

The production structure in this model is designed to allow for a straightforward mapping between the model and disaggregated trade flow statistics. For simplicity and comparability with the large set of models considered by Arkolakis et al. (2012), I consider a model economy with a single tradeable goods sector. A multiple sector extension is considered in Section 6.2. The tradeable goods sector is comprised of a finite number of products, $k = 1, ..., K$, which is each made up of a continuum of unique varieties, $\omega \in [0, 1]$. Thus, a given variety is identified by the pair (k, ω). Because a product in the model is mapped into a 6-digit Harmonized System product category in the international trade data, I use the terms "product" and "product category" interchangeably.

The non-tradeable final good is produced by a representative firm using capital, labor, and a composite tradeable good according to the following Cobb-Douglas production function:

$$Y_n = \left((L_i^f)^\alpha (K_i^f)^{1-\alpha}\right)^\gamma \left(Q_n^f\right)^{1-\gamma}.$$

The composite tradeable good is a CES aggregate of products, given by

$$Q_n = \left(\sum_{k=1}^{K} (q_n^k)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}},$$

where $\sigma > 1$ is the elasticity of substitution across products, and each product is a CES aggregate if its component varieties, given by

$$q_n^k = \left(\int_0^1 q_n^k(\omega)^{\frac{\eta^k - 1}{\eta^k}} d\omega \right)^{\frac{\eta^k}{\eta^k - 1}},$$

where $\eta^k > 1$ is the elasticity of substitution across varieties of product k.[6]

2.2 International Trade

Individual varieties can, in principle, be produced in any country and can be shipped anywhere in the world. International shipments, however, are subject to trade barriers, which are assumed to take the convenient "iceberg" form, as in Samuelson (1954), meaning that delivering 1 unit of a good from i to n requires shipping $d_{ni} > 1$ units. Note that, as in other one-sector quantitative trade models, bilateral trade costs are constant across products. This assumption greatly simplifies the analytical expressions that follow and allows for comparisons with these models. In the multiple-sector extension in Section 6.2, trade costs are allowed to vary across sectors, as in Caliendo and Parro (2012) and Levchenko and Zhang (2013).[7]

Each variety is produced using a Cobb-Douglas combination of capital, labor, and the composite tradeable good. Denoting the efficiency of producers in country i in producing (k, ω) as $Z_i^k(\omega)$, the cost of delivering one unit of the variety from i to n is given by

$$c_{ni}^k(\omega) = \frac{\tilde{\alpha}(w_i^\alpha r_i^{1-\alpha})^\beta P_i^{1-\beta} d_{ni}}{Z_i^k(\omega)} \equiv \frac{c_i d_{ni}}{Z_i^k(\omega)}, \tag{1}$$

where w_i is the wage in i, r_i is the rental rate of capital, and P_i is the price of the composite tradeable good.[8]

I assume that $Z_i^k(\omega)$ consists of a deterministic component, which reflects the overall state of technology in i for producing all varieties of k, and a stochastic component, which allows for idiosyncratic differences in productivity across varieties. Specifically, I assume that

$$\ln\left(Z_i^k(\omega)\right) = \ln\left(Z_i^k\right) + \varepsilon_i^k(\omega).$$

Assuming that $\varepsilon_i^k(\omega)$ is distributed Gumbel (type-I extreme value), as in McFadden (1974), with shape parameter θ, implies that $Z_i^k(\omega)$ follows a Fréchet distribution, as in an EK model, with a CDF given by

$$F_i^k(z) = e^{-T_i^k(\hat{\gamma}^k z)^{-\theta}}, \tag{2}$$

[6] The assumption that a continuum of varieties exists for each product, which is admittedly strong, is purely for analytical convenience. With a finite number of varieties, all the results that follow hold in expectation.

[7] Since this paper is focused on the aggregate effects of trade barriers, this assumption is innocuous if both (a) the distribution of trade costs across products is not systematically related to countries' patterns of comparative advantage and (b) product-specific components of trade costs, which are relegated to the error term in the estimation, are uncorrelated with the gravity variables used to proxy for bilateral trade costs.

[8] The constant $\tilde{\alpha} = (\alpha^\alpha(1-\alpha)^{1-\alpha}\beta)^{-\beta}(1-\beta)^{\beta-1}$.

where $T_i^k = (Z_i^k)^\theta$, and $\theta > 1$ is inversely related to the dispersion of efficiency across varieties.[9] The average level of T_i^k over all products reflects country i's absolute advantage, variation in T_i^k across products reflects i's deterministic pattern of comparative advantage, and θ governs the degree of comparative advantage across varieties of a each product due to idiosyncratic productivity differences.

2.2.1 Product-Level Trade Flows

Perfect competition implies that the price of variety (k, ω) charged in n by producers from i is equal to $c_{ni}^k(\omega)$, and buyers in n will purchase the variety from the source that offers the lowest price, so the realized price of the variety in n is

$$p_n^k(\omega) = \min_i \{c_{ni}^k(\omega)\}.$$

Following the analysis of Eaton and Kortum (2002), the probability that country i is the low-cost provider of a given variety to n is given by

$$\pi_{ni}^k = \frac{T_i^k (c_i d_{ni})^{-\theta}}{\Phi_n^k}, \tag{3}$$

where

$$\Phi_n^k = \sum_{i=1}^N T_i^k (c_i d_{ni})^{-\theta}. \tag{4}$$

The distribution of the realized price of a variety in n is independent of the source that provides it at the lowest cost, so π_{ni}^k is also the share of expenditure by n on varieties of product k that come from i. The parameter Φ_n^k features heavily in the analysis that follows. It summarizes the state of technology for producing product k in every country in the world from the perspective of an importer in n. It is higher if n has access – through low trade barriers – to sources with high values of T_i^k and low input costs.

Equation (3) also highlights the role of θ in the determination of trade flows. Note that (3) can be rewritten as

$$\pi_{ni}^k = \frac{(c_i d_{ni}/Z_i^k)^{-\theta}}{\sum_{i=1}^N (c_i d_{ni}/Z_i^k)^{-\theta}}.$$

Thus, θ governs the importance of the deterministic component of costs in determining the allocation of expenditure by n across potential sources of product k. The larger is the value of θ – i.e. the smaller the degree of dispersion in the distribution of productivity across varieties – the higher the probability that country n buys a given variety of k from the country with the highest deterministic

[9] The CDF of the Gumbel distribution is $F(\varepsilon) = e^{-e^{-\theta\varepsilon}}$. The constant $\hat{\gamma}^k = \Gamma(1-(\eta^k-1)/\theta)^{-\frac{1}{\eta^k-1}}$, where $\Gamma(\cdot)$ is the gamma function. This constant is included in (2) purely for notational convenience, as it eliminates constants in the expressions for price indexes and relative expenditure across products that would appear otherwise. The only role that $\hat{\gamma}^k$ plays is in the mapping between relative productivity across products and relative sales, and it is irrelevant to the analysis of this paper. Note that $\hat{\gamma}^k$ is unrelated to the parameter, γ, in the final goods production function.

productivity and lowest production and trade costs. In terms of total trade flows within a product category, θ governs the sensitivity of the share imports by n to these variables. That is, θ serves as the elasticity of substitution in demand for a given product across source countries.

2.2.2 The Allocation of Expenditure

Cost minimization by the final goods producer implies that expenditure on a particular product is given by

$$X_n^k = \left(\frac{P_n^k}{P_n}\right)^{1-\sigma} X_n, \tag{5}$$

where $X_n \equiv P_n Q_n$ is total expenditure on tradeable goods, and the price indexes are given by

$$P_n^k = \left(\int_0^1 p_n^k(\omega)^{1-\eta^k} d\omega\right)^{\frac{1}{1-\eta^k}} = (\Phi_n^k)^{-\frac{1}{\theta}} \tag{6}$$

and

$$P_n = \left(\sum_{k=1}^{K} \left(P_n^k\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}} = \Phi_n^{-\frac{1}{\theta}}, \tag{7}$$

where $\Phi_n = \left(\sum_k (\Phi_n^k)^{\frac{\sigma-1}{\theta}}\right)^{\frac{\theta}{\sigma-1}}$.

Equations (5) - (7) imply that that the share of n's expenditure that is devoted to product k is given by $X_n^k/X_n = (\Phi_n^k/\Phi_n)^{(\sigma-1)/\theta}$. That is, relative expenditure on k depends on the level of the technology index, Φ_n^k, with an elasticity given by the ratio $(\sigma-1)/\theta$. In the same way that θ serves as the elasticity of π_{ni}^k with respect to i's deterministic productivity and costs, it also governs the elasticity of P_n^k with respect to Φ_n^k, which is an index of these values across the set of potential exporters, as is clear from (4). Thus, a greater value of θ implies a smaller response in the price index of k to a change in technology somewhere in the world, and a larger value of σ implies a greater response in expenditure on k to a given change in the price index, so it is the relative value of the two that determines the responsiveness of expenditure to the state of technology.

2.2.3 Aggregate Trade Flows

Combining equations (3) - (7) yields an expression relating aggregate bilateral trade flows to production costs, trade costs, and states of technology in every country, which is given in the proposition that follows.

Proposition 1. *Given that productivity is distributed according to (2), production and transport costs are given by (1), and demand for each product in each destination is given by (5), then the aggregate share of total expenditure on tradeable goods by n that originated in i can be expressed as*

$$\pi_{ni} \equiv \frac{X_{ni}}{X_n} = \frac{T_i(c_i d_{ni})^{-\theta}}{\Phi_n}\tilde{T}_{ni} \tag{8}$$

where

$$T_i = \left(\sum_{k=1}^{K} (T_i^k)^{\frac{\sigma-1}{\theta}} \right)^{\frac{\theta}{\sigma-1}} \tag{9}$$

and

$$\tilde{T}_{ni} = \sum_{k=1}^{K} \left(\frac{\Phi_n^k}{\Phi_n} \right)^{\frac{\sigma-1}{\theta}-1} \frac{T_i^k}{T_i}. \tag{10}$$

Equation (8) is nearly an aggregate version of (3) with one key difference: the presence of the *bilateral* term $\tilde{T}_{ni}$. The rest of this equation is in common with other EK models, where exports from i to n are greater the greater is T_i, which governs the average productivity in i; the lower are production costs in i; and the lower are trade costs between i and n. Thus, Proposition 1 demonstrates that, when average productivity varies across products and countries, aggregate bilateral trade flows cannot be expressed as a function of only aggregate variables. Rather, they depend on the interaction among the patterns of comparative advantage of every country.

To understand how this interaction is embodied in $\tilde{T}_{ni}$, recall that Φ_n^k summarizes the state of technology for product k in every potential source country from the perspective of an importer in n. Now, suppose that $\theta > \sigma - 1$, which implies that the allocation of expenditure across source countries within a product category is more responsive to changes in production and trade costs than the allocation of expenditure across products.[10] Then, $\tilde{T}_{ni}$ is greater if T_i^k tends to be relatively large for products for which Φ_n^k is relatively small. In other words, if country i is relatively productive for products which are relatively difficult for n to obtain at a low cost, then, all else equal, i will export relatively more to n. The presence of trade barriers implies that Φ_n^k differs across countries, so $\tilde{T}_{ni}$ is country-pair specific and thus cannot be absorbed into an importer-specific or exporter-specific variable.

2.3 The Trade Cost Elasticity

How do countries' patterns of comparative advantage across products shape the effects of trade barriers on aggregate trade flows? To answer this question, first note that the partial elasticity of $\tilde{T}_{ni}$ with respect to the cost of exporting from i to n, holding production costs constant, can be expressed as

$$\frac{\partial \ln(\tilde{T}_{ni})}{\partial \ln d_{ni}} = [\theta - (\sigma - 1)] \left(\sum_{k=1}^{K} \frac{X_{ni}^k}{X_{ni}} \pi_{ni}^k - \pi_{ni} \right). \tag{11}$$

[10] This assumption is not mathematically necessary. However, if $\theta \leq \sigma - 1$, then $\lim_{K \to \infty} Y_i = \infty$, for all i, even if $K \cdot T_i^k$ is held constant for all i and k, which implies that the number of products with productivity greater than a given threshold is also constant. Further, this implies the counterintuitive result that an increase in T_i^k causes an increase in $X_{ni'}^k$, for $i' \neq i$. As a result, in the remaining discussion, I consider only the case where $\theta > \sigma - 1$, which according to the analysis that follows, appears to be consistent with the data.

Using this result, the partial elasticity of n's relative imports from i with respect to d_{ni} is given by

$$\begin{aligned}\varepsilon_{ni} \equiv \frac{\partial \ln(\pi_{ni}/\pi_{nn})}{\partial \ln d_{ni}} &= -\theta + \frac{\partial \ln \tilde{T}_{ni}}{\partial \ln d_{ni}} - \frac{\partial \ln \tilde{T}_{nn}}{\partial \ln d_{ni}} \\ &= -\theta + [\theta - (\sigma - 1)] \sum_{k=1}^{K} \pi_{ni}^{k} \left(\frac{X_{ni}^{k}}{X_{ni}} - \frac{X_{nn}^{k}}{X_{nn}} \right),\end{aligned} \tag{12}$$

where X_{ni}^k is the value of exports to n from i of k. This expression highlights the role of across-product comparative advantage in ameliorating the effect of trade barriers on aggregate trade flows. In general, $|\varepsilon_{ni}| \in [\sigma - 1, \theta]$, and it lies closer to $\sigma - 1$ the more i's exports tend to be concentrated in products for which it has a relatively large market share.

When trade costs from i to n increase, all of the varieties that n was importing from i become more expensive. This induces n to reallocate its expenditure both across sources of varieties within each product category and across products. If i has a strong comparative advantage in the products which it exports intensively, then an increase in trade costs results in a small reallocation away from i to other sources for the those products – because other sources are not particularly competitive – and a larger reallocation toward other products. Thus, it is the elasticity of substitution across products, σ, that largely governs the trade cost elasticity. On the other hand, if i's comparative advantage is weak in the products it exports intensively – because its pattern of relative productivity across products is similar to that of other countries from which n imports – then it is θ which largely determines the trade cost elasticity.

Since the second term in (12) results from the presence of $\tilde{T}_{ni}$ in (8), $\tilde{T}_{ni}$ can be interpreted as a measure of the degree to which the strength of i's comparative advantage across products, compared with the other potential sources of goods for n, reduces the negative effect of d_{ni} on i's exports to n.

Also, note that this elasticity is a function of endogenous variables, which implies that this model does belong to the class of models for which Arkolakis et al. (2012) show that the welfare effects of trade barriers depend only on aggregate trade flows and the trade cost elasticity.

2.4 Aggregate Productivity and the Gains from Trade

To see how countries' patterns of comparative advantage affect the welfare gains from trade, first note that welfare in this model is equivalent to real final output (GDP) per worker, Y_i/L_i. Using (8), along with the optimality conditions of the firms' cost-minimization problems, gives a useful expression for a country's welfare:[11]

$$y_i = \left(\frac{T_i}{\pi_{ii}} \right)^{\frac{1-\gamma}{\beta\theta}} \tilde{T}_{ii}^{\frac{1-\gamma}{\beta\theta}} k_i^{1-\alpha}. \tag{13}$$

[11] For simplicity, constant terms that are common across countries have been omitted.

As in Waugh (2010), this expression resembles that for real GDP per worker based on a standard neoclassical growth model in that y_i depends on the capital-labor ratio and a term representing total factor productivity.

To understand how TFP is determined, first consider a country in autarky. In this case, π_{ii} and $\tilde{T}_{ii}$ are equal to one, so aggregate productivity is equal to $T_i^{\frac{1-\gamma}{\beta\theta}}$. From (9), we can see that, if $\theta > \sigma - 1$, holding constant the average value of T_i^k, T_i is maximized if T_i^k is constant across products.[12] International trade, then, can benefit the economy two ways. First, it gives producers of the composite good access to additional sources for each variety of each intermediate input, which lowers the cost of obtaining some fraction of these varieties and has a similar effect on TFP as an increase in the average value of T_i^k. Second, if i's profile of T_i^k across products differs from the profiles of other countries with which it trades, then the country also benefits because it becomes relatively less costly to obtain varieties of the products for which it has a comparative disadvantage, which has a similar effect on TFP as a reduction in the dispersion of T_i^k across products.

In terms of the arguments of (13), π_{ii} is a weighted average of the fraction of varieties that are produced domestically, so the inverse of this term is a measure of the degree to which producers in i specialize according to comparative advantage. Given π_{ii}, the value of θ determines the gains from trade due to idiosyncratic (within-product) comparative advantage, where a smaller value of θ, i.e. more dispersion in productivity across varieties, implies larger gains from trade. The gains from trade due to deterministic (across-product) comparative advantage are determined by $\tilde{T}_{ii}$, which measures the degree to which i's technological profile across products differs from that of the rest of the world, summarized by the profile of Φ_i^k. Given these profiles, the value of $\tilde{T}_{ii}$ also depends on the value of the ratio $(\sigma - 1)/\theta$, where a smaller value of σ, which implies that products are less substitutable in production of the composite intermediate good, leads to greater gains from access for foreign-produced goods for which i has a comparative disadvantage.[13]

Again, it is the term $\tilde{T}_{ni}$ that summarizes the way in which this model deviates from aggregate quantitative trade models, in which changes in welfare due to changes in trade barriers depend only on changes in π_{ii} and the trade cost elasticity. To further highlight this role of $\tilde{T}_{ni}$, I next discuss a few special cases in which this term disappears, implying that this model is consistent with an aggregate quantitative trade model.

2.5 Some Special Cases

Before using the model to quantitatively assess the welfare effects of trade barriers, it is useful to more precisely explore the circumstances under which countries' patterns of comparative advantage

[12] This is because a greater value of θ implies greater diminishing returns to T_i^k in decreasing P_n^k, which can be seen in (6), while a greater value of $\sigma - 1$ implies a greater change in P_n associated with a given change in P_n^k as P_n^k decreases. So, $\theta > \sigma - 1$ implies that the economy benefits relatively more from having access to a large set of intermediate inputs than from having access to one very cheaply produced input.

[13] The parameters γ and β appear due to the role of intermediates in the production of both intermediate and final goods. Because intermediates are used to produce intermediates, this effect is amplified by to a degree determined by $1/\beta$. And, $1 - \gamma$ is the share of the composite intermediate good used in production of the final good.

influence aggregate trade flows and welfare. To this end proposition 2 lists three cases in which aggregate trade flows and welfare depend only on aggregate variables.

Proposition 2. *Suppose that any of the following hold:*

1. $T_i^k = T_i T^k, \forall i, k,$
2. $\frac{T_i^k}{\sum_k T_i^k} \in \{0, 1\}, \forall i, j,$
3. $\theta = \sigma - 1.$

Then, aggregate trade flows from source i to destination n can be expressed as

$$\pi_{ni} = \left(\frac{T_i (c_i d_{ni})^{-\theta}}{\Phi_n} \right)^{\frac{\rho}{\theta}}; \tag{14}$$

the trade cost elasticity $\varepsilon_{ni} = -\rho$, for all n and i; and welfare in a given country, i, is given by

$$y_i = \left(\frac{T_i^{\frac{\rho}{\theta}}}{\pi_{ii}} \right)^{\frac{1-\gamma}{\beta\rho}} k_i^{1-\alpha}.$$

In case 1, $\rho = \theta$; in case 2, $\rho = \sigma - 1$; and, in case 3, $\rho = \theta = \sigma - 1$.

Proofs are provided in Appendix C. In the first case, the model reduces to an EK model. Because there is no deterministic comparative advantage across products, trade occurs only because of idiosyncratic productivity differences across producers within product categories. This implies that $\tilde{T}_{ni} = 1$, and the trade cost elasticity is equal to θ. The second case is the opposite extreme in which each country produces a unique set of products. As result, there is no intra-product trade, $\tilde{T}_{ni} = \left(T_i (c_i d_{ni})^{-\theta} / \Phi_n \right)^{(\sigma-1)/\theta - 1}$, and trade costs only effect the relative price of one country's products versus another's, so the trade cost elasticity is equal to $\sigma - 1$. In this case, the model closely resembles that of Krugman (1980) or an Armington model, such as Anderson and van Wincoop (2003). In the final case, the parameters align such that expected benefit to an importer of gaining access to a new foreign variety is independent of the product category to which that variety belongs, so it is only aggregate stocks of technology, T_i, that matter for aggregate trade flows and welfare, not the distribution of technology parameters across products.

In all of these cases, the model reduces to one that fits within the framework of Arkolakis et al. (2012), in which trade flows follow an aggregate gravity equation, and the gains from trade depend only on π_{ii} and the trade cost elasticity. This provides a useful way to demonstrate how this model, in general, differs from these aggregate quantitative trade models. As discussed above, $\tilde{T}_{ii}$ is always equal to one in autarky, so the total welfare gain from trade relative to autarky is equal to $(\tilde{T}_{ii}/\pi_{ii})^{(1-\gamma)/\beta\theta}$. In case 1, the gains from trade arise only due to idiosyncratic productivity differences, so $\tilde{T}_{ii}$ disappears. In case 2, across-product productivity differences are so extreme that they swamp the effects idiosyncratic productivity differences within product categories. Thus, $\tilde{T}_{ni}$ negates the other terms in (8), replacing θ with $\sigma - 1$ as the trade cost elasticity and, as a result, the elasticity of welfare with respect to π_{ii}. And, in case 3 across-product productivity differences make no difference for aggregate outcomes, so, as in case 1, $\tilde{T}_{ii} = 1$.

In general, however, the value of $\tilde{T}_{ni}$ matters for the gains from trade, and the term's response to changes in trade barriers influence the responses of aggregate trade flows and welfare. Thus, while the gains from trade can be concisely summarized by (13), because there is no way to measure $\tilde{T}_{ii}$ using aggregate data, we must turn to disaggregated data to measure and predict changes in the gains from trade.

2.6 Equilibrium

To close the model, I assume that trade is balanced, i.e. total imports equal total exports for every country. Due to the Cobb-Douglas production functions of final and intermediate goods, each country devotes a constant share of labor and capital to each activity. This, combined with the balanced trade condition, implies that the set of wages obey the following conditions:

$$w_i L_i = \sum_{n=1}^{N} \pi_{ni} w_n L_n. \tag{15}$$

Alvarez and Lucas (2007) show that (15) defines a contraction mapping on $\{w_i\}$. Thus, these conditions define a unique set of wages for which the world economy is in equilibrium, given labor and capital endowments, trade costs, and the set of product-level technology parameters.

3 Comparative Advantage in the Data

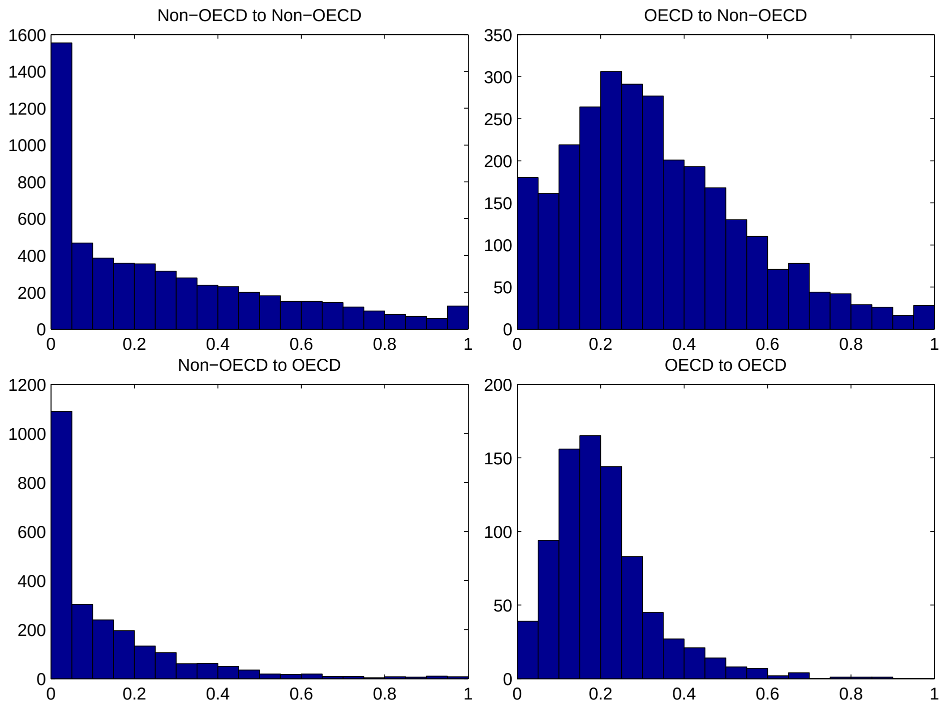
Proposition 1 makes clear that, in general, the effect of trade barriers on aggregate trade flows and welfare depends on the patterns of comparative advantage of every country in the world. However, Proposition 2 indicates that there are particular states of the world in which aggregate variables are sufficient to summarize the effects of trade barriers. To determine whether the world is approximately described by one of these cases, I investigate some patterns of the product-level trade data before going on to formally quantify the role of the composition of trade flows in determining the aggregate effects of trade barriers.

I use data from the U.N. Comtrade database, at the 6-digit level of Harmonized System – the most disaggregated classification for which bilateral trade data is widely available. The data comprise bilateral trade flows for 132 countries in 4,608 HS-6 product categories. Details are in Appendix B. To assess whether the data is roughly consistent with any of the cases of Proposition 2, I define the following Elasticity Index, based on (11):

$$\mathrm{EI}_{ni} = \frac{\sum_{k=1}^{K} \frac{X_{ni}^k}{M_n^k} \cdot \frac{X_{ni}^k}{X_{ni}} - \frac{X_{ni}}{M_n}}{1 - \frac{X_{ni}}{M_n}}$$

where k refers to an HS-6 product category. This index differs somewhat from (11) because of the

Figure 1: Histograms of Elasticity Index by Group



lack of product-level data on domestic trade flows. However, like the summation term in (12), the index is equal to zero in case 1 of Proposition 2 and equal to one in case 2.

Figure 1 plots histograms of the values of EI_{ni} for each country pair with positive trade flows, sorted by whether the source or destination is a high-income OECD country.[14] It is evident that there is a great deal of heterogeneity in the index, both within and across country groups, with OECD exports tending to have a higher EI_{ni}, while OECD imports tend to have a lower EI_{ni}. This indicates that there is variation in the patterns of comparative advantage across countries taking a form that the model predicts is important in determining aggregate trade flows and welfare – i.e. the world is not close to either case 1, in which EI_{ni} would be zero for every country pair, or case 2, in which EI_{ni} would be one for every country pair. In addition, the fact that EI_{ni} tends to be asymmetric – higher when the source country is a developed country rather than a developing country – suggests that trade barriers affect the trade flows of developed and developing countries differently, which may have further implications for the cross-country distribution of income.

It is still possible, however, that, as in case 3, non-trivial patterns of comparative advantage exist in the data but have little effect on aggregate trade flows and welfare. For some direct, albeit crude, evidence as to whether this is the case, consider the expression, based on (8), for n's relative imports from i:

$$\ln\left(\pi_{ni}/\pi_{nn}\right) = \ln(T_i c_i^{-\theta}) - \ln(T_n c_n^{-\theta}/\tilde{T}_{nn}) - \theta \ln(d_{ni}) + \ln(\tilde{T}_{ni}).$$

[14]High-income is defined according the World Bank classification used in the World Development Indicators database.

This value is a function of source-specific and destination-specific terms, bilateral trade costs, and $\tilde{T}_{ni}$. If we are in case 3, then $\tilde{T}_{ni} = 1$, and only country-specific terms and trade costs affect relative imports.[15] However, in general, bilateral trade flows will also depend on countries' patterns of comparative advantage, which are summarized by $\tilde{T}_{ni}$. If it were possible to measure $\tilde{T}_{ni}$ in the data, we could use this relationship to evaluate its importance for aggregate trade flows. While this is not the case, because EI_{ni} is based on the expression for $\partial \ln(\tilde{T}_{ni})/\partial \ln(d_{ni})$, (11), it can be utilized as part of the following linear approximation of $\tilde{T}_{ni}$

$$\ln(\tilde{T}_{ni}) \approx \ln(\tilde{T}_{ii}^{\mathrm{FT}}) + [\theta - (\sigma - 1)] \ln(d_{ni})\tilde{\mathrm{EI}}_{ni},$$

where $\tilde{\mathrm{EI}}_{ni} = (1 - X_{ni}/M_n) \times \mathrm{EI}_{ni}$ and $\tilde{T}_{ii}^{\mathrm{FT}}$ is the value of $\tilde{T}_{ni}$ under frictionless trade, using the fact that, at this point, $\tilde{T}_{ni}$ does not vary across destination because prics are equal everywhere.

Given this measure of $\ln(\tilde{T}_{ni})$, to evaluate it's effect on trade flows, a measure of bilateral trade costs is still required. In keeping with the gravity literature, I use distance as a rough proxy. Thus, I regress the log of relative imports on a set of source and destination fixed effects, the log of distance, and log distance interacted with $\tilde{\mathrm{EI}}_{ni}$. If $\theta > \sigma - 1$, then the coefficient on the last term should be positive, and it should be equal to zero if the world is described by case 3. Table 1 presents the coefficient estimates. The coefficient on the interaction term is positive and strongly significant, and the magnitude indicates that countries' patterns of comparative advantage are important in modulating the effects of trade barriers on aggregate bilateral trade flows.

Table 1: Correlation Between Trade Flows and Elasticity Index

Variable	Coefficient	Standard Error
$\ln(\mathrm{dist}_{ni})$	−1.978	0.023
$\ln(\mathrm{dist}_{ni}) \times \tilde{\mathrm{EI}}_{ni}$	0.430	0.009
No. of Obs.	11,588	
R^2	0.81	

Of course, since the approximation of $\ln(\tilde{T}_{ni})$ ignores the contributions of trade costs between other pairs of countries, and since non-distance-related bilateral trade barriers are ignored, this relationship could also be influenced by omitted variable bias. However, since there is no ex ante reason to suspect that these variables would be correlated with $\tilde{\mathrm{EI}}_{ni}$, the relationship evident in Table 1 can be taken as a strong indicator that countries' patterns of comparative advantage across products are important in determining the effect of trade barriers on aggregate trade flows.

Further, while this lends credence to the predictions of the model, it does not rule out that this relationship depends on other factors outside both this model and aggregate quantitative trade models. For instance, if trade costs in the model are allowed to vary by importer, exporter, and

[15] Relative imports also depend only on country-specific terms and trade costs in the other cases of Proposition 2, but in these cases, EI_{ni} is constant across county pairs, so the relationship between EI_{ni} and normalized import shares would be degenerate.

product, it is possible to construct any pattern of product-level and aggregate trade flows, which implies that Figure 1 and Table 1 could be replicated exactly even if the world is described by one of the special cases of Proposition 2. To address this issue, in addition to estimating trade costs that vary by sector in Section 6.2, I show in Appendix C that trade barriers that take the form $d_{ni}^k = d_{ni} d_n^k$ cannot generate the relationship evident in Table 1, which gives some indication that it would take a very special set of product-specific trade costs to generate these patterns in the data in the absence of non-trivial patterns of comparative advantage.[16]

4 Quantitative Implementation

To fully quantify the degree to which the effects of trade barriers on welfare depend on the composition trade flows, I compare the counterfactual changes in welfare due to changes in trade costs predicted by the baseline model with those of a model restricted to conform with case 1. In the latter case, the model reduces to a one-sector Eaton-Kortum model. Specifically, given the assumptions on production in the tradeable and nontradeable sectors, it reduces to the model of Waugh (2010), which is based on that of Alvarez and Lucas (2007). Because trade flows in the restricted model are consistent with an aggregate quantitative trade model, I fit the model only to aggregate data, and refer to it as the "aggregate model". The unrestricted model requires product-level trade data to discipline its parameters, so I refer to it as the "product-level" model. In the remainder of this section, I discuss the choice of parameter values and the estimation of trade costs necessary to compute the counterfactual results. Appendix B provides a detailed description of the data employed.

4.1 Parameter Values

Countries' endowments of labor and capital are computed directly from Heston et al. (2012). The capital stocks were computed based on the perpetual inventory method based on PPP-adjusted investment rates.

4.1.1 Elasticities

The elasticity of substitution across products, σ, is important for the quantitative predictions of models which feature monopolistic competition, such as Krugman (1980), and the gravity models based on them. For Ricardian EK models and models featuring heterogeneous firms, such as Chaney (2008), the value of θ is important. As a result, there have been many attempts to estimate the values of both.

For σ, I rely on the estimation of Broda and Weinstein (2006), which estimates the elasticity of substitution using the method developed by Feenstra (1994) and disaggregated US import data. In this model, σ is the elasticity of substitution *across* HS-6 categories, as opposed to that across

[16] For example, this type of variation in trade costs would be consistent with variation in tariffs across products that is consistent with the Most Favored Nation rule of WTO.

varieties *within* disaggregated categories. As a result, I set $\sigma = 2.2$, the median value estimated within 3-digit SITC categories for the period 1990-2001. This also happens to be very close to the value of 2, which, in the model of Ruhl (2004), is consistent with both estimates based on macroeconomic time series and those based on trade liberalization episodes.

To fix the value of θ, I turn to Simonovska and Waugh (2013), which develops a procedure to consistently estimate θ in the context of an EK model using international price data. However, their estimate requires some adjustment to be applicable to this model. Specifically, note that (8) implies the following relationship:

$$-\tilde{\theta} \equiv \frac{1}{N(N-1)} \sum_{n=1}^{N} \sum_{i \neq n} \frac{\ln(\pi_{ni}/\pi_{ii})}{\ln(d_{ni}P_i/P_n)} = -\theta + \frac{1}{N(N-1)} \sum_{n=1}^{N} \sum_{i \neq n} \frac{\ln(\tilde{T}_{ni}/\tilde{T}_{ii})}{\ln(d_{ni}P_i/P_n)}.$$

Simonovska and Waugh (2013) consistently estimate a value of $\tilde{\theta} = 4.1$, which is equal to θ under the assumption that case 1 of proposition 2 holds. Since we have seen that this is not the case in the data, the final term of the expression above must be subtracted from $\tilde{\theta}$ to obtain the value of θ that is consistent with the model. This is possible after estimating trade costs and recovering the technology parameters, T_i^k, which is done below. After doing this, I find that a value of $\theta = 6.0$ is consistent with a measured value of $\tilde{\theta} = 4.1$. In Section 6.1, I explore the implications of using different values of σ and θ.

4.1.2 Production Function Parameters

Labor's share in total value added is determined by α. Following Gollin (2002), I set $\alpha = 2/3$. The share of value added in total output in manufacturing is governed by β. I set $\beta = 0.3$ to match the value of this share over all countries in the OECD STAN database in 2003. The parameter γ serves two roles. Like β, it determines the share of value added in output in nontradeables, and it also determines the share of nontradeables in total value added. In the STAN database, for 2003, the average value of the former is 0.77, and the average value of the latter is 0.50. I choose an intermediate value, and set $\gamma = 0.65$.[17]

4.2 Estimating Trade Costs

Many quantitative trade models take advantage of the gravity-like structure of the model to estimate trade costs as a function of observable variables expected to influence barriers to trade, using bilateral trade data. Following this literature, I parameterize trade costs in the following way:

$$\ln d_{ni} = \ln d_i + dist_{ni}^m + bord_{ni} + lang_{ni} + col_{ni} + rta_{ni} + ex_i, \tag{16}$$

[17]The values of α, β, and γ are important for the model's predictions of both the baseline level of income per worker and the response of income to changes in trade barriers. However, they matter little for the differences in predictions between the product-level and aggregate models. Therefore, for the sake of brevity, I do not conduct any sensitivity analyses with regard to these parameters.

where $dist_{ni}^{m}$ is the effect of the distance between n and i lying in one of six distance intervals; $bord_{ni}$ is the effect of countries n and i sharing a common border; $lang_{ni}$ is the effect of sharing a common language; col_{ni} is the effect of having a colonial relationship; rta_{ni} is the effect of n and i being part of a regional trade agreement; and d_i is an exporter-specific border cost.[18]

In the special cases of Proposition 2, aggregate bilateral trade flows follow a standard gravity equation, based on (14), so standard techniques from the gravity literature can be used to estimate the coefficients of (16) using aggregate trade data. In general, however, trade costs cannot be identified from aggregate bilateral trade flows separately from the effect of composition, via the $\tilde{T}_{ni}$ term in (8). Fortunately, though, (3) is consistent with a gravity equation at the product level, which implies that trade costs can be identified from disaggregated trade flows.

Using product-level data to estimate trade costs introduces two complications that are not present in gravity estimations based on aggregate trade flows. First, at the HS-6 level of aggregation, the typical practice of using source and destination fixed effects to control for endogenous country-specific variables becomes infeasible. Such an estimation, based on (3), would require employing more than one million source-product and destination-product fixed effects, which is well beyond the abilities of a typical computer using standard software and techniques. The second issue is that there is a lack of data on output or expenditure at a level of disaggregation comparable to the trade data for the vast majority of countries. This is problematic because domestic trade is not reported in the trade data and is required to estimate the trade costs associated with international borders, which account for a very large share of trade costs and which Waugh (2010) argues are important for understanding cross-country differences in income.

4.2.1 Product-Level Gravity

To overcome these challenges, I employ the method of French (2014), which uses the structure of the model to avoid the need for fixed effects, as in Anderson and van Wincoop (2003), and requires only aggregate data on output to recover country-specific border costs. This method relies on the fact that, based on (3), product-level, bilateral trade flows can be expressed as the solution to the following system:

$$\begin{aligned} X_{ni}^{k} &= \frac{M_n^k E_i^k}{E^k} \frac{\tilde{d}_{ni}^{-\theta}}{\tilde{\Phi}_n^k \tilde{\Psi}_i^k} \\ \tilde{\Phi}_n^k &= \sum_{i \neq n} \frac{\tilde{d}_{ni}^{-\theta}}{\tilde{\Psi}_i^k} \frac{E_i^k}{E^k} \\ \tilde{\Psi}_i^k &= \sum_{n \neq i} \frac{\tilde{d}_{ni}^{-\theta}}{\tilde{\Phi}_n^k} \frac{M_n^k}{E^k} \end{aligned} \tag{17}$$

[18] I assume that this effect is exporter-specific, rather than importer-specific, following Waugh (2010), which argues that this specification is more consistent with data on the prices of tradable goods.

where $\tilde{d}_{ni} = d_{ni}/d_i$, M_n^k and E_i^k are product-level imports and exports, respectively, to and from the rest of the world, and $E^k = \sum_i E_i^k$. Importantly, (17) makes clear that the product-level technology indexes, $\tilde{\Phi}_n^k$, and $\tilde{\Psi}_i^k$ – which Anderson and van Wincoop (2003) refer to as outward multilateral resistance – can be calculated as functions of only observable data and trade costs.[19] Further, because everything in (17) is expressed as a function of imports and exports, rather than output and expenditure, the parameters of the trade cost function can be estimated using only data on international trade.

Note that the value of d_i has no effect on X_{ni}^k in (17). This is due to the insight of Anderson and van Wincoop (2003) that only *relative* trade costs matter in determining international trade flows, and since (17) expresses bilateral trade flows conditional on total imports and exports – i.e. conditional on the goods having already crossed an international border – border costs are irrelevant. Thus, d_i is not directly identified by the product-level gravity estimation. The procedure for recovering these parameters is discussed below.

I estimate the parameters of (16), net of the value of d_i, via structural Poisson pseudo-maximum likelihood. In other words, given a set of parameters and data on total product-level imports and exports, (17) yields predicted bilateral product-level trade flows. The estimation procedure finds the set of parameter values that maximizes the Poisson likelihood function, given the observed trade flows. The Poisson likelihood function is chosen due to the attractive properties discussed in Santos Silva and Tenreyro (2006) and its current widespread use in gravity estimations. In fact, French (2014) shows that, when aggregate data is used rather than product-level data, this procedure gives identical coefficient estimates as the currently standard fixed-effects Poisson PML estimation procedure. So, the coefficient estimates based on product-level data can be directly compared to the estimates based on aggregate data in the literature.[20]

4.2.2 Recovering Technology Parameters and Border Costs

Given estimated values of all the variables of (17), it is possible to recover values for the interaction of technology parameters and production and border costs. However, because (17) depends only on product-level data – in the same way that only the ratio T_i^k/Φ_n^k matters for the value of π_{ni}^k in (3) – only relative values of left-hand of these variables across countries are identified. Thus, I define $T_i^k \equiv T^k \tilde{T}_i^k$ and impose that $\sum_i \tilde{T}_i^k = 1$. Then, I recover the interaction of technology and cost variables from the the following identity:

$$\tilde{T}_i^k (c_i d_i)^{-\theta} = \frac{E_i^k}{E^k} (\tilde{\Psi}_i^k)^{\theta}. \tag{18}$$

[19] The term $\tilde{\Phi}_n^k$ is the analogue of Φ_n^k calculated over all countries except n. Thus, $\Phi_n^k = \tilde{\Phi}_n^k + T_n^k c_n^{-\theta}$.

[20] French (2014) also shows that, while the coefficient from product-level estimations based on this technique vary little across the choice of objective function, estimates based on aggregate data vary much more widely. However, the aggregate and product-level estimates based on the Poisson likelihood function are remarkably similar, indicating that this estimator is more robust to misspecification than alternative ones, which lends further support for its wide use in all types of gravity estimations.

Given estimates of $\tilde{d}_{ni}$ and $\tilde{T}_i^k(c_i d_i)^{-\theta}$, having estimates of $d_i^{-\theta}$ and T^k would make it possible to recover the values of $d_{ni}^{-\theta}$ and $T_i^k c_i^{-\theta}$. With values of $d_{ni}^{-\theta}$ and $T_i^k c_i^{-\theta}$, the values of π_{nn}^k can be calculated according to (3). In addition, given the values of σ and θ and data on total manufacturing expenditure X_n, product-level expenditure can be calculated according to (5). Using these results, I take the set of values of T^k to be those for which the model's implied world trade flows for each product match the data, i.e.

$$E^k = \sum_{n=1}^{N}(1-\pi_{nn}^k)\left(\frac{\Phi_n^k}{\Phi_n}\right)^{\frac{\sigma-1}{\theta}} X_n,$$

and I take border costs, d_i, to be the values for which the model's predicted domestic trade shares match the data, i.e.

$$\frac{X_{nn}}{X_n} = \sum_{k=1}^{K}\pi_{nn}^k\left(\frac{\Phi_n^k}{\Phi_n}\right)^{\frac{\sigma-1}{\theta}}.$$

Because domestic trade shares depend on the set of T^k's, and world trade in each product depends on the set of d_i's, they must be solved for jointly.

Finally, to recover the values of T_i^k, it is necessary to remove domestic production costs from $T_i^k c_i^{-\theta}$. To do this, I use data on aggregate bilateral trade flows and the model's equilibrium conditions (15) to solve for wages, i.e.

$$w_i = \sum_{n=1}^{N}\frac{X_{ni}}{X_n}\frac{w_n L_n}{L_i}.$$

Given wages and using price levels computed according to (7) and capital-labor ratios from the data, domestic costs are given by

$$c_i = \tilde{\beta}(w_i k_i^{\alpha-1})^{\beta} P_i^{1-\beta},$$

which uses the fact that, due to the Cobb-Douglas production functions for final output and intermediate varieties, in equilibrium, $r_i = (1-\alpha)w_i/(\alpha k_i)$.[21]

4.2.3 Estimation Results

Table 2 reports the estimates of the coefficients of (16) based both on aggregate trade flow data data, consistent with the the restricted model, and on product-level data, consistent with the unrestricted model. As noted above, the estimation based on aggregate data is identical to a Poisson PML estimation using source and destination fixed effects to proxy for endogenous variables. As is well known in the gravity literature, the elasticity of trade costs with respect to the independent variables is not separately identified from the value of θ. Because the restricted model implies a different value of θ, and thus a different mapping between the coefficient estimates and the associated effects on trade costs, than the unrestricted model, the percentage effect of each variable on trade costs is

[21] The constant $\tilde{\beta} = (\alpha\beta)^{-\beta}(1-\beta)^{\beta-1}$.

also reported.

Table 2: Trade Cost Coefficient Estimates

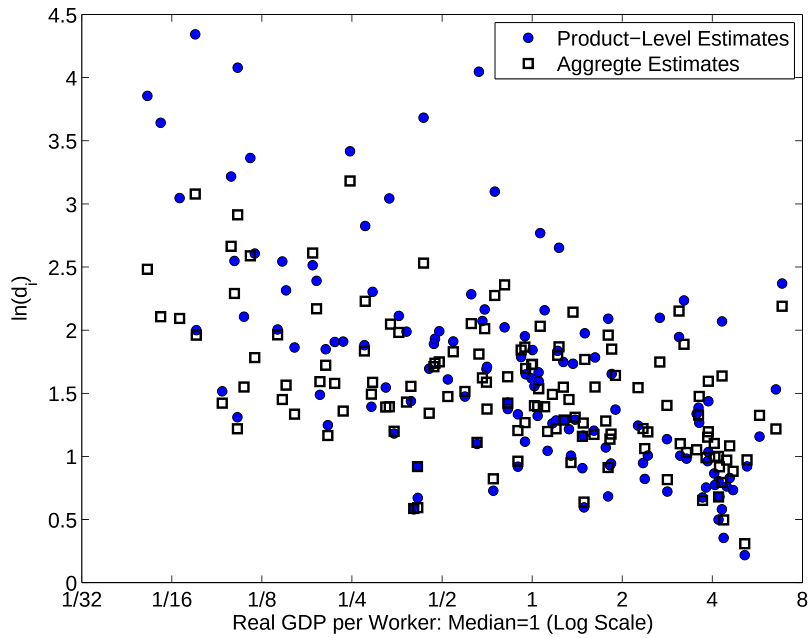
	Coefficients		Percentage Effect	
	Aggregate	Product Level	Aggregate	Product Level
median($\ln(d_i)$)	-6.04	-9.30	336.72	377.56
625 - 1,250 km	-0.30	-0.40	7.70	6.98
	(0.21)	(0.23)		
1,250 - 2,500 km	-0.65	-0.82	17.22	14.87
	(0.28)	(0.32)		
2,500 - 5,000 km	-1.01	-1.24	28.08	23.20
	(0.34)	(0.43)		
5,000 - 10,000 km	-1.80	-2.25	55.21	45.86
	(0.34)	(0.42)		
>10,000 km	-1.93	-2.59	60.10	54.48
	(0.39)	(0.46)		
Shared Border	0.54	0.56	-12.32	-8.91
	(0.11)	(0.13)		
Common Language	0.32	0.37	-7.52	-6.02
	(0.08)	(0.08)		
Colonial Ties	0.14	0.15	-3.33	-2.48
	(0.12)	(0.10)		
RTA	0.81	0.84	-17.88	-13.23
	(0.10)	(0.11)		
No. of Obs.	17,292	79,681,536		
Value of θ			4.10	5.75

Notes: Standard errors, clustered by source country, are in parentheses. Coefficients reported are multiplied by $-\theta$, as the effects of the independent variable of interest and the trade elasticity are not separately identified by the gravity estimation. The implied percentage effect of each coefficient on the ad valorem tariff equivalent trade cost is $100 \times (e^{\text{-}coeff/\theta} - 1)$, where *coeff* is the reported coefficient.

As is discussed in more detail in French (2014), in the case of Poisson PML, the coefficient estimates based on aggregate and product-level data are quite similar, though the product-level distance coefficients are somewhat greater in absolute value. The largest difference is in the average level of exporter-specific border costs. The intuition for this is straightforward. Since the restricted model assumes no role for across-product comparative advantage, countries tend to trade less of their output for a given level of trade barriers when compared with the unrestricted model, so higher values of $\theta \ln(d_i)$ are required for the unrestricted model to match observed trade flows. However, because the unrestricted model implies a larger value of θ – i.e. less comparative advantage across varieties within product categories – the effect of the higher coefficients on implied trade costs is dampened. Taken together, the estimates based on the aggregate model imply somewhat lower border costs and somewhat higher trade costs due to distance and bilateral relationships than those based on the product-level model.

To get a sense of the distribution of border costs across countries, Figure 2 plots the estimated values of $\ln(d_i)$ against income per worker for each model, showing that border costs are generally higher for low-income countries, and this difference tends to be greater in the product-level estimates.

Figure 2: Border Costs Estimates



4.3 Quantitative Implications

4.3.1 Income Per Worker

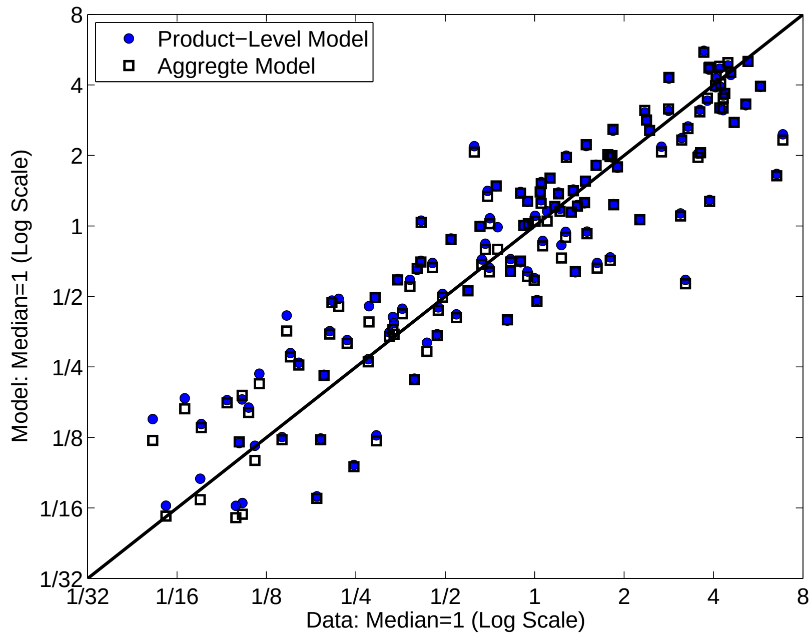
Figure 3 reports the predicted values of final output per worker in the both the aggregate and product-level models against real GDP per worker in the data. Clearly both models do quite well at matching the cross-country distribution of income. This is not surprising, as Waugh (2010) has shown models that feature exporter-specific border costs perform quite well in this regard. However, it is reassuring that the predictions of the product-level model do not deviate far from those of the aggregate model in an area where the aggregate model is known to perform well, and it lends credence to the counterfactual experiments that follow, which focus on the effects of trade barriers on relative income levels across countries.

Both models slightly under-predict the dispersion of income across countries. The variance of log GDP per worker in the data is 1.53, and the ratio of the 90th to the 10th percentile is 31.3. The aggregate model predicts values of 1.30 and 23.3, respectively, while the product-level model predicts respective values of 1.24 and 22.3. I find it reasonable that both under-predict the level of dispersion as the model makes the simplifying assumption that productivity in the non-tradeable sector is constant across countries.

4.3.2 Tradeable Goods Prices

The aggregate model predicts an elasticity of the tradeable goods price, P_n, with respect to income – measured as the coefficient estimated by regressing the log of P_n on the log of the country's level of real GDP per worker – of -0.058, which is highly statistically significantly different from zero.

Figure 3: Real GDP per Worker: Model vs. Data



The product-level model predicts an elasticity of 0.007, which is not significantly different from zero. Waugh (2010) reports an elasticity 0.15 estimated from benchmark price data from the Penn World Tables. Thus, the product-level model performs relatively well in predicting the relationship between tradeable goods prices and levels of income across countries.

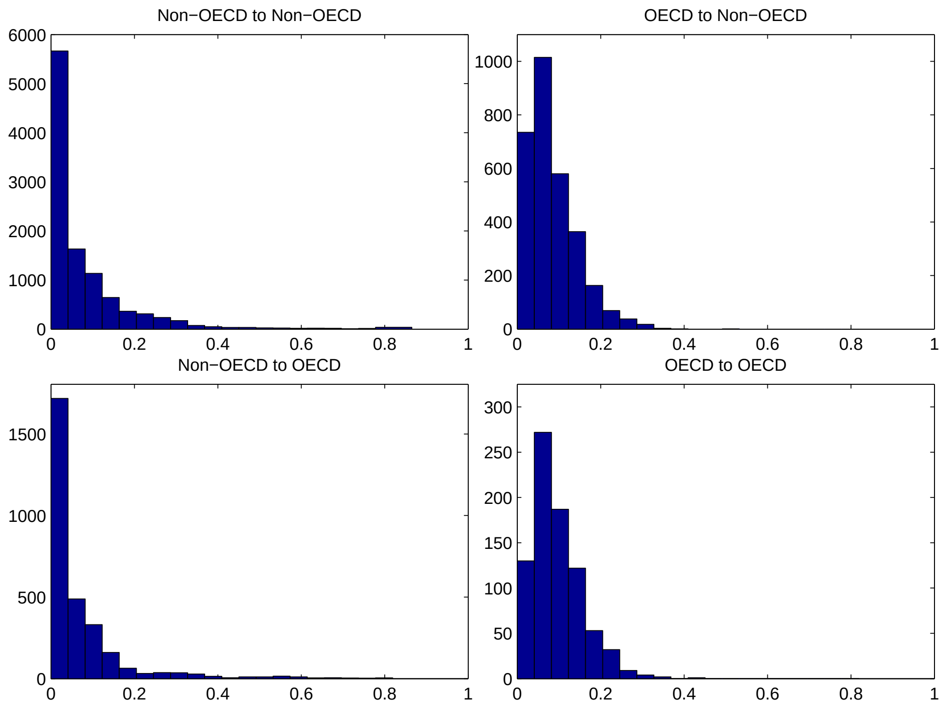
4.3.3 The Elasticity Index and Bilateral Trade Flows

To test the ability of the product-level model to accurately predict relevant moments of the product-level trade data, I replicate the calculations behind Figure 1 and Table 1 using the values of product-level trade flows predicted by the model. Figure 4 depicts histograms of EI_{ni}, calculated using predicted product-level trade flows, and Table 3 presents the coefficient estimates of the regression of predicted relative imports on source and destination fixed effects, log distance, and log distance interacted with the predicted values of $\tilde{\mathrm{EI}}_{ni}$.

Overall, the product-level model does rather well in replicating these patterns. As in the data, the elasticity index tends to be smaller when the source is a non-OECD country, though the predicted values tend to be smaller than those in the data for all sets of countries. Table 3 shows that the model almost exactly matches the measured effect of comparative advantage, proxied by $\tilde{\mathrm{EI}}_{ni}$, on the relationship between trade barriers and aggregate trade flows.[22] Keeping in mind that the aggregate model predicts no variation in EI_{ni} and thus zero correlation between EI_{ni} and aggregate trade flows, this indicates that the product-level model, despite maintaining most of

[22]Because the relationship between distance and trade flows predicted by the model is a function of the Poisson PML gravity estimation, the under-prediction of the coefficient on log distance is consistent with the fact that, as is discussed in French (2014), Poisson PML tends to estimate a smaller coefficient on distance than does log-linear OLS.

Figure 4: Histograms of Elasticity Index by Group



the restrictions of the aggregate model, goes quite far in accurately predicting the composition of bilateral trade flows.

Table 3: Correlation Between Trade Flows and Elasticity Index

Variable	Coefficient	Standard Error
$\ln(\text{dist}_{ni})$	-0.808	0.005
$\ln(\text{dist}_{ni}) \times \text{EI}_{ni}$	0.431	0.009
No. of Obs.	17,292	
R^2	0.98	

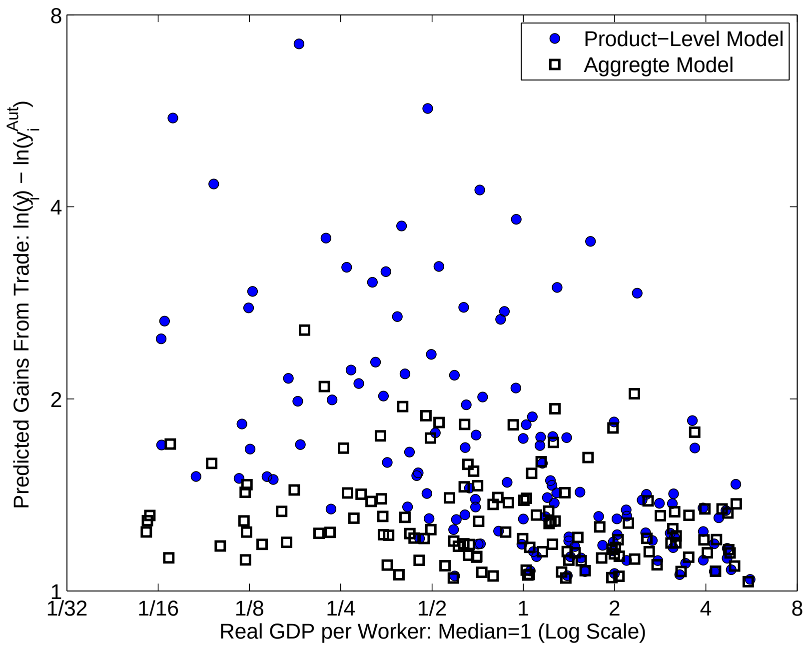
4.3.4 The Importance of Composition

In addition to the relationship depicted in Table 3, because the product-level model predicts values of $\tilde{T}_{ni}$, it provides a more direct way to measure the importance of countries' patterns of comparative advantage in determining aggregate trade flows. One way of doing this is to quantify the degree to which the variance in $\tilde{T}_{ni}$ is important in predicting these flows.

The log-linear form of (8) provides a particularly straightforward way to measure the importance of $\tilde{T}_{ni}$ in explaining the cross-sectional variation in bilateral trade flows:

$$\log \pi_{ni} = \ln(T_i c_i^{-\theta}) - \ln(\Phi_n) + \ln(d_{ni}^{-\theta}) + \log(\tilde{T}_{ni})$$

Figure 5: Gains From Trade: Aggregate vs. Product-Level Model



Because this is an identity, regressing the left-hand side of this expression on all the right-hand side variables yields an R^2 of one. Performing the same regression but omitting the final term yields an R^2 of 0.75, indicating that 25 percent of the variation in predicted aggregate bilateral trade flows is due to the variation in $\tilde{T}_{ni}$, which is to say that it is due to the interaction among countries' patterns of comparative advantage across products.

4.3.5 The Gains from Trade

As is discussed above, in the restricted model, the gains from trade relative to autarky are given by $\pi_{ii}^{-\frac{1-\gamma}{\beta\theta}}$, while in the product-level model they also depend on patterns of across-product comparative advantage, embodied in $\tilde{T}_{ii}$. Thus, before conducting any formal counterfactual experiments, it is possible compare the predictions of these two models on this front.

Figure 5 depicts the predicted gains from trade relative to autarky for each model. Table 4, described below, also depicts the average change in log income and the change in the dispersion of income across countries. Because both models match the data on aggregate domestic trade shares, this component of the predicted gains from trade is identical. The predictions differ for two reasons. First, the aggregate model assumes no across-product comparative advantage, i.e. $\tilde{T}_{ii} = 1$, so it will automatically predict lower gains from trade arising from such. Second, however, because the value of θ needed for the product-level model to predict the value of $\tilde{\theta}$ estimated by Simonovska and Waugh (2013) is larger than that for the aggregate model, the product-level model will predict lower gains from trade due to idiosyncratic productivity differences within product categories.

As is clear in Figure 5, the gains from trade predicted by the product-level model are generally larger and are more skewed in favor of low-income countries.[23] This indicates that, in general, the gains from trade due to comparative advantage across products is much greater than the over-prediction of the gains from trade due to comparative advantage within product categories by the restricted model.

Using the results of the product-level model, the log-linear form of (13) allows for a simple decomposition of the gains from trade into across-product and within-product components, i.e.

$$\ln(y_i/y_i^{Aut}) = -\frac{1-\gamma}{\beta\theta}\ln(\pi_{ii}) + \frac{1-\gamma}{\beta\theta}\ln(\tilde{T}_{ii}),$$

where the first term represents the within-product component of the gains from trade, and the second term represents the across-product component. The average value of the first term across countries is 0.16, and the average value of the second term is 0.30, which implies that roughly two-thirds of the gains from trade are due to comparative advantage across products. However, as Figure 5 suggests, this result is largely driven by low-income countries. If we look only at high-income OECD countries, only 38% of the gains from trade are due to across-product comparative advantage, which explains why the discrepancy between the models' predictions is larger for low-income countries.

To better understand what drives this result, consider the effect of a uniform worldwide change in trade costs on $\tilde{T}_{ii}$. Formally, denoting trade costs as $d_{ni} \equiv \tilde{d}_{ni}\bar{d}$, for $n \neq i$, the partial elasticity of $\tilde{T}_{ii}$ with respect to $\bar{d}$ can be expressed as

$$\frac{\partial \ln(\tilde{T}_{ii})}{\partial \ln(\bar{d})} = -\left[\theta - (\sigma - 1)\right]\left(\sum_{k=1}^{K}\frac{X_{ii}^k}{X_{ii}}\pi_{ii}^k - \pi_{ii}\right).$$

This implies that, compared to autarky, as trade costs fall, $\tilde{T}_{ii}$ increases faster for countries whose domestic market share is relatively high for the products which make up a relatively large share of domestic trade. Put another way, countries whose domestic trade flows are concentrated in relatively few products are predicted to find autarky more costly. It turns out that this tends to be the case for low-income countries.

This result may seem surprising given the previous finding that EI_{ni} tends to be smaller for low-income exporters, which requires that their market shares tend to be small for the products that they export intensively. The reason these two results hold simultaneously is that low-income countries tend to have similar profiles of T_i^k and, as a result, tend to specialize in a similar set of products. On the other hand, high-income countries tend to have profiles are relatively unique, especially when compared with low-income countries, but also relative to other high-income countries. Thus, low-income countries' exports tend to be highly responsive to trade barriers, since they are competing closely with other low-income countries for market share. On the other hand, as the only country that does not face trade barriers in the domestic market, they face little competition from other

[23]The product-level model's predicted gains from trade are smaller than the aggregate model's for 13 mostly large, developed countries, including the US, Germany, UK, France, and Korea.

low-income countries in their comparative advantage products, while they benefit from access to their comparative disadvantage products, which tend to be imported from high-income countries.

5 Counterfactual Predictions

In addition to comparing the gains from trade relative to autarky predicted by each model, since both models are fully parameterized, it is possible to consider the effects of other counterfactual changes in trade barriers. I choose two counterfactual experiments which allow for comparison with the results of similar experiments based on aggregate quantitative trade models in the literature. The first additional experiment is the elimination of all asymmetric barriers to trade, i.e. $d_{ni} = \min\{d_{ni}, d_{in}\}$, which Waugh (2010) argues likely reflects policy-related barriers, as most natural impediments to trade, such as distance, are symmetric by nature, as opposed to asymmetric barriers, such as border costs. The second experiment is a move to entirely frictionless trade, i.e. $d_{ni} = 1$.

Table 4 depicts the predicted changes in welfare in both models for each counterfactual experiment. For each scenario, the table reports the average change in log income per worker across all countries relative to the baseline case as well as two measures of the dispersion of income across countries, the variance of log income per worker and the ratio of the 90th to the 10th percentile. Table 5 decomposes changes in welfare into a component determined by within-product comparative advantage and one determined by across-product comparative advantage. The average change in log income per worker is decomposed into the separate effects of changes in π_{ii} and $\tilde{T}_{ii}$, as in the previous section. I decompose changes in the variance of log income per worker using the following identity derived from (13):

$$\begin{aligned}
&\Delta \mathrm{var}\left(\ln(y_i)\right) = \\
&\quad\left[\left(\frac{1-\gamma}{\beta\theta}\right)^2 \mathrm{var}\left(\Delta\ln(\tilde{\pi}_{ii})\right) - 2\frac{1-\gamma}{\beta\theta}\mathrm{cov}(\ln(y_i), \Delta\ln(\tilde{\pi}_{ii})) - \left(\frac{1-\gamma}{\beta\theta}\right)^2 \mathrm{cov}(\Delta\ln(\tilde{T}_{ii}), \Delta\ln(\tilde{\pi}_{ii}))\right] \\
&\quad + \left[\left(\frac{1-\gamma}{\beta\theta}\right)^2 \mathrm{var}(\Delta\ln(\tilde{T}_{ii})) + 2\frac{1-\gamma}{\beta\theta}\mathrm{cov}(\ln(y_i), \Delta\ln(\tilde{T}_{ii})) - \left(\frac{1-\gamma}{\beta\theta}\right)^2 \mathrm{cov}(\Delta\ln(\tilde{T}_{ii}), \Delta\ln(\tilde{\pi}_{ii}))\right]
\end{aligned}$$

Thus, the first expression in brackets represents the change in the variance of log income related to within-product comparative advantage, and the second expression presents the change related to across-product comparative advantage.

The basic finding of all of these experiments is that the welfare gains from trade tend to be larger and more skewed toward low-income countries in the product-level model. Consistent with the discussion in the previous section, the cost of moving to autarky is nearly twice as large, on average, in the product-level model, and the effect on the dispersion of income is much larger. The variance of log income increases by 46% in the product-level model, compared with only 8% in aggregate model, which is similar to Waugh (2010), in which moving to autarky has a very small effect on the dispersion of income.

The gains from removing asymmetric trade barriers follow a similar pattern to the gains from moving from autarky to the baseline case. The average change in log income is significantly higher

Table 4: Counterfactual Income per Worker: Aggregate vs. Product-Level Models

Variable	Model	Baseline	Autarky	$\min(d_{ni}, d_{in})$	$d_{ni} = 1$
$\text{mean}(\Delta \log(y_i))$	Aggregate	–	-0.23	0.36	1.58
	Product-Level	–	-0.46	0.52	1.60
$\text{var}(\log(y_i))$	Aggregate	1.30	1.41	0.92	0.66
	Product-Level	1.24	1.81	0.74	0.55
y_{90}/y_{10}	Aggregate	23.29	27.27	16.52	9.48
	Product-Level	22.34	36.51	11.28	7.35

Table 5: Decomposition of Changes in Income per Worker

	Autarky	$\min(d_{ni}, d_{in})$	$d_{ni} = 1$
Contribution to $\text{mean}(\Delta \log(y_i))$			
π_{ii}	-0.16	0.22	1.04
$\tilde{T}_{ii}$	-0.30	0.29	0.55
Contribution to $\Delta\text{var}\,(\log(y_i))$			
π_{ii}	0.08	-0.16	-0.31
$\tilde{T}_{ii}$	0.48	-0.35	-0.38

in the product-level model, and the variance of log income falls by 40% in the product-level model compared with only 29% in the aggregate model. However, in this scenario, the intuition behind the result is slightly different. Because border costs tend to be higher for low-income countries, these countries tend to benefit relatively more from the elimination of these barriers. The major reason that low-income countries gain relatively more from the elimination of border costs in the product-level model is that the estimates of these costs based on the product-level model are generally greater for these countries.

A similar line of reasoning is behind the differences between the model's predictions of the gains from eliminating all trade barriers. Because the product-level model estimates generally lower symmetric trade barriers, the additional gains from removing these barriers, above the gains from removing border costs, is predicted to be lower, and the additional gain to low-income countries is also relatively lower. While the two models predict nearly identical average gains from moving to frictionless trade from the baseline, there is no fundamental reason why this must be the case. However, Table 5 provides some insight into why the gains from trade predicted by the product-level model taper off, relative the predictions of the aggregate model, as the world moves from autarky to free trade. When the world is close to autarky, the gains from trade due to across-product comparative advantage, which are missing from the aggregate model, dominate, but as the trade barriers continue to fall, the gains due to within-product comparative advantage become more important. Since the latter gains are lower in the product-level model, due to the higher value of θ, the product-level model predicts smaller gains from reducing trade barriers when the world is close to free trade.

Put another way, because the product-level model takes into account the gains from trade due to across-product comparative, which is precluded by the aggregate model, and because these gains are relatively large when the world is closer to autarky, the product-level model predicts that, in terms of the possible gains from trade, the current state of the world is much farther from autarky, and the elimination of asymmetric trade barriers would move the world much closer to free trade, than the aggregate model would suggest.[24]

Finally, it is worth noting that the summary statistics reported in Tables 4 and 5, mask a great deal of heterogeneity in the predictions of the two models for individual countries. For example, though the overall effects of moving from the baseline to free trade are fairly similar, the average absolute difference between the change in log income predicted by the two models is 0.15. In particular, the product-level model predicts much larger gains for many African countries and much smaller gains for many former Soviet Republics. For example, for the eastern African countries of Eritrea, Ethiopia, and Sudan, the average difference between the product-level and aggregate models' predicted changes in log income is 0.69, while for the Baltic states of Estonia, Latvia, and Lithuania, the average difference is -0.21. This indicates that the former group have much stronger patterns of across-product comparative advantage, which may reflect their colonial history, compared to the latter group, whose economic structure was influenced by central planning and a period of relative isolation from potential trading partners in Western Europe.

6 Robustness

Computing the baseline counterfactual results required making several choices of parameter values and relied on the assumption of a single tradeable goods sector. In this section, I evaluate the restrictiveness of those choices, first conducting the counterfactual experiments using alternative parameter values and then doing so using a multi-sector extension to the baseline model.

6.1 Alternative Elasticity Values

I consider three alternative sets of values of the elasticities θ and σ. First, to evaluate the effect of having different values of θ in the aggregate and product-level models, I set $\theta = 4.1$ in the product-level model to match the value used in the the aggregate model. Second, I set the $\theta = 8.3$, the value estimated by Eaton and Kortum (2002) which has been used in many subsequent papers. Finally, to measure the sensitivity of the model to the value of σ, I set $\sigma = 3.3$, the median value estimated by Broda and Weinstein (2006) for the elasticity of substitution across products within 10-digit U.S. HTS product categories.

Table 6 reports the same measures of income as Table 4 for each scenario. For the sake of comparision, the first row associated with each measure reports the results based on the baseline set of parameters. The most striking result is that the value of θ does not matter much for the

[24] I am thankful to a referee for pointing out this interpretation of the results.

Table 6: Counterfactual Real Income per Worker: Alternative Parameters

Variable	θ	σ	Baseline	Autarky	$\min(d_{ni}, d_{in})$	$d_{ni} = 1$
mean($\Delta \log(y_i)$)	6.0	2.2	–	-0.46	0.52	1.60
	4.1	2.2	–	-0.50	0.59	2.08
	8.3	2.2	–	-0.43	0.46	1.28
	6.0	3.1	–	-0.31	0.45	1.47
var($\log(y_i)$)	6.0	2.2	1.24	1.81	0.74	0.55
	4.1	2.2	1.28	1.84	0.73	0.51
	8.3	2.2	1.22	1.81	0.76	0.60
	6.0	3.1	1.27	1.54	0.81	0.60
y_{90}/y_{10}	6.0	2.2	22.34	36.51	11.28	7.35
	4.1	2.2	23.14	35.24	10.63	6.99
	8.3	2.2	21.89	37.71	11.68	8.38
	6.0	3.1	22.80	29.63	14.03	9.04

welfare effects of going from the baseline to autarky and removing asymmetric trade barriers. In the case of autarky, this is because an increase in the θ decreases the gains from trade due to within-product comparative advantage but increases the gains due to across-product comparative advantage. In the case of removing asymmetric trade barriers, the result also depends on the fact that, when θ is larger, a greater value of $\theta \ln(d_i)$ is required for the model to match domestic trade shares, especially for low-income countries, for whom domestic trade shares are less responsive to trade costs. It turns out that in both cases, the offsetting effects of θ nearly cancel out.

Where θ does appear to be relatively important is in regard to the welfare gains of moving to frictionless trade. This is not surprising given the result that, in moving from a world with only symmetric trade barriers to one of free trade, the gains from trade are primarily due to within-product comparative advantage, which is governed by θ. Further, the result that low-income countries benefit relatively more from across-product comparative advantage is consistent with the observation that the dispersion of income is less affected by θ in the move to free trade trade than is the average increase in income.

Finally, keeping in mind that when $\theta = \sigma - 1$, the aggregate and product-level models are equivalent, increasing σ has the expected effect of shifting all of the predictions of the product-level model closer to those of the aggregate model.

6.2 Multiple Sectors

In this section, I extend the model to allow for multiple manufacturing sectors, or "industries", in similar fashion to the models of Caliendo and Parro (2012) and Levchenko and Zhang (2013). The extension is analytically straightforward, though the expressions are slightly more complex. The major drawback of this framework is that the data requirements are much greater. In particular, identifying border costs that vary by sector requires data on sector-level manufacturing output. I use data on 18 manufacturing sectors, which are roughly equivalent to 2-digit ISIC industies. Availability of this data decreases the sample size from 132 to 60 countries (details are in Appendix

B), with the poorest countries among those that must be excluded. Thus, the major tradeoff in allowing for cross-sector heterogeneity in trade costs is that the ability of this framework to address issues related to the cross-country income distribution is somewhat limited.

There are $j = 1, .., J$ manufacturing sectors, each of which is comprised of $k = 1, ..., K^j$ products. The structure of the model is otherwise the same. Now, a given variety is identified by the triple (j, k, ω), and the product category to which it belongs is defined by the pair (j, k). In this setup, the composite tradeable good is a Cobb-Douglas aggregate of sectoral composite goods given by

$$Q_n = \prod_{j=1}^{J} (Q_n^j)^{\delta^j}, \tag{19}$$

where $\delta^j > 0$ is the share of sector j in total tradeable expenditure, and $\sum_j \delta^j = 1$. The aggregation of varieties into products and products into sectoral composite goods mirror their counterparts from the single-sector version of the model.

Iceberg trade costs, $d_{ni}^j > 0$, are allowed to vary by sector. I define average bilateral trade costs, d_{ni}, as the sector-share-weighted geometric mean of d_{ni}^j, i.e.

$$d_{ni} = \prod_{j=1}^{J} (d_{ni}^j)^{\delta^j}.$$

Then, as in the single-sector version of the model, aggregate bilateral trade flows are given by (8), except that, now,

$$T_i = \prod_{j=1}^{J} \left(\sum_{k=1}^{K^j} (T_i^{jk})^{\frac{\sigma-1}{\theta}} \right)^{\delta^j \frac{\theta}{\sigma-1}},$$

and

$$\begin{aligned} \tilde{T}_{ni} &= \sum_{j=1}^{J} \delta^j \frac{\Phi_n}{\Phi_n^j} \left(\frac{d_{ni}^j}{d_{ni}} \right)^{-\theta} \sum_{k=1}^{K^j} \left(\frac{\Phi_n^{jk}}{\Phi_n^j} \right)^{\frac{\sigma-1}{\theta}-1} \frac{T_i^{jk}}{T_i} \\ &= \sum_{j=1}^{J} \delta^j \frac{\Phi_n}{\Phi_n^j} \left(\frac{d_{ni}^j}{d_{ni}} \right)^{-\theta} \tilde{T}_{ni}^j, \end{aligned}$$

where $\Phi_n^j = \left(\sum_{k=1}^{K^j} (\Phi_n^{jk})^{\frac{\sigma-1}{\theta}} \right)^{\frac{\theta}{\sigma-1}}$, and $\Phi_n = \prod_j (\Phi_n^j)^{\delta^j}$.

The interpretation of T_i – the determinant of aggregate productivity in autarky – remains the same, as does the sector-level version of $\tilde{T}_{ni}$, denoted $\tilde{T}_{ni}^j$. The key difference is that relative trade costs across sectors now interact with the forces of across-product comparative advantage to influence trade flows. That is, if trade costs from i to n are relatively low in the sectors for which the covariance between T_i^{jk} and Φ_n^{jk} is also relatively low, then i will export relatively more to n.

Because the expression for π_{ni} in the multi-sector version of the model is identical to that for the single-sector version, so is the expression for welfare (13). Thus, the basic intuition for the welfare effects of trade barriers is preserved, with the only major difference being that, through

T_{ii}, as with aggregate trade flows, the welfare gains from trade, conditional on the domestic import share, are now affected by relative trade costs across sectors. Because labor is mobile across all sectors, equilibrium is defined by the same set of conditions as the single-sector model (15).

6.2.1 Parameter Values

I use the same set of common parameter values as with the single-sector version of the model to quantify the multi-sector version. Trade costs are also estimated and technology parameters recovered in the same way, except that the estimations are performed sector-by-sector and employ data on sector-level manufacturing output. As above, I fit both aggregate and product-level versions of the multi-sector model to the data, where the aggregate multi-sector model again assumes that the world economy is characterized by case 1 of Proposition 2, so only data on total sector-level trade flows are required to estimate the model's parameters.

Table 7: Average Trade Cost Coefficient Estimates: Single vs. Multi-Sector Models

	Single-Sector		Multi-Sector	
	Aggregate	Product Level	Aggregate	Product Level
median($\ln(d_i)$)	−4.85	−6.41	−5.00	−6.09
625 - 1,250 km	−0.29	−0.38	−0.43	−0.48
	(0.20)	(0.24)		
1,250 - 2,500 km	−0.62	−0.77	−0.85	−0.97
	(0.29)	(0.33)		
2,500 - 5,000 km	−0.95	−1.16	−1.23	−1.41
	(0.34)	(0.43)		
5,000 - 10,000 km	−1.77	−2.16	−2.13	−2.46
	(0.34)	(0.43)		
>10,000 km	−1.86	−2.43	−2.29	−2.78
	(0.39)	(0.47)		
Shared Border	0.55	0.57	0.56	0.60
	(0.11)	(0.14)		
Common Language	0.27	0.31	0.27	0.31
	(0.09)	(0.09)		
Colonial Ties	0.07	0.09	0.19	0.18
	(0.13)	(0.10)		
RTA	0.76	0.81	0.76	0.83
	(0.10)	(0.13)		
No. of Obs.	3,540	16,294,620	3,540	16,294,620

Notes: Standard errors, clustered by source country, are in parentheses. Coefficients reported are multiplied by $-\theta$, as the effects of the independent variable of interest and the trade elasticity are not separately identified by the gravity estimation. Coefficients reported for the multi-sector estimations are sector-share-weighted average values, i.e. $\mathit{coeff} = \sum_j \delta^j \mathit{coeff}^j$.

Table 7 reports the sector-weighted average coefficient estimates for the multi-sector estimations along with those from single-sector estimations, comparable to those reported in Table 2, based on the sample of countries for which sectoral-level manufacturing output data is available, for the sake of comparison. Because the reported coefficients for the multi-sector estimations are averages across separate estimations, standard errors are not reported. The sector-specific coefficient estimates and standard errors are reported in Table A3. The average multi-sector estimates are generally

Table 8: Counterfactual Real Income per Worker: Single vs. Multi-Sector Models

Variable	Sectors	Model	Baseline	Autarky	$\min(d_{ni}, d_{in})$	$d_{ni} = 1$
mean($\Delta \log(y_i)$)	Single	Agg.	–	-0.17	0.19	1.24
		P.L.	–	-0.26	0.26	1.17
	Multi	Agg.	–	-0.27	0.20	1.27
		P.L.	–	-0.35	0.44	1.46
var($\log(y_i)$)	Single	Agg.	0.83	0.88	0.67	0.47
		P.L.	0.78	0.99	0.56	0.41
	Multi	Agg.	0.85	0.98	0.70	0.47
		P.L.	0.80	1.15	0.56	0.41
y_{90}/y_{10}	Single	Agg.	11.38	12.83	8.50	6.07
		P.L.	9.90	11.79	6.70	5.21
	Multi	Agg.	11.59	17.19	8.32	5.73
		P.L.	11.10	21.22	6.44	4.34

in line with their single-sector counterparts. For the aggregate estimations, the average multi-sector coefficient estimates tend to be larger in absolute value than their single-sector counterparts, indicating that the implied trade costs are somewhat larger. A similar pattern holds for the product-level case, except that the differences tend to be smaller, and the border cost estimates are smaller in absolute value, indicating that the implied overall trade costs are much more similar in the product-level estimation.

6.2.2 Counterfactual Experiments

I consider the same set of counterfactual experiments as above using the multi-sector version of the model. Table 8 presents the results for both the aggregate and product-level versions of the multi-sector model as well as both versions of the single-sector model with the sample restricted to the set of countries for which sectoral-level manufacturing output data is available.

By and large, the differences in the predictions of the product-level model relative to the aggregate model are very similar in the multi-sector and single-sector versions of the models. Namely, the gains from trade relative to autarky and the gains from removing border costs are larger and more skewed toward low-income countries in the product-level model. Further, it is only in regard to the gains from trade relative to autarky in which the predictions of the multi-sector, aggregate model differ substantially from the single-sector aggregate model. This suggests that, in general, to measure the effects of composition on aggregate trade flows and welfare, they must be accounted for at a level of disaggregation much lower than the industry level.

While the effect of trade barriers on aggregate trade flows and welfare appears to depend much more on countries' patterns of across-product comparative advantage than on differences in trade costs or average productivity across industries, accounting for the latter does lead to different predictions of the product-level model in some cases. The fact that low-income countries benefit relatively more from trade relative to autarky in the multi-sector, product-level model indicates that trade barriers tend to be lower in the sectors in which these countries have relatively strong comparative advantage versus high-income countries, such as the textiles industry. The other notable difference is that the average change in income from eliminating asymmetric trade barriers is larger in the multi-sector, product-level model, though the predicted effect on the dispersion of income is nearly identical to the single-sector, product-level model. This result is largely driven by the food, beverage, and tobacco industry, in which the gains due to across-product comparative advantage are relatively large and uncorrelated with development and for which trade barriers are estimated to be particularly high.

Taken together, the overall pattern that emerges from these counterfactual experiments is that, once across-product comparative advantage has been taken into account, allowing for heterogeneity in trade costs and average productivity across industries either makes little difference for the effects of trade barriers on welfare or amplifies the differences between the predictions of the product-level and aggregate single-sector models. Thus, given that product-level trade data is available for a much larger set of countries than industry-level output data, one could conclude that studies interested in the welfare effects of changes in bilateral trade barriers should consider the effect of countries' patterns of across-product comparative advantage, even when data requirements preclude the consideration of other forms of sectoral heterogeneity, such as variation in border costs.

A final point of concern for some may be the assumption that θ is constant across sectors. While some papers, such as Caliendo and Parro (2012), have attempted to estimate θ at the sectoral level, due to data requirements, such estimates are likely to be less reliable than aggregate estimates such as that of Simonovska and Waugh (2013).[25] More importantly, the robustness exercises reported in Table 6 indicate that the welfare effects of border-type trade barriers are not particularly sensitive to the value of θ. This implies that sectoral variation in θ would also be relatively unimportant in measuring the welfare effects of changes in such barriers. In the case of geographic barriers, where a higher value of θ led to somewhat smaller gains from the removal of geographic barriers to trade, then sectoral variation in θ could lead to an under-prediction of the welfare gains from lowering such barriers if θ tended to be lower for sectors in which these barriers tend to be high.

7 Conclusion

This paper has developed a framework in which to quantitatively assess the role of the composition of trade flows in determining the aggregate welfare effects of trade barriers. One advantage of this framework is that it maintains the tractability, relatively low data requirements, and parsimony in expressing the gains from trade of many aggregate quantitative trade models while also allowing the effects of the interactions among countries' patterns of comparative advantage across products to be taken into account. The framework nests models that belong to the class delineated by Arkolakis et al. (2012), in which the gains from trade are a function of only aggregate trade flows and the trade cost elasticity, and makes clear that, more generally, the gains from trade also depend on a third term which succinctly summarizes the effect of across-product comparative advantage.

[25] As Levchenko and Zhang (2014) point out, it is not clear the extent to which sectoral estimates depend on the structure that must be imposed on the estimation or how sensitive they are to measurement error.

Quantifying this effect requires using data on product-level trade flows, which are now reported by nearly every country for which aggregate trade flow data are available, but it imposes no further data requirements except where other simplifying restrictions of this class of aggregate trade models are relaxed. Thus, the insights and methods of this paper can be applied in a wide range of contexts in which the effects of trade barriers were inferred from aggregated trade flows, including industry-level studies.

Applying these methods to questions of the aggregate gains from trade compared to autarky and the gains from removing asymmetric trade barriers, this paper has shown that, compared with the predictions of an aggregate EK model, falling trade barriers lead to significantly larger increases in income per worker on average and larger increases for developing relative to developed countries than the aggregate model would predict. Thus, accounting for the way in which the effect of trade barriers on aggregate welfare depends on patterns of comparative advantage across products is quantitatively important, which suggests that such patterns should not be ignored in questions of the aggregate causes and consequences of worldwide international trade flows, especially in regard to questions of the role of international trade in economic development.

References

Alvarez, Fernando and Robert E. Lucas Jr., "General Equilibrium Analysis of the Eaton-Kortum Model of International Trade," *Journal of Monetary Economics*, 2007, *54* (6), 1726 – 1768.

Anderson, James E., "A Theoretical Foundation for the Gravity Equation," *The American Economic Review*, 1979, *69* (1), 106–116.

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

_ **and** _ , "Trade Costs," *Journal of Economic Literature*, 2004, *42* (3), 691–751.

_ **and Yoto V. Yotov**, "Specialisation: Pro and Anti-Globalizing 1990-2002," 2011, *Working Paper.*

Arkolakis, Costas, Arnaud Costinot, and Andres Rodriguez-Clare, "New Trade Models, Same Old Gains?," *American Economic Review*, 2012, *102* (1), 94–130.

Broda, Christian and David E. Weinstein, "Globalization and the Gains from Variety," *Quarterly Journal of Economics*, 2006, *121* (2), 541–585.

Caliendo, Lorenzo and Fernando Parro, "Estimates of the Trade and Welfare Effects of NAFTA," Technical Report 2012.

Caselli, Francesco, "Accounting for Cross-Country Income Differences," *Handbook of Economic Growth*, 2005, *1*, 679–741.

Chaney, Thomas, "Distorted Gravity: the Intensive and Extensive Margins of International Trade," *American Economic Review*, 2008, *98* (4), 1707–1721.

Chen, Natalie and Dennis Novy, "Gravity, trade integration, and heterogeneity across industries," *Journal of International Economics*, 2011, *85* (2), 206 – 221.

Costinot, Arnaud and Andres Rodriguez-Clare, "Trade Theory with Numbers: Quantifying the Consequences of Globalization," in Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff, eds., *Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff, eds.*, Vol. 4 of *Handbook of International Economics*, Elsevier, 2013.

_ **, Dave Donaldson, and Ivana Komunjer**, "What Goods Do Countries Trade? A Quantitative Exploration of Ricardo's Ideas," *Review of Economic Studies*, 2012, *79* (2), 581–608.

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

Feenstra, Robert C., "New Product Varieties and the Measurement of International Prices," *The American Economic Review*, 1994, *84* (1), pp. 157–177.

Fieler, Ana Ceclia, "Nonhomotheticity and Bilateral Trade: Evidence and a Quantitative Explanation," *Econometrica*, 2011, *79* (4), 1069–1101.

French, Scott, "Biased Gravity Estimates: Heteroskesticity or Misspecification?," 2014, *Working Paper.*

Gollin, Douglas, "Getting Income Shares Right," *Journal of Political Economy*, 2002, *110* (2), pp. 458–474.

Helpman, Elhanan, Marc Melitz, and Yona Rubinstein, "Estimating Trade Flows: Trading Partners and Trading Volumes," *Quarterly Journal of Economics*, 2008, *123* (2), 441–487.

Heston, Alan, Robert Summers, and Bettina Aten, *Penn World Table Version 7.1,* Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, November 2012.

Krugman, Paul, "Scale Economies, Product Differentiation, and the Pattern of Trade," *The American Economic Review*, 1980, *70* (5), pp. 950–959.

Levchenko, Andrei A. and Jing Zhang, "The Evolution of Comparative Advantage: Measurement and Welfare Implications," 2013, *Working Paper.*

_ **and** _ , "Ricardian Productivity Differences and the Gains from Trade," *European Economic Review*, 2014, *65* (0), 45 – 65.

Mayer, Thierry and Soledad Zignago, "Notes on CEPIIs distances measures (GeoDist)," 2011, *CEPII Working Paper 2011-25.*

McFadden, Daniel, "Conditional Logit Analysis of Qualitative Choice Behavior," in Paul Zarembka, ed., *Paul Zarembka, ed.*, Frontiers in Econometrics, Academic Press, 1974.

Melitz, Marc, "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 2003, *71*, 1695–1725.

Ruhl, Kim J., "The International Elasticity Puzzle," *Working Paper*, 2004.

Samuelson, Paul A., "The Transfer Problem and Transport Costs, II: Analysis of Effects of Trade Impediments," *The Economic Journal*, 1954, *64* (254), pp. 264–289.

Silva, J. M. C. Santos and Silvana Tenreyro, "The Log of Gravity," *The Review of Economics and Statistics*, 2006, *88* (4), pp. 641–658.

Simonovska, Ina and Michael E. Waugh, "The Elasticity of Trade: Estimates and Evidence," 2013, *Working Paper.*

Waugh, Michael E., "International Trade and Income Differences," *American Economic Review*, 2010, *100* (5), 2093–2124.

A Additional Tables

Table A1: ISIC Rev. 3 Sectors

ISIC code	Sector Description	HS-6 Products	δ^j
15A	Food, beverages, and tobacco	427	0.145
17	Textiles	541	0.023
18	Wearing apparel, fur	241	0.017
19	Leather, leather products, and footwear	67	0.007
20	Wood products (excluding furniture)	69	0.019
21	Paper and paper products	119	0.030
22	Printing and publishing	36	0.047
23	Coke, refined petroleum products, nuclear fuel	20	0.053
24	Chemicals and chemical products	877	0.102
25	Rubber and plastics products	121	0.039
26	Non-metallic mineral products	170	0.032
27	Basic metals	359	0.061
28	Fabricated metal products	221	0.055
29C	Office, accounting, computing machinery; Other machinery	565	0.093
31A	Electrical machinery; Communication equipment	235	0.085
33	Medical, precision and optical instruments	211	0.020
34A	Transport equipment	135	0.133
36	Furniture, other manufacturing	189	0.036

Table A2: Countries and Sources of Manufacturing Output Data

Country	Source	Country	Source	Country	Source
Albania	INDSTAT	Gambia	WDI	Panama	INDSTAT(int.)
Argentina	WDI	Georgia*	INDSTAT	Papua New Guinea	WDI
Australia*	INDSTAT	Germany*	STAN	Paraguay	WDI
Austria*	STAN	Ghana	INDSTAT	Peru	INDSTAT
Azerbaijan*	INDSTAT	Greece*	STAN	Philippines	INDSTAT
Bahamas	WDI	Guatemala	WDI	Poland*	STAN
Bangladesh	WDI	Honduras	WDI	Portugal*	STAN
Barbados	WDI	Hungary*	STAN	Qatar	INDSTAT
Belarus	WDI	Iceland*	STAN	Rep. of Korea*	STAN
Belize	WDI	India*	INDSTAT	Rep. of Moldova*	INDSTAT
Benin	WDI	Indonesia*	INDSTAT	Romania	INDSTAT
Bolivia	WDI	Iran*	INDSTAT	Russian Federation*	INDSTAT
Bosnia Herzegovina	WDI	Ireland*	STAN	Rwanda	WDI
Botswana	INDSTAT	Israel*	STAN	St. Lucia	WDI
Brazil*	INDSTAT	Italy*	STAN	St. Vinc. and Gren.	WDI
Brunei Darussalam	WDI	Jamaica	WDI	Samoa	WDI
Bulgaria*	INDSTAT	Japan*	STAN	Sao Tome and Princ.	WDI
Burkina Faso	WDI	Jordan*	INDSTAT	Saudi Arabia	INDSTAT(int.)
Burundi	WDI	Kazakhstan*	INDSTAT	Senegal	WDI
Cambodia	WDI	Kenya*	INDSTAT	Slovakia*	STAN
Cameroon	WDI	Kyrgyzstan*	INDSTAT	Slovenia*	STAN
Canada*	STAN	Latvia*	INDSTAT	South Africa*	INDSTAT
Cape Verde	WDI	Lebanon	WDI	Spain*	STAN
Central African Rep.	WDI	Lithuania*	INDSTAT	Sri Lanka	INDSTAT(int.)
Chile*	INDSTAT	Madagascar*	INDSTAT	Sudan	WDI
China*	INDSTAT	Malawi	WDI	Swaziland	WDI
Colombia*	INDSTAT	Malaysia*	INDSTAT	Sweden*	STAN
Costa Rica	WDI	Maldives	WDI	Switzerland*	STAN
Croatia	WDI	Malta*	INDSTAT	Syria	INDSTAT
Cuba	WDI	Mauritania	WDI	TFYR of Macedonia	INDSTAT
Cyprus	INDSTAT	Mauritius	INDSTAT	Thailand*	INDSTAT(int.)
Czech Rep.*	STAN	Mexico*	STAN	Togo	WDI
Cte d'Ivoire	WDI	Morocco	INDSTAT	Trinidad and Tobago*	INDSTAT
Denmark*	STAN	Mozambique	WDI	Tunisia*	INDSTAT
Dominican Rep.	WDI	Namibia	WDI	Turkey*	INDSTAT
Ecuador*	INDSTAT	Nepal	WDI	USA*	STAN
El Salvador	WDI	Netherlands*	STAN	Uganda	WDI
Eritrea	INDSTAT	New Zealand*	STAN	Ukraine*	INDSTAT
Estonia*	STAN	Nicaragua	WDI	United Kingdom*	STAN
Ethiopia	INDSTAT	Niger	WDI	Utd. Rep. of Tanzania	INDSTAT
Fiji	INDSTAT	Nigeria*	INDSTAT	Uruguay*	INDSTAT
Finland*	STAN	Norway*	STAN	Venezuela	WDI
France*	STAN	Oman	INDSTAT	Viet Nam	INDSTAT
Gabon	WDI	Pakistan	INDSTAT(int.)	Zambia	WDI

* Sector-level manufacturing output data available.

Notes: INDSTAT(int.) indicates that output data was interpolated based on INDSTAT data for years before and after 2003.

Table A3: Multi-Sector Trade Costs Coefficient Estimates

(a) Aggregate Model

	Food	Textiles	Apparel	Leather	Wood	Paper	Printing	Petr/Coal	Chemicals	Rub/Plstc	Minerals	Bas. Metal	Fab. Metal	Computing	Electrical	Medical	Transport	Furniture
median(ln(d_i))	−5.19	−3.71	−4.14	−3.70	−4.95	−4.80	−7.13	−4.03	−4.67	−4.57	−5.21	−4.07	−5.32	−4.38	−3.95	−5.19	−5.44	−5.46
625 - 1,250 km	−0.59	−0.81	−0.97	−1.21	−0.60	−0.48	−0.55	−1.08	−0.02	−0.58	−0.79	−0.15	−0.58	−0.18	−0.50	0.25	−0.22	−0.43
	(0.15)	(0.37)	(0.25)	(0.45)	(0.31)	(0.23)	(0.30)	(0.30)	(0.37)	(0.20)	(0.21)	(0.25)	(0.19)	(0.28)	(0.29)	(0.39)	(0.40)	(0.28)
1,250 - 2,500 km	−1.24	−1.03	−1.25	−1.69	−1.36	−1.09	−0.87	−1.60	−0.39	−1.07	−1.45	−0.76	−1.15	−0.37	−0.55	0.04	−0.54	−1.03
	(0.19)	(0.50)	(0.30)	(0.45)	(0.38)	(0.31)	(0.42)	(0.32)	(0.51)	(0.29)	(0.31)	(0.30)	(0.27)	(0.43)	(0.39)	(0.53)	(0.53)	(0.38)
2,500 - 5,000 km	−1.89	−2.06	−2.02	−2.33	−1.72	−1.37	−1.27	−1.74	−0.78	−1.61	−1.94	−1.03	−1.53	−0.58	−1.15	−0.26	−0.51	−1.47
	(0.25)	(0.62)	(0.47)	(0.54)	(0.69)	(0.49)	(0.63)	(0.65)	(0.65)	(0.40)	(0.40)	(0.31)	(0.39)	(0.48)	(0.43)	(0.60)	(0.74)	(0.40)
5,000 - 10,000 km	−2.71	−2.75	−2.79	−2.64	−2.67	−2.49	−2.02	−3.79	−1.70	−2.64	−2.91	−2.04	−2.50	−1.46	−2.08	−1.19	−1.11	−1.69
	(0.29)	(0.49)	(0.39)	(0.51)	(0.71)	(0.52)	(0.60)	(0.46)	(0.63)	(0.42)	(0.38)	(0.31)	(0.42)	(0.52)	(0.42)	(0.64)	(0.74)	(0.53)
>10,000 km	−2.72	−3.10	−3.03	−2.59	−3.24	−2.66	−2.04	−4.50	−2.06	−2.64	−3.30	−2.53	−2.56	−1.49	−1.90	−1.45	−1.36	−1.49
	(0.34)	(0.56)	(0.48)	(0.59)	(0.88)	(0.57)	(0.79)	(0.65)	(0.73)	(0.50)	(0.48)	(0.34)	(0.46)	(0.59)	(0.50)	(0.69)	(0.83)	(0.59)
Shared Border	0.49	0.44	0.59	0.37	0.99	0.65	0.47	0.90	0.40	0.62	0.65	0.79	0.64	0.40	0.45	0.38	0.57	0.68
	(0.10)	(0.22)	(0.18)	(0.21)	(0.22)	(0.12)	(0.19)	(0.31)	(0.17)	(0.11)	(0.12)	(0.11)	(0.10)	(0.16)	(0.23)	(0.21)	(0.17)	(0.16)
Common Language	0.25	0.49	0.52	0.11	−0.23	−0.07	0.59	−0.26	0.53	0.24	0.20	0.20	0.29	0.31	0.06	0.48	0.34	0.39
	(0.16)	(0.18)	(0.24)	(0.23)	(0.23)	(0.20)	(0.20)	(0.31)	(0.21)	(0.10)	(0.12)	(0.12)	(0.11)	(0.12)	(0.09)	(0.20)	(0.14)	(0.17)
Colonial Ties	0.50	0.10	0.20	0.41	0.35	0.20	0.35	0.64	−0.01	0.30	0.19	0.36	0.33	0.19	0.01	0.01	−0.28	0.18
	(0.12)	(0.24)	(0.26)	(0.24)	(0.12)	(0.14)	(0.15)	(0.26)	(0.17)	(0.13)	(0.11)	(0.14)	(0.08)	(0.14)	(0.13)	(0.16)	(0.22)	(0.15)
RTA	0.68	0.90	0.78	0.77	0.59	0.88	1.25	−0.08	0.38	0.94	0.48	0.67	0.86	0.77	0.53	0.09	1.38	1.20
	(0.18)	(0.20)	(0.27)	(0.30)	(0.34)	(0.30)	(0.36)	(0.37)	(0.15)	(0.21)	(0.18)	(0.19)	(0.18)	(0.18)	(0.20)	(0.23)	(0.22)	(0.22)
No. of Obs.	3540.00	3540.00	3540.00	3540.00	3540.00	3540.00	3540.00	3540.00	3540.00	3540.00	3540.00	3540.00	3540.00	3540.00	3540.00	3540.00	3540.00	3540.00

(b) Product-Level Model

	Food	Textiles	Apparel	Leather	Wood	Paper	Printing	Petr/Coal	Chemicals	Rub/Plstc	Minerals	Bas. Metal	Fab. Metal	Computing	Electrical	Medical	Transport	Furniture
median(ln(d_i))	−8.54	−4.28	−4.72	−4.59	−6.29	−5.80	−8.27	−4.01	−6.19	−5.39	−7.50	−5.48	−6.12	−5.22	−4.64	−6.06	−6.23	−6.55
625 - 1,250 km	−0.73	−0.78	−0.91	−0.95	−0.64	−0.54	−0.55	−1.06	−0.08	−0.57	−0.81	−0.48	−0.62	−0.18	−0.46	0.02	−0.19	−0.62
	(0.12)	(0.25)	(0.23)	(0.24)	(0.23)	(0.20)	(0.28)	(0.34)	(0.36)	(0.19)	(0.19)	(0.18)	(0.20)	(0.21)	(0.25)	(0.37)	(0.34)	(0.32)
1,250 - 2,500 km	−1.61	−1.10	−1.32	−1.50	−1.64	−1.33	−0.89	−1.61	−0.41	−1.09	−1.48	−1.18	−1.19	−0.47	−0.58	−0.05	−0.52	−1.18
	(0.17)	(0.35)	(0.30)	(0.21)	(0.28)	(0.27)	(0.42)	(0.37)	(0.52)	(0.29)	(0.27)	(0.23)	(0.29)	(0.31)	(0.35)	(0.56)	(0.41)	(0.45)
2,500 - 5,000 km	−2.30	−2.05	−2.14	−2.04	−2.35	−1.93	−1.28	−1.95	−0.79	−1.65	−1.92	−1.58	−1.62	−0.61	−1.30	−0.29	−0.68	−1.64
	(0.29)	(0.43)	(0.42)	(0.38)	(0.45)	(0.41)	(0.66)	(0.74)	(0.70)	(0.40)	(0.34)	(0.27)	(0.42)	(0.42)	(0.36)	(0.69)	(0.59)	(0.64)
5,000 - 10,000 km	−3.30	−3.12	−3.17	−2.75	−3.72	−3.21	−2.08	−4.23	−1.75	−2.75	−2.92	−2.84	−2.66	−1.70	−2.30	−1.20	−1.42	−2.20
	(0.30)	(0.38)	(0.37)	(0.40)	(0.44)	(0.42)	(0.62)	(0.55)	(0.69)	(0.40)	(0.36)	(0.30)	(0.44)	(0.44)	(0.40)	(0.70)	(0.58)	(0.62)
>10,000 km	−3.49	−3.57	−3.40	−2.79	−4.57	−3.82	−2.27	−5.21	−2.19	−2.91	−3.46	−3.58	−2.84	−1.86	−2.27	−1.46	−1.73	−2.23
	(0.43)	(0.42)	(0.44)	(0.55)	(0.56)	(0.49)	(0.78)	(0.66)	(0.77)	(0.45)	(0.44)	(0.31)	(0.47)	(0.47)	(0.45)	(0.74)	(0.57)	(0.74)
Shared Border	0.63	0.56	0.57	0.43	1.05	0.71	0.53	0.96	0.47	0.64	0.78	0.76	0.66	0.41	0.38	0.43	0.62	0.69
	(0.12)	(0.16)	(0.18)	(0.17)	(0.18)	(0.13)	(0.20)	(0.34)	(0.19)	(0.11)	(0.13)	(0.08)	(0.11)	(0.15)	(0.21)	(0.25)	(0.16)	(0.21)
Common Language	0.38	0.52	0.48	0.36	0.01	0.19	0.64	−0.17	0.44	0.34	0.37	0.07	0.37	0.34	0.22	0.48	0.32	0.17
	(0.09)	(0.16)	(0.26)	(0.19)	(0.21)	(0.16)	(0.22)	(0.32)	(0.12)	(0.09)	(0.10)	(0.10)	(0.11)	(0.10)	(0.09)	(0.14)	(0.11)	(0.18)
Colonial Ties	0.43	0.17	0.25	0.47	0.43	0.21	0.33	0.56	0.09	0.25	0.17	0.37	0.35	0.12	−0.01	0.02	−0.32	0.22
	(0.13)	(0.16)	(0.25)	(0.22)	(0.11)	(0.11)	(0.13)	(0.24)	(0.11)	(0.11)	(0.12)	(0.08)	(0.09)	(0.11)	(0.10)	(0.15)	(0.17)	(0.12)
RTA	0.98	0.80	0.82	0.89	0.22	0.97	1.24	−0.21	0.68	0.98	0.70	0.98	0.81	0.88	0.50	0.36	1.22	1.06
	(0.17)	(0.16)	(0.26)	(0.37)	(0.20)	(0.20)	(0.34)	(0.37)	(0.19)	(0.18)	(0.17)	(0.15)	(0.17)	(0.20)	(0.21)	(0.29)	(0.22)	(0.25)
No. of Obs.	1,511,580	1,915,140	853,140	237,180	244,260	421,260	127,440	70,800	3,104,580	428,340	601,800	1,270,860	782,340	2,000,100	831,900	746,940	477,900	669,060

Notes: Standard errors, clustered by source country, are in parentheses. Coefficients reported are multiplied by $-\theta$, as the effects of the independent variable of interest and the trade elasticity are not separately identified by the gravity estimation. The implied percentage effect of each coefficient on the ad valorem tariff equivalent trade cost is $100 \times (e^{-coeff/\theta} - 1)$, where *coeff* is the reported coefficient.

B Data

B.1 Trade Data

Product-level, bilateral trade data is taken from the U.N. Comtrade database classified into six-digit Harmonized System (HS) product codes. For 2003, the database contains trade flow data for 155 reporting countries classified according to the HS1996 classification system.[26] These 155 reporting countries report trade with an additional 74 non-reporting countries and territories. However, to ensure a complete trade flow matrix, only reporting countries are included in the sample.

For pairs of reporting countries, bilateral trade flows are typically reported in both directions by both countries. Trade flows reported by the exporting country were used because these flows are more likely to be consistent with the manufacturing output data, which is reported by the producing country, and because exports are typically reported "free on board", as opposed to "cost, insurance, and freight", and the former is consistent with the measure of trade flows in the model. This results in a dataset of 155 countries, 5,122 product codes, and 4,481,143 non-zero bilateral, product-level trade flow observations.

To combine the trade flow data with manufacturing output data, trade in non-manufacturing HS codes was dropped from the dataset. These are identified using the mapping from HS1996 codes to ISIC (revision 3) codes available from the U.N. Statistics Division.[27] This concordance was developed by the U.N. Statistics Division based on the mapping between the HS1996 classification and the CPC 1.0 classification and the mapping between the CPC 1.0 and the ISIC rev. 3. All HS codes not mapped to ISIC 2-digit industries 15-37 are dropped. This reduces the number of HS codes in the sample to 4,608 and the number of observations to 4,255,517.

B.2 Gravity Variables

The bilateral relationship variables used to estimate trade costs are from the Gravity dataset available from CEPII (see Mayer and Zignago, 2011). The variables used in the estimation are population-weighted distance (*distw*), whether countries share a border (*contig*), whether they share a common official language (*comlang_off*), whether they have ever had a colonial link (*colony*), and whether they are currently members of a common regional trade agreement (*rta*).

B.3 GDP, Labor Force, and Capital Stock Statistics

Data on real GDP, the size of the labor force, and real capital stocks are derived from version 7.1 of the Penn World Tables (Heston et al., 2012). The measure of real GDP used is total PPP-converted

[26] The year was chosen to maximize the number of countries for which both product-level trade data from Comtrade and manufacturing gross output data form INDSTAT were available. Of these 155 reporting countries, 105 originally reported their trade data using the HS2002 system, and the data was subsequently converted to the HS1996 system by Comtrade. To evaluate whether this conversion is likely to have affected the results of this paper, I also conducted the analysis using data for 2001, when nearly all reporting countries reported in the HS1996 system, and that the results were very similar.

[27] This is available for free download from the following url: http://unstats.un.org/unsd/cr/registry/regdntransfer.asp?f=183.

GDP, based on the Geary-Khamis method, at current prices in 2003. The size of the labor force for each country is computed, as in Caselli (2005), as RGDPCH*POP/RGDPWOK in 2003, where RGDPCH is PPP converted GDP per capita, computed using the chain method, at 2005 constant prices; RGDPWOK is PPP converted GDP per worker in the same units; and POP is population.

The real value of countries' capital stock is computed, as in Caselli (2005), using the perpetual inventory method. Real aggregate investment is computed as RGDPL*POP*KI, where RGDPL is PPP converted GDP per capita, computed using the Laspeyres method, at 2005 constant prices, and KI is the investment share of GDP. I assume a depreciation rate of capital of 0.06.

B.4 Value Added Shares

Sectoral value added as a share of total value added and value added as a share of gross output in the manufacturing and non-tradeable sectors is calculated from data obtained from the STAN database available from the OECD for 2003. Manufacturing is defined as ISIC (Rev. 3) industries 15-37, and non-tradables is defined as industries 40-99, which includes electricity, gas, and water supply; construction; wholesale and retail trade; and services.

B.5 Manufacturing Output

Data on gross manufacturing output is is obtained from three sources. Where it is available, the data is taken from the OECD STAN database. For countries not in this database, data is obtained from the Industrial Statistics Database (INDSTAT4), 2011 Edition, CD-ROM available from the United Nations Industrial Development Organization. Where data for 2003 is not available but is available for other years both before and after 2003, the log of 2003 output is taken as the linear interpolation between the values of log output from the most recent year pre- and post-2003. Where no data is available from either of these sources, gross manufacturing output is imputed from total manufacturing value added obtained from the World Development Indicators database of the World Bank. Manufacturing value added is scaled up by a factor of 3.04 based on a cross-sectional regression of gross output on value added with no constant term, which has an R^2 of 0.99.

Industry-level data on gross manufacturing is also obtained from the STAN database, where available, and the INDSTAT4 database, otherwise. Both sources report data using the ISIC Revision 3 system. STAN reports data at the 2-digit industry level, and INDSTAT4 at the 4-digit level. However, in the INDSTAT database, many countries report data using combinations of categories, and many appear to report data for related industries using either one or the other industry code but not both. In addition different countries report data only in more aggregated categories. Because of such issues, the data was aggregated to the 2-digit level, and several 2-digit industries were combined. Table A1 lists the industries that are used, their definitions, the number of 6-digit HS-1996 codes within each industry, and the industry's share in total world manufacturing expenditure. As with the aggregate data, industry-level output data was interpolated for observations for which data was not available for 2003 but was available for years before and after 2003.

B.6 Constructing the Sample

To estimate trade costs, recover technology parameters, and compute the model equilibrium, data on product-level trade flows, total manufacturing output, the size of its labor force and the size of its capital stock are required for each country. Thus, the required data must be available for a country from the Comtrade database, at least one of the STAN, INDSTAT, or WDI databases, and the Penn World Tables for a country to be included in the sample.

Beginning with the 155 countries that make up the sample of product-level trade data, lack of manufacturing output data reduces the sample size to 144 countries. Lack of data from the Penn World Tables further reduces the sample size to 141 countries. To avoid problems related to entrepot trade, China, Hong Kong, and Macao are merged into a single country. There were also several other cases in which there were apparent problems of entrepot trade – i.e. reported exports exceeded reported gross output – which resulted in 7 countries being dropped from the sample.[28] These two steps together reduced the sample to 132 countries. Once the trade and manufacturing data were merged, domestic absorbtion of domestic manufacturing output, X_{ii}, was then calculated as total manufacturing output minus total manufacturing exports to all countries (including non-reporters), and total manufacturing absorbtion, X_i, was calculated as X_{ii} plus total imports from countries in the sample, yielding an internally consistent bilateral trade flow matrix.

In constructing the sample of industry-level output and trade flows, great care was taken to ensure the quality and consistency of the data, which included inspecting the data line-by-line for many countries in the sample. Countries with significant discrepancies, for instance between the sum of industry-level output and reported total output, were excluded from the sample. Even after excluding these countries, for about 12% of observations, reported exports exceeded reported gross output. For these observations, output was imputed based on the value of exports and the country's overall ratio of exports to output for the entire manufacturing sector. When this resulted in an imputed measure of industry-level output that exceeded the reported value by more than 30%, the country was removed from the sample. This resulted in a final sample of 60 countries, 18 manufacturing industries, and 2,360,978 observed product-level bilateral trade flows. The set of countries that make up the aggregate and industry-level samples, along with the source of output data, is reported in Table A2.

C Mathematical Appendix

C.1 Proof of Proposition 2

C.1.1 Case 1

Given that $T_i^k = T_i T^k$, equation (9) implies that $\sum_k (T^k)^{\frac{\sigma-1}{\theta}} = 1$, which is without loss of generality because the general equilibrium admits one normalization over the set of values of T_i^k. This implies

[28] These included Armenia, Belgium, Guyana, Luxembourg, Mali, Mongolia, and Singapore.

that

$$
\begin{aligned}
\Phi_n &= \left(\sum_k \left(\sum_i T_i^k (c_i d_{ni})^{-\theta} \right)^{\frac{\sigma-1}{\theta}} \right)^{\frac{\theta}{\sigma-1}} \\
&= \sum_i T_i (c_i d_{ni})^{-\theta} \left(\sum_k \left(T^k \right)^{\frac{\sigma-1}{\theta}} \right)^{\frac{\theta}{\sigma-1}} \\
&= \sum_i T_i (c_i d_{ni})^{-\theta}.
\end{aligned}
$$

This further implies that $\Phi_n^k = T^k \Phi_n$. Applying these results to (10) gives that

$$
\begin{aligned}
\tilde{T}_{ni} &= \sum_{k=1}^{K} \left(\frac{\Phi_n^k}{\Phi_n} \right)^{\frac{\sigma-1}{\theta}-1} \frac{T_i^k}{T_i} \\
&= \sum_{k=1}^{K} (T^k)^{\frac{\sigma-1}{\theta}} \left(\frac{\Phi_n}{\Phi_n} \right)^{\frac{\sigma-1}{\theta}-1} \frac{T_i}{T_i} \\
&= 1.
\end{aligned}
$$

Substituting these values into (8) and (13) yields the first and third results of Proposition 2. The second result follows by applying the following result, based on (3), to (12):

$$
\begin{aligned}
\pi_{ni}^k &= \frac{T_i^k (c_i d_{ni})^{-\theta}}{\Phi_n^k} \\
&= \frac{T^k}{T^k} \frac{T_i (c_i d_{ni})^{-\theta}}{\Phi_n} \\
&= \pi_{ni}.
\end{aligned}
$$

□

C.1.2 Case 2

Define $\Omega_i^k = \{k | T_i^k > 0\}$. Given that $\frac{T_i^k}{\sum_k T_i^k} \in \{0, 1\}$, for all i and j, equation (4) implies that

$$
\Phi_n^k = \begin{cases} T_i^k (c_i d_{ni})^{-\theta} & \text{if } k \in \Omega_i^k \\ 0 & \text{otherwise.} \end{cases}
$$

This implies that

$$\begin{aligned}
\Phi_n &= \left(\sum_k \left(\sum_i T_i^k (c_i d_{ni})^{-\theta}\right)^{\frac{\sigma-1}{\theta}}\right)^{\frac{\theta}{\sigma-1}} \\
&= \left(\sum_i (c_i d_{ni})^{-(\sigma-1)} \sum_{k\in\Omega_i^k} \left(T_i^k\right)^{\frac{\sigma-1}{\theta}}\right)^{\frac{\theta}{\sigma-1}} \\
&= \left(\sum_i T_i^{\frac{\sigma-1}{\theta}} (c_i d_{ni})^{-(\sigma-1)}\right)^{\frac{\theta}{\sigma-1}},
\end{aligned}$$

where the last equality results from the definition of T_i in (9). This implies that

$$\begin{aligned}
\tilde{T}_{ni} &= \sum_{k=1}^{K} \left(\frac{\Phi_n^k}{\Phi_n}\right)^{\frac{\sigma-1}{\theta}-1} \frac{T_i^k}{T_i} \\
&= \frac{(c_i d_{ni})^{\theta-(\sigma-1)}}{T_i \Phi_n^{\frac{\sigma-1}{\theta}-1}} \sum_{k\in\Omega_i^k} \left(T_i^k\right)^{\frac{\sigma-1}{\theta}} \\
&= \left(\frac{T_i}{\Phi_n}\right)^{\frac{\sigma-1}{\theta}-1} (c_i d_{ni})^{\theta-(\sigma-1)}.
\end{aligned}$$

Substituting these values into (8) yields the first result of Proposition 2. The second result follows because, if $k \in \Omega_i^k$, then, according to (3), $\pi_{ni}^k = 1$ and $X_{nn}^k = 0$. Thus, (12) implies that $\varepsilon_{ni} = -(\sigma-1)$. The third result follows from rearranging (8).

□

C.1.3 Case 3

If $\theta = \sigma - 1$, then, from (9), $T_i = \sum_k T_i^k$. This implies that $\tilde{T}_{ni} = 1$. Substituting this value into (8) and (13) yields the first and third results of Proposition 2. The second result holds trivially

□

C.2 Product-Varying Trade Costs and the Elasticity Index

Suppose that trade costs vary by product and take the form $d_{ni}^k = d_{ni} d_n^k$, for $n \neq i$, and $d_{nn}^k = 1$.[29] I will discuss the model's predictions for the patterns depicted in Figure 1 and Table 1 under each of the special cases of Proposition 2 to demonstrate that this form of product-varying trade costs, without non-trivial patterns of comparative advantage, cannot produce the patterns observed in the data.

Under cases 1 and 2, the value of EI_{ni} is unaffected by this form of variation in trade costs

[29] Note that if $d_{nn}^k = d_n^k$ for all n and k, this specification would be isomorphic to the baseline model.

across products. In case 1, it is straightforward to show that

$$\frac{X_{ni}^k}{M_n^k} = \frac{T_i(c_i d_{ni})^{-\theta}}{\sum_{i \neq n} T_i(c_i d_{ni})^{-\theta}} = \frac{X_{ni}}{M_n},$$

which implies that $\mathrm{EI}_{ni} = 0$, for all n and i. In case 2, $\pi_{ni}^k \in \{0,1\}$, regardless of the value of d_n^k, which implies that $\mathrm{EI}_{ni} = 1$, for all n and i.

In case 3, arbitrary patterns of comparative advantage may exist, but they do not affect aggregate trade flows. Thus, EI_{ni} can potentially take on any value between zero and one, but, in the baseline setup, the coefficient on $\ln(\mathrm{dist}_{ni}) \times \tilde{\mathrm{EI}}_{ni}$ reported in Table 1 should be equal to zero. So, what remains to be shown is that this form of variation in trade costs cannot cause variation in aggregate trade flows such that a positive coefficient would be estimated even though $\theta = \sigma - 1$. To see this, first note that, given this form of trade costs, aggregate trade flows continue to follow (8), except that

$$\tilde{T}_{ni} = \sum_{k=1}^{K} (d_n^k)^{-\theta} \left(\frac{\Phi_n^k}{\Phi_n} \right)^{\frac{\sigma-1}{\theta}-1} \frac{T_i^k}{T_i},$$

and $\Phi_n^k = \sum_i T_i^k (c_i d_{ni} d_n^k)^{-\theta}$. In case 3,

$$\tilde{T}_{ni} = \sum_{k=1}^{K} (d_n^k)^{-\theta} \frac{T_i^k}{T_i},$$

which implies that any variation in π_{ni}/π_{nn} associated with patterns of comparative advantage is orthogonal to bilateral trade costs. So, once country-specific factors and distance are controlled for in the regression reported in Table 1, the estimated coefficient on $\ln(\mathrm{dist}_{ni}) \times \tilde{\mathrm{EI}}_{ni}$ should be zero.