UNSW Business School

Working Paper

Never Stand Still | Business School | Economics

UNSW Business School Research Paper No. 2015 ECON 24

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*

October, 2015

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 relax this assumption, developing a framework in which to analyze the role of interactions among countries' patterns of comparative advantage in determining the aggregate effects of trade barriers. My 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 my 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 trade barriers. A key finding is that the welfare gains from trade tend to be larger and more skewed in favor of low-income countries than an aggregate model would suggest.

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

Keywords: International Trade; Welfare; Composition; Product Level; Comparative Advantage; Trade Barriers; Gravity; Trade Cost Elasticity; Developing Countries

*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. Some of the content of this paper has been adapted from an unpublished paper previously circulated under the title "The Composition of Exports and Gravity"

1 Introduction

The workhorse quantitative models of international trade imply that the aggregate effects of trade barriers and the welfare gains from trade 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 desirable feature that the amount of data required to make predictions regarding aggregate variables – such as income, welfare, and trade flows – is quite low. As Arkolakis et al. (2012) have shown, for a large class of such models, the welfare gains from trade are a function of only the share of domestic goods in aggregate expenditure and the elasticity of bilateral trade flows with respect to variable 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 require the implicit assumption 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 these restrictions, developing a model with the flexibility to allow for arbitrary patterns of comparative advantage across products for every country, while maintaining much of the analytical tractability of aggregate models. I show that these patterns of product-level comparative advantage can interact in non-trivial ways to influence the effects of trade barriers on aggregate bilateral trade flows and welfare. Using data on product-level bilateral trade flows, I find countries' patterns of comparative advantage are quantitatively important in determining aggregate bilateral trade flows and the welfare gains from trade.[2] For example, I find that one quarter of the variation in aggregate bilateral trade flows and two thirds of the average country's welfare gains from trade relative to autarky are related to countries' patterns of comparative advantage. In addition, I find that, after taking these patterns into account, the welfare gains from trade are significantly larger and more skewed in favor of developing countries than an aggregate model would conclude, with the welfare gains being more than twice that of an aggregate model for the average non-OECD country.

The model I employ is an extension of the Ricardian trade model of Eaton and Kortum (2002) – henceforth EK. As in the EK model, there is a continuum of product varieties, and international trade occurs due to countries' idiosyncratic differences in productivity across varieties. However, I allow countries' expected productivity to differ across product categories into which varieties are grouped, in contrast with the EK model, for which every variety is ex-ante identical. This setup maintains much of the analytical tractability of the EK model, while also allowing for any pattern of product-level comparative advantage to be incorporated into the model.

[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). For the sake of brevity, in the remainder of the paper, I refer to this class of models as "aggregate trade models".

[2] In particular, I use data from the UN Comtrade database at 6-digit level of Harmonized System, which includes bilateral trade flows within more than 4,500 manufactured product categories.

My model provides a succinct way to summarize and quantify the strength of a basic Ricardian force that is absent from standard aggregate quantitative trade models. Specifically, 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). Except in some very special cases in which my model collapses to an aggregate model, aggregate trade flows from i to n depend on the strength of i's product-level comparative advantage in n, vis-à-vis the rest of the world. By contrast, the aggregate models delineated by Arkolakis et al. (2012) all assume that there is either no scope for comparative advantage across products or that each country completely specializes in a unique set of products. Under such restrictions, trade barriers have no effect on the relative prices of an exporter's products in any market, so the elasticity of aggregate trade flows with respect to trade costs is constant and identical for every bilateral country pair.

In the presence of non-trivial patterns of comparative advantage, the welfare effects of trade barriers are also non-homogeneous. The trade cost elasticity differs across country pairs and depends on countries' patterns of product-level comparative advantage, which implies that the welfare effects of trade barriers also depend on these patterns. In my model, the magnitude of this effect is fully summarized by an endogenous, country-specific term which measures the effect of a country's comparative advantage on its domestic trade share. This captures the insight that, 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. Further, as external trade barriers fall, a country's domestic trade share will fall relatively slowly if its product-level comparative advantage is relatively strong, despite the fact that it benefits relatively more from specializing in its comparative advantage products. Thus, the tight link between the domestic trade share and welfare of aggregate models is broken.

I use data on product-level trade flows to infer countries' patterns of product-level comparative advantage and consider how they influence the welfare effects of trade barriers under several counterfactual scenarios. As trade barriers fall, the model predicts that countries with relatively strong patterns of comparative advantage will specialize more fully in the production of their comparative advantage products. In the case of the welfare gains from trade relative to autarky, this implies that countries whose domestic trade flows are concentrated in relatively few products experience greater gains from trade. It turns out that this tends to be the case for low-income countries.

I also consider the welfare effects of the growth of Chinese exports and find that the gains from trade are highly dependent on the similarities of countries' patterns of comparative advantage with China's in foreign markets. By contrast, an aggregate model predicts that the gains from China's growth are driven by countries' geographical proximity to China because, if the relative prices of countries' exports are assumed to be affected uniformly, countries benefit from the lower prices of Chinese exports in proportion to the share of their expenditure devoted to Chinese goods.

This paper builds on the previous literature which uses quantitative trade models to determine the effects of trade barriers on aggregate bilateral trade flows, income, and welfare – including

Eaton and Kortum (2002), Anderson and van Wincoop (2003), Alvarez and Lucas (2007), and Helpman et al. (2008) – and more recent papers that address discrepancies between more traditional quantitative trade models and the data.[3] The main contribution of my paper to this literature is that it demonstrates how the workhorse class of quantitative trade models can be generalized to account for the aggregate effects of non-trivial patterns of product-level comparative advantage. It does so in a way that maintains, to a large extent, the tractability and parsimony of this class of models while utilizing the wealth of information contained in product-level trade data, which is available for most of the world's countries. It also provides succinct and intuitive expressions relating the gains from trade to countries' patterns of product-level comparative advantage, allowing for a straightforward decomposition of the gains from trade into across-product and within-product components.

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 we 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. My 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. I also show that, when the patterns that exist in the data are taken into account using a more general framework, the role of such patterns is quantitatively important.

A recent branch of the literature, to which my paper is highly complementary, is focused on the effects of trade barriers in multi-sector models. Most closely related are Caliendo and Parro (2015) and Levchenko and Zhang (2014), which take into account sectoral heterogeneity at the industry level in measuring the gains from trade.[4] Although my model shares many features with the models in these papers, my paper is distinct in two important ways. First, because I focus on the effects of patterns of comparative advantage at a level of aggregation much lower than the broad sectors considered by these papers, I relax the assumption of constant sectoral expenditure shares.[5] This implies that the welfare formulas they derive, based on sectoral domestic trade shares, are no longer valid for arbitrary changes in trade barriers. Instead, I show that the aggregate effects of potentially complex patterns of comparative advantage can be summarized by a single additional term in otherwise standard expressions for aggregate trade flows and welfare. Second, because neither product-level domestic expenditure shares nor this comparative advantage term

[3]Examples of the latter include Waugh (2010), Fieler (2011), and Caron et al. (2014). 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. Costinot and Rodríguez-Clare (2014) review recent advances in this literature in measuring the welfare gains from international trade relative to autarky.

[4]Arkolakis et al. (2012) also derive an expression for welfare in a multi-sector extension to their aggregate framework. Other notable recent papers that consider the effects of trade barriers in multi-sector models include Anderson and Yotov (2011), Costinot et al. (2012), and Chen and Novy (2011).

[5]Levchenko and Zhang (2015) also utilize a multi-sector model with endogenous sectoral expenditure shares. However, their focus is on estimating sectoral productivity for 19 manufacturing sectors. Thus, they do not derive expressions for aggregate trade flows and welfare, nor do they consider how patterns of comparative advantage shape the effects of trade barriers on these variables.

are directly observable in the data, I demonstrate how these values, and thus the aggregate effects of product-level comparative advantage, can be inferred from data on product-level international trade flows.

To make the distinction between my model of product-level comparative advantage and multi-sector trade models clear, I extend my model to a multi-sector environment with constant sectoral expenditure shares and in which trade costs are allowed to vary by sector. The basic result is that accounting for product-level comparative advantage within the sectors of a multi-sector model has much the same implications as in the single-sector version, and if anything, the effects of product-level comparative advantage are quantitatively larger in the multi-sector specification. Thus, even in models that feature sectoral heterogeneity, it is quantitatively important to account for product-level patterns of comparative advantage. In addition, modelling product-level comparative advantage as in my model involves relaxing some of the restrictions imposed by aggregate and multi-sector trade models, and the data required to quantify the effects of product-level comparative advantage does not typically limit the sample of countries that can be studied. As a result, it is straightforward to incorporate these effects into such environments.

In Section 2, I outline the basic theoretical framework for this analysis. In Section 3, I derive the basic results regarding the role of patterns of product-level comparative advantage in shaping the effects of trade barriers on aggregate trade flows and welfare, first via a simple example and then for the full model. In Section 4, I use product-level trade data to estimate trade barriers and infer countries' patterns of productivity across products. Section 5 describes the main quantitative results concerning the aggregate implications of countries' patterns of comparative advantage. Section 6 presents the results of additional counterfactual experiments which highlight the differences between the predictions of my model and a standard aggregate trade model. The implications of extending the model to a multi-sector environment are presented in Section 7, and Section 8 concludes.

2 A Model of Product-Level Comparative Advantage and Trade

Here, I present my baseline model of trade which is driven by countries' patterns of comparative advantage across a large number of products. 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 multi-sector extension is considered in Section 7.

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, each endowed with k_i units of capital and one unit of labor, which is supplied inelastically. Each household obtains utility from the consumption of a non-tradeable final good.

2.1 Production

The tradeable goods sector consists 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). 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 of 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 $\eta^k > \sigma$ 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 single-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 multi-sector extension in Section 7, trade costs are allowed to vary across sectors, as in Caliendo and Parro (2015) and Levchenko and Zhang (2015).[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

[6] The assumption of a continuum of varieties of each product is purely for analytical convenience. With a finite number of varieties, all the results that follow hold in expectation. The assumption that $\eta^k > \sigma$ – i.e., that varieties within a product category are more substitutable than are varieties in different product categories – is not mathematically necessary with a finite number of products. However, it simplifies the intuition for the key results that follow, and appears to be consistent with the data (see, e.g., Broda and Weinstein, 2006).

[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.

cost of delivering one unit of the variety from i to n is given by

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

where w_i is the wage in i, r_i is the rental rate of capital, P_i is the price of the composite tradeable good, and $\tilde{\alpha} = (\alpha^\alpha(1-\alpha)^{1-\alpha}\beta)^{-\beta}(1-\beta)^{\beta-1}$.

I assume that $Z_i^k(\omega)$ is a random variable that follows a Fréchet (type-II extreme value) distribution:

$$F_i^k(z) = e^{-T_i^k(\bar{\Gamma}^k z)^{-\theta}}. \quad (2)$$

The form of (2) can be motivated by assuming that $Z_i^k(\omega)$ can be decomposed into a deterministic component and an idiosyncratic component as follows: $\ln\left(Z_i^k(\omega)\right) = \ln\left(Z_i^k\right) + \varepsilon_i^k(\omega)$, where $\varepsilon_i^k(\omega)$ is distributed Gumbel (type-I extreme value), as in McFadden (1974).[8]

The expected value of $Z_i^k(\omega)$ is increasing in T_i^k, which reflects the overall state of technology in i for producing all varieties of k, and the variance of $Z_i^k(\omega)$ across varieties of k in i is inversely proportional to θ. Thus, the average level of T_i^k over all products reflects country i's overall absolute advantage, variation in T_i^k across products reflects i's deterministic pattern of comparative advantage across products, and θ governs the degree of comparative advantage across varieties of a each product due to idiosyncratic productivity differences.

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. Thus, 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

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

where

$$(P_n^k)^{-\theta} = \sum_{i=1}^N T_i^k(c_i d_{ni})^{-\theta}, \quad (4)$$

which is well-defined if $\theta > \eta^k - 1$, which I maintain henceforth.

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

[8] The constant $\bar{\Gamma}^k = \Gamma(1-(\eta^k-1)/\theta)^{-\frac{1}{\eta^k-1}}$, 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 $\bar{\Gamma}^k$ plays is in the mapping between relative productivity across products and relative sales, and it is irrelevant to the analysis of this paper.

come from i. The price index P_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. The price index is lower if n has access – through low trade barriers – to sources with relatively high levels of productivity (larger T_i^k) and low input costs.

Equation (3) also highlights the role of the within-product comparative advantage parameter, θ, in the determination of trade flows. 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 relatively low costs of producing and delivering that variety. Thus, θ governs the importance of these costs in determining the allocation of expenditure by n across potential sources of product k.

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 aggregate price index is given by

$$P_n^{1-\sigma} = \sum_{k=1}^{K} \left(P_n^k\right)^{1-\sigma}. \tag{6}$$

Recalling that the price index P_n^k embodies the worldwide state of technology for producing varieties of k from the perspective of an importer in n, (5) implies that expenditure is relatively higher on products which are produced relatively efficiently in the source countries which can deliver products to n relatively cheaply.

Aggregate Trade Flows Combining equations (3) - (6) 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*

$$\pi_{ni} \equiv \frac{X_{ni}}{X_n} = \frac{T_i(c_i d_{ni})^{-\theta}}{P_n^{-\theta}} \tilde{T}_{ni} \tag{7}$$

where

$$T_i = \left(\sum_{k=1}^{K} (T_i^k)^{\frac{\sigma-1}{\theta}}\right)^{\frac{\theta}{\sigma-1}} \tag{8}$$

and

$$\tilde{T}_{ni} = \sum_{k=1}^{K} \left(\frac{P_n^k}{P_n} \right)^{\theta-(\sigma-1)} \frac{T_i^k}{T_i}. \tag{9}$$

Equation (7) 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, the lower are production costs in i, and the lower are trade costs between i and n.

The term $\tilde{T}_{ni}$ embodies the effect of the interaction among countries' patterns of product-level comparative advantage on aggregate trade flows. To understand how this interaction is embodied in $\tilde{T}_{ni}$, note that the price index P_n^k plays two roles in allocating n's expenditure. First, according to (3), a higher value of P_n^k, for a given value of T_i^k, implies that n devotes a higher share of its expenditure to varieties of k produced in i, with an elasticity equal to θ, because it implies that i is more likely to be the low-cost producer of a given variety of k. Second, according to (5), a higher value of P_n^k implies that n allocates relatively less of its total expenditure to product k, with an elasticity equal to $(\sigma - 1)$. If $\theta > \sigma - 1$, the former effect dominates, and i exports relatively more to n if T_i^k is positively correlated with P_n^k – that is, if country i is relatively productive for products that country n cannot purchase relatively cheaply from other sources (including domestic producers in n).[9] Thus, $\tilde{T}_{ni}$ measures the degree to which n's total expenditure on products from i is enhanced by the strength of i's pattern of comparative advantage across products.

Proposition 1 implies 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.

2.3 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:

$$w_i L_i = \sum_{n=1}^{N} \pi_{ni} w_n L_n. \tag{10}$$

Alvarez and Lucas (2007) show that (10) 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.

[9] This parameter restriction is implied by the assumption that $\eta^k > \sigma$ along with the restriction that $\theta > \eta^k$, which was necessary to ensure that P_n^k is well-defined. If it were not the case, 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. This would also imply the counterintuitive result that an increase in T_i^k causes an increase in $X_{ni'}^k$, for $i' \neq i$. Moreover, according to the analysis the follows, this restriction seems to be consistent with the data.

Table 1: Values of Key Terms in Simple Example

	Case 1	Case 2
$P_h = P_f$	$T^{-\frac{1}{\theta}}(1+d^{-\theta})^{-\frac{1}{\theta}}$	$T^{-\frac{1}{\theta}}(1+d^{1-\sigma})^{\frac{1}{1-\sigma}}$
$\tilde{T}_{hh} = P_{ff}$	1	$(1+d^{1-\sigma})^{\frac{\theta}{\sigma-1}-1}$
$\tilde{T}_{hf} = \tilde{T}_{fh}$	1	$(1+d^{\sigma-1})^{\frac{\theta}{\sigma-1}-1}$
$\pi_{hh} = \pi_{ff}$	$\frac{1}{1+d^{-\theta}}$	$\frac{1}{1+d^{1-\sigma}}$
$\pi_{hf} = \pi_{fh}$	$\frac{d^{-\theta}}{1+d^{-\theta}}$	$\frac{d^{1-\sigma}}{1+d^{1-\sigma}}$

3 Aggregate Implications of the Composition of Trade Flows

In this section, I utilize the model developed in Section 2 to explore the implications of countries' patterns of comparative advantage for the effects of trade barriers on aggregate trade flows and welfare. Before discussing general results, I provide a simple 2-country example to illustrate how these patterns shape bilateral trade flows and the gains from trade.

3.1 A Simple Example

Suppose that there are two symmetric countries (h and f), two products (1 and 2), and symmetric trade costs ($d_{hf} = d_{fh} \equiv d$). Further, suppose that $\alpha = 1$ and $\beta = 1$, so that labor is the only factor of production. Now consider two cases. In the first, expected productivity levels are the same in both countries and both sectors: $T_h^1 = T_h^2 = T_f^1 = T_f^2 \equiv T/2^{\theta/(\sigma-1)}$. In the second case, each country can only produce one of the two products: $T_h^1 = T_f^2 \equiv T$ and $T_f^1 = T_h^2 = 0$, an extreme form of comparative advantage. I have constructed the values of the T's so that, in both cases, $T_h = T_f = T$. Because the countries are symmetric, $w_h = w_f$ in equilibrium, and I normalize their values to unity.

Table 1 lists the expressions for the price level, $\tilde{T}$ terms, and aggregate trade shares under the two cases. In the first case, which is representative of an aggregate EK model, neither country has a comparative advantage in either product, so the $\tilde{T}$ terms are equal to 1. The countries trade only because of idiosyncratic productivity differences across varieties within each product category. Thus, their trade shares depend only on the level of trade barriers, d, and the elasticity θ, which governs the dispersion of productivity across varieties and the strength of idiosyncratic, within-product comparative advantage.

In the second case, the countries trade due to comparative advantage across products, the strength of which is measured by the $\tilde{T}$ terms. This case highlights several important properties of $\tilde{T}_{ni}$. First, it is weakly greater than unity – i.e., comparative advantage is trade-enhancing. Second, $\tilde{T}_{hh}$ is decreasing in d, while $\tilde{T}_{hf}$ is increasing in d – i.e., the forces of comparative advantage ameliorate the effect of trade barriers, which is to shift expenditure toward domestic sources. In this case of extreme comparative advantage, aggregate trade shares depend only on the level of trade barriers and the elasticity $(\sigma - 1)$, which is governed by the elasticity of substitution across products in demand.

It may seem counterintuitive that $\tilde{T}_{hh} \geq 1$, implying that domestic trade flows are enhanced by comparative advantage. However, it is important to note that $\tilde{T}_{ni}$ measures the strength of country i's comparative advantage in serving market n, *relative to every other source country*, which depends on the correlation between T_i^k and P_n^k, rather than that between T_i^k and T_n^k. In autarky, $\tilde{T}_{ii} = 1$, which confirms the intuition that a country cannot have a comparative advantage in any product relative to itself, while with frictionless trade, the values of P_n^k, and thus $\tilde{T}_{ni}$, are equalized across destinations. Both of these features are easily verified in case 2.

Finally, note that the aggregate price level is weakly lower in case 2 compared to case 1, with the two being equal in autarky. Because wages have been normalized to 1, welfare in these cases is equal to the inverse of the price level. Thus, given that $\theta > \sigma - 1$, the welfare gains from trade will be larger the greater is the strength of countries' product-level comparative advantage. This result continues to hold even when aggregate trade flows are held constant, which requires d to be larger in case 2 than in case 1. In case 1, $P = T^{-1/\theta}\pi_{ii}^{1/\theta}$, while in case 2, $P = T^{-1/\theta}\pi_{ii}^{1/(\sigma-1)}$. This demonstrates that the welfare gains from trade depend not only on how much countries trade with the rest of the world but also on the patterns of comparative advantage across products that give rise to the observed aggregate trade flows.

3.2 Special Cases of the Full Model

The two cases from the simple example, while informative of the role of $\tilde{T}_{ni}$ in measuring the strength and effect of countries' patterns of comparative advantage, turn out to correspond to two special cases of the full model. In these cases, the partial elasticity of relative trade flows with respect to trade costs, $\varepsilon_{ni} \equiv \partial \ln(\pi_{ni}/\pi_{nn})/\partial \ln d_{ni}$, is constant and equal for all countries, and given this elasticity, the effect of trade barriers on welfare can be summarized by the aggregate domestic expenditure share (π_{ii}). Proposition 2 formalizes this observation.

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

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

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

$$\pi_{ni} = \frac{T_i^{\frac{\rho}{\theta}} (c_i d_{ni})^{-\rho}}{P_n^{-\rho}}; \tag{11}$$

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

In case 1, like its counterpart in the simple example, there is no comparative advantage across products. Trade occurs only because of idiosyncratic productivity differences across producers within product categories. Thus, the model reduces to an EK model. Case 2 is the opposite extreme in which each country produces a unique set of products. As result, there is no intra-product trade, $\tilde{T}_{ni} = \left(T_i(c_i d_{ni}/P_n)^{-\theta}\right)^{(\sigma-1)/\theta-1}$, and trade costs only affect the relative price of one country's products versus another's, so the trade cost elasticity is equal to $1 - \sigma$. In this case, the model closely resembles that of Krugman (1980) or an Armington model, such as Anderson and van Wincoop (2003). In case 3, the parameters align such that the product category to which a variety belongs is irrelevant, and the model collapses to an EK model.

In all three of these special 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 (constant) trade cost elasticity. Conversely, if any of the assumptions underlying the special cases is violated, trade barriers have heterogeneous effects across country pairs, and the welfare gains from trade cannot be inferred from observable aggregate variables alone. The quantitative exercises that follow utilize product-level trade flow data to assess the degree to which the effect of trade barriers on aggregate trade flows and the gains from trade depend on patterns of comparative advantage that deviate from the assumptions of these special cases.

3.3 Comparative Advantage and Aggregate Trade Flows

I now demonstrate how countries' patterns of comparative advantage shape the effects of trade barriers on aggregate trade flows when the world economy is not restricted to satisfy the assumptions of Proposition 2. Consider the partial elasticity of $\tilde{T}_{ni}$ with respect to d_{ni} (holding constant production costs):

$$\frac{\partial \ln \tilde{T}_{ni}}{\partial \ln d_{ni}} = [\theta - (\sigma - 1)] \frac{1}{\pi_{ni}} \sum_{k=1}^{K} \frac{X_n^k}{X_n} (\pi_{ni}^k - \pi_{ni})^2. \tag{12}$$

The summation term is the variance of i's product-level trade shares in n, weighted by the products' shares of total expenditure in n. This expression gives us a useful and intuitive measure of the strength of a country's comparative advantage – which is market-specific – and demonstrates its role in ameliorating the effect of trade barriers on bilateral trade flows.[10]

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

$$\begin{aligned}\varepsilon_{ni} \equiv \frac{\partial \ln (\pi_{ni}/\pi_{nn})}{\partial \ln d_{ni}} &= -\theta + \frac{\partial \ln \tilde{T}_{ni}}{\partial \ln d_{ni}} - \frac{\partial \ln \tilde{T}_{nn}}{\partial \ln d_{ni}} \\ &= -\theta + [\theta - (\sigma - 1)] \sum_{k=1}^{K} \frac{X_n^k}{X_n} \left(\frac{(\pi_{ni}^k)^2}{\pi_{ni}} - \frac{\pi_{ni}^k \pi_{nn}^k}{\pi_{nn}} \right).\end{aligned} \tag{13}$$

The summation term is equivalent to the difference between the variance of π_{ni}^k, relative to π_{ni}, and the covariance of π_{ni}^k and π_{nn}^k, relative to π_{nn}. This implies that the stronger is i's comparative advantage in n (and the less correlated are i's and n's patterns of comparative advantage), the closer to zero is the trade cost elasticity of i's exports to n. In general, $\varepsilon_{ni} \in [-\theta, 1 - \sigma]$. Thus, cases 1 and 2 of Proposition 2 represent limiting cases in terms of the effect of trade barriers on aggregate trade flows.

Intuitively, the trade cost elasticity depends on the strength of i's comparative advantage because, as d_{ni} increases, the varieties that n imports 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's comparative advantage in n is relatively strong, it will have a relatively large market share in the products which it tends to export intensively, and the increase in d_{ni} 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. Conversely, if i's comparative advantage in n is relatively weak, then a change in d_{ni} will not affect the relative prices of products, and the share of varieties for which i is the low-cost producer falls for every product with an elasticity equal to θ, which largely determines the trade cost elasticity.

This variation of the trade cost elasticity across country pairs highlights the way in which this model generalizes 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, under the macro-level restriction that this elasticity is constant and equal across countries.

3.4 Aggregate Productivity and the Gains from Trade

Countries' patterns of comparative advantage also affect the welfare gains from trade. Note that welfare in this model is equivalent to real final output (GDP) per worker, $y_i \equiv Y_i/L_i$. Using (7), along with the optimality conditions of the firms' cost-minimization problems, and abstracting from

[10] Levchenko and Zhang (2014) use the coefficient of variation of $T_i^k / \sum_i T_i^k$ as a heuristic measure of the strength of a country's comparative advantage. Equation (12) constitutes a theoretically-founded measure that is similar in spirit but is country-pair specific and based on equilibrium trade shares, not the underlying technology parameters alone.

constant terms that are common across countries, a country's welfare can be expressed as

$$y_i = \left(\frac{T_i}{\pi_{ii}}\right)^{\frac{1-\gamma}{\beta\theta}} \tilde{T}_{ii}^{\frac{1-\gamma}{\beta\theta}} k_i^{1-\alpha}. \tag{14}$$

As in Waugh (2010), this expression resembles that for real GDP per worker based on a standard neoclassical growth model as y_i depends on the capital-labor ratio and a term representing total factor productivity. Like in a standard EK model, TFP depends on the aggregate productivity index, T_i, and the domestic trade share, π_{ii}, and is increasing the level of dispersion in idiosyncratic productivity (smaller θ) . The departure from an EK model lies in the presence of $\tilde{T}_{ii}$.

In autarky, when π_{ii} and $\tilde{T}_{ii}$ are equal to unity, TFP is determined by solely T_i. As in a standard EK model, international trade increases productivity by giving producers of the composite intermediate good access to additional sources for each variety of each product. Since π_{ii} is an expenditure-weighted average of the share of varieties for which i is the low-cost producer – i.e., $\pi_{ii} = \sum_k \frac{X_i^k}{X_i}\pi_{ii}^k$ – the inverse of this measure represents the average reduction in the cost of obtaining the varieties that are produced more efficiently abroad.

Given the value of π_{ii}, trade increases productivity by more if i has a greater degree of comparative advantage across products, vis-à-vis the rest of the world, in its domestic market. This effect is embodied in the value of $\tilde{T}_{ii}$. In autarky, $(P_n^k)^{-\theta} = T_i^k c_i^{-\theta}$, which is perfectly correlated with T_i^k, and $\tilde{T}_{ii} = 1$. As trade barriers fall, as long as T_i^k is not perfectly correlated for all countries – as in case 1 of Proposition 2 – the correlation between $(P_n^k)^{-\theta}$ and T_i^k weakens, and since $\theta > \sigma - 1$, the value of $\tilde{T}_{ii}$ rises, according to (9). This results in an increase in TFP because it implies that P_n^k falls faster for the products for which it was highest in autarky, which lowers the cost of the composite intermediate good because products are imperfect substitutes.[11]

This effect is further illustrated by rewriting (9), using (3) to substitute for the equilibrium product-level price indexes, P_n^k:

$$\tilde{T}_{ii} = \left[\sum_{k=1}^{K} \left(\frac{T_i^k}{T_i}\right)^{\frac{\sigma-1}{\theta}} \left(\frac{\pi_{ii}^k}{\pi_{ii}}\right)^{1-\frac{\sigma-1}{\theta}}\right]^{\frac{\theta}{\sigma-1}},$$

This expression highlights the way in which $\tilde{T}_{ii}$ embodies the effect of countries' patterns of comparative advantage in shaping the gains from trade. The value of $\tilde{T}_{ii}$, and thus TFP, is greater the more able is a country to concentrate its production in the products for which it is relatively most productive. This is the case if, compared to autarky, its domestic trade shares (and prices) fall by more for the products for which it is relatively unproductive.

[11] Here, the restriction that $\theta > \sigma - 1$ ensures that products are not so substitutable that a decrease in trade barriers causes demand for i's comparative disadvantage products to increase by so much that domestic production of these products increases, lowering TFP.

3.5 Comparison to Other Multi-Sector Models

Arkolakis et al. (2012), in a multi-sector extension to a standard EK model, derive an expression for the welfare gains from trade based on sectoral domestic trade shares. While (14) expresses welfare in terms of the aggregate domestic trade share and the strength of a country's comparative advantage, embodied in $\tilde{T}_{ii}$, it is straightforward to derive an expression, similar to that of Arkolakis et al. (2012), relating changes in welfare to changes in product-level domestic trade shares:

$$\begin{aligned} d\ln y_i &= -\frac{1-\gamma}{\beta\theta}\sum_{k=1}^{K}\frac{X_i^k}{X_i}d\ln\pi_{ii}^k \\ &= -\frac{1-\gamma}{\beta\theta}\left(\sum_{k=1}^{K}\frac{X_i^k}{X_i}\left[d\ln\pi_{ii}^k - d\ln\pi_{ii}\right] + d\ln\pi_{ii}\right) \\ &= \frac{1-\gamma}{\beta\theta}\left(d\ln\tilde{T}_{ii} - d\ln\pi_{ii}\right) \end{aligned} \tag{15}$$

The key distinction between the first line of (15) and the expression derived by Arkolakis et al. (2012) is that this expression is only analytically tractable for marginal changes in product-level trade shares and welfare, while theirs is valid for an arbitrary shock. This is because the upper-level aggregator is CES in my model, which implies that X_n^k/X_n is not constant and depends on equilibrium prices. By contrast, Arkolakis et al. (2012) and other notable multi-sector models, such as Caliendo and Parro (2015) and Levchenko and Zhang (2014), assume this aggregator to be Cobb-Douglas.

However, relaxing the Cobb-Douglas assumption is important given my primary focus on the effects of patterns of comparative advantage across products. Because these products are defined at a much lower level of aggregation than the sectors of most multi-sector models, the assumption of a unitary elasticity of substitution across products is much less plausible. For example, Levchenko and Zhang (2014) utilize data on 19 manufacturing sectors, in contrast to the 4,608 products in the dataset I employ. Further, while product-level data on *international* trade flows is available for a large set of countries, comparably disaggregated data on domestic trade flows is not available for most countries. So, while an expression such as the first line of (15) can be quite useful in measuring the effects of trade flows in the absence of systematic patterns of comparative advantage or in the presence of productivity differences across only broad sectors – i.e., the level of aggregation for which data on domestic trade shares is generally available for many countries – in what follows, I make use of the form of (14), instead, because it is globally valid as well as due to its relative parsimony, tractability, and intuitive content, even though it involves a term, $\tilde{T}_{ii}$, that is not directly observable.

The last two lines of (15) demonstrate the connection between the expression for the welfare gains from trade based on product-level trade shares and (14), which is based on the value of $\tilde{T}_{ii}$. Changes in $\tilde{T}_{ii}$ summarize the effect of relative changes in product-level trade shares on aggregate welfare. The welfare gains from a fall in trade barriers are larger, relative to a standard EK model, the greater the fall in π_{ii}^k in sectors for which X_i^k is relatively small, which is the case for products for which P_i^k is relatively high. This accords with the intuition from the previous section.

4 Quantitative Implementation

Having established that countries' patterns of product-level comparative advantage can be important in shaping the effect of trade barriers on aggregate trade flows and welfare, I use my model to evaluate the quantitative importance of this effect. In this section, I describe the procedure used to recover the model's production function, technology, and trade cost parameters from the data.

I parameterize both the full model of Section 2 as well as a version of the model restricted to conform with case 1 of Proposition 2. In the restricted case, the model reduces to a single-sector EK 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 product-level model, all differences in relative productivity across countries and varieties of a given product are purely idiosyncratic. Since this assumption is more plausible the more narrowly-defined is a product in the data, I map a product in the model into a 6-digit Harmonized System (HS-6) product category in the international trade data. This is the most disaggregated classification for which bilateral trade data is available for a large number of countries. The product-level trade flow dataset, obtained from the U.N. Comtrade database, comprises bilateral trade flows for 132 countries in 4,608 HS-6 manufactured product categories for 2003. Appendix B provides a detailed description of the data employed in this section.

4.1 Parameter Values

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

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, which governs the degree of dispersion in productivity across varieties of a product, 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 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 (7) 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 Appendix E, I explore the implications of using different values of σ and θ.

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 $\gamma = 0.65$.

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:

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

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.[12]

In the special cases of Proposition 2, aggregate bilateral trade flows follow a standard aggregate gravity equation, based on (11), 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 (7). Fortunately, (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 data is required to identify the trade costs associated with international borders, which account for a very large share of trade costs.

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:

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

where

$$
\ln \tilde{d}_{ni}^{-\theta} \equiv -\theta \ln(d_{ni}/d_i) = -\theta \left[dist_{ni}^m + bord_{ni} + lang_{ni} + col_{ni} + rta_{ni}\right], \tag{18}
$$

and where M_n^k and E_i^k are product-level imports and exports to and from the rest of the world, and $E^k = \sum_i E_i^k$.[13] I have defined $\tilde{d}_{ni}$ inclusive of the elasticity θ to make clear that, as in most gravity models, the parameters of (18) cannot be separately identified from the value of θ.

Importantly, (17) makes clear that the product-level indexes, $\tilde{P}_n^k$ and $\tilde{\Psi}_i^k$, which Anderson and

[12] 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.

[13] The term $\tilde{P}_n^k$ is the analogue of P_n^k calculated over all countries except n. Thus, $(P_n^k)^{-\theta} = (\tilde{P}_n^k)^{-\theta} + T_n^k c_n^{-\theta}$.

van Wincoop (2003) refer to as multilateral resistance terms, can be calculated as functions of only observable data (which does not include domestic trade flows) and trade costs. Thus, (17) and (18) define product-level trade flows as a nonlinear function, $X_{ni}^k = g(\mathbf{I}_{ni}, \mathbf{M}_n^k, \mathbf{E}_i^k; \mathbf{\Omega})$, of the parameters of (18), $\mathbf{\Omega}$; the vector of indicators of bilateral relationships in (18), $\mathbf{I}_{ni}$; and the vectors of product-level total imports and exports, $\mathbf{M}_n^k$ and $\mathbf{E}_i^k$.

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. Because (17) expresses bilateral trade flows conditional on total imports and exports, 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 (18), via structural Poisson pseudo-maximum likelihood. Specifically, this involves the following procedure:

1. Choose an initial set of parameter values, $\mathbf{\Omega}_0$.
2. Given $\mathbf{\Omega}_0$ and data on bilateral relationships, $\mathbf{I}_{ni}$, and product level imports and exports, $\mathbf{M}_n^k$ and $\mathbf{E}_i^k$, compute $\hat{X}_{ni}^k = g(\mathbf{I}_{ni}, \mathbf{M}_n^k, \mathbf{E}_i^k; \mathbf{\Omega}_0)$.
3. Repeat step 2, searching over the set of possible values of $\mathbf{\Omega}$ until the value, $\hat{\mathbf{\Omega}}$, which maximizes the Poisson likelihood function, is found.[14]

I chose the Poisson likelihood function due to the attractive properties discussed in Santos Silva and Tenreyro (2006) and Gourieroux et al. (1984) and its current widespread use in gravity estimations. In addition, French (2014) shows that, when aggregate data is used rather than product-level data, due to the adding-up properties of Poisson PML described by Arvis and Shepherd (2013) and Fally (2015), 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.

4.3 Recovering Technology Parameters and Border Costs

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

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

At this point, given estimates of $\tilde{d}_{ni}$ and $\tilde{T}_i^k (c_i d_i)^{-\theta}$, 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

[14] Specifically, $\hat{\mathbf{\Omega}}$ is the value of $\mathbf{\Omega}$ for which following first order conditions are satisfied: $\sum_{n,i,k} g_{\mathbf{\Omega}}(\mathbf{I}_{ni}, \mathbf{M}_n^k, \mathbf{E}_i^k; \hat{\mathbf{\Omega}})(X_{ni}^k - \hat{X}_{ni}^k)/\hat{X}_{ni}^k = 0$, where $g_{\mathbf{\Omega}}(\cdot)$ is the gradient of $g(\cdot)$ with respect to $\mathbf{\Omega}$.

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.

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

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.

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

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 (10) 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 (6) and capital-labor ratios from the data, domestic production costs are given by

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

where $\tilde{\beta} = (\alpha\beta)^{-\beta}(1-\beta)^{\beta-1}$, which uses the fact that, due to the Cobb-Douglas production functions for final output and intermediate varieties, in equilibrium, $r_i = (1-\alpha)w_i/(\alpha k_i)$.

4.4 Trade Cost Estimates

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. 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 also reported.

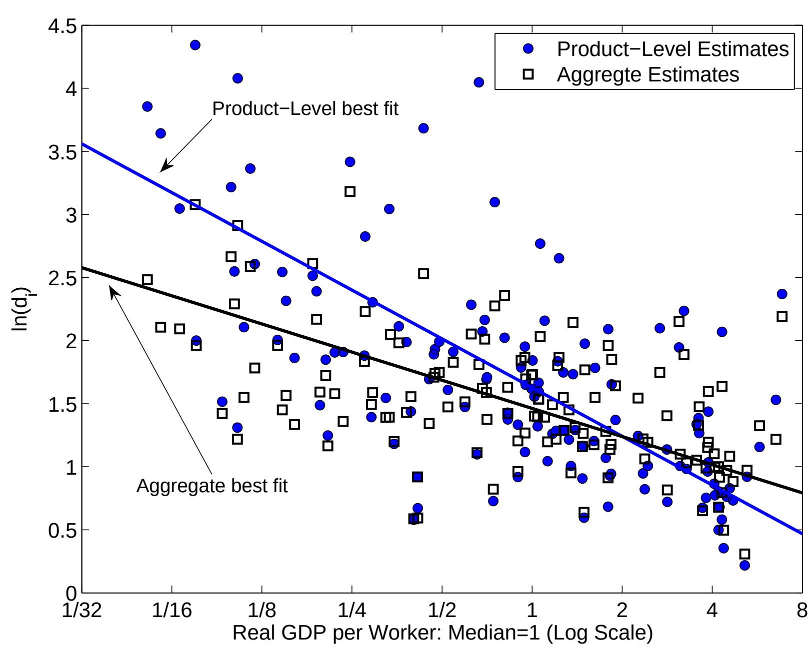
The coefficient estimates based on aggregate and product-level data are similar, although the product-level distance coefficients are greater in absolute value. The largest difference is in the median level of exporter-specific border costs. The intuition for this is straightforward. Because the

Table 2: Trade Cost Coefficient Estimates

	Coefficients		Percentage Effect	
	Aggregate	Product Level	Aggregate	Product Level
median($\ln(d_i^{-\theta})$)	−6.04	−9.30	337	371
625 - 1,250 km	−0.30	−0.40	7.70	6.98
	(0.13)	(0.12)		
1,250 - 2,500 km	−0.65	−0.82	17.22	14.87
	(0.17)	(0.17)		
2,500 - 5,000 km	−1.01	−1.24	28.08	23.20
	(0.21)	(0.22)		
5,000 - 10,000 km	−1.80	−2.25	55.21	45.86
	(0.20)	(0.20)		
>10,000 km	−1.93	−2.59	60.10	54.48
	(0.23)	(0.22)		
Shared Border	0.54	0.56	−12.32	−8.91
	(0.09)	(0.09)		
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)	(0.09)		
RTA	0.81	0.84	−17.88	−13.23
	(0.09)	(0.08)		
No. of Obs.	17,292	79,681,536		
Value of θ			4.1	6.0

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.

Figure 1: Border Costs Estimates



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 1 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.

5 The Importance of the Composition of Trade Flows

In this section, I explore the quantitative implications of the patterns of comparative advantage inferred from the product-level trade data for aggregate bilateral trade flows and the welfare gains from trade. I show in Appendix E that the product-level model outperforms the aggregate model in accounting for the cross-country distribution of income and prices of tradeable goods.

5.1 Aggregate Trade Flows

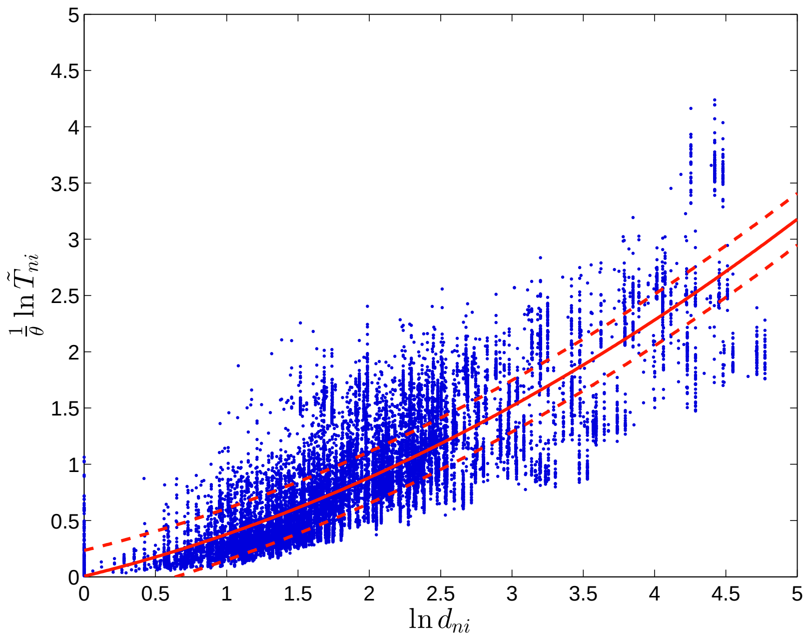
In the product-level model, the effects of trade barriers between i and n are ameliorated by the strength of i's product-level comparative advantage in n, which is measured by the $\tilde{T}_{ni}$ term in (7). The log-linear form of this expression provides a particularly straightforward way to measure the importance of $\tilde{T}_{ni}$ in explaining the cross-sectional variation in bilateral trade flows. Consider a regression of the form:

$$\ln \pi_{ni} = \xi_0 + \xi_1 \ln T_i c_i^{-\theta} - \xi_2 \ln P_n^{-\theta} + \xi_3 \ln d_{ni}^{-\theta} + \xi_4 \ln \tilde{T}_{ni} + \varepsilon_{ni}$$

Because this is based on an identity, this regression yields an R^2 of unity. Performing the same regression but omitting the $\tilde{T}_{ni}$ term yields an R^2 of 0.755, indicating that 24.5 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.

There is a great deal of heterogeneity in the value of $\tilde{T}_{ni}$ across country pairs. To illustrate this, Figure 2 plots the value of $\tilde{T}_{ni}$ for each country pair against their bilateral trade costs. For comparability with trade costs, the values of $\ln \tilde{T}_{ni}$ are normalized by θ so that the magnitudes are equivalent in terms of their effect on bilateral trade flows. The aggregate model restricts all values of $\ln \tilde{T}_{ni}$ to be equal to zero. At the opposite extreme, if all of the points were aligned with the 45

Figure 2: Trade Costs and $\tilde{T}_{ni}$



Note: Quadratic curve of best fit (solid line) and interquartile prediction interval (dashed lines) shown. The exports of Mauritania lie outside of the figure range.

degree line, it would imply that the forces of comparative advantage are so strong that trade costs are irrelevant.[15]

Consistent with (12), $\tilde{T}_{ni}$ is greater for country pairs for which d_{ni} is relatively large. This means that the fraction of the effect of trade costs on bilateral trade flows that is offset by countries' patterns of comparative advantage tends to be greater for countries with relatively large trade costs. The slope of the best-fit curve is always less than unity, implying that higher trade costs are never associated with *greater* trade flows, but it is convex, indicating that the effects of the largest bilateral trade barriers tend to be offset to a relatively greater extent by the strength of countries comparative advantage. It is also clear from both the visible dispersion of points and the interquartile prediction interval plotted in Figure 2 that there there is a great deal of heterogeneity in the effect of product-level comparative advantage across country pairs with similar levels of bilateral trade costs.

5.2 The Welfare Gains from Trade

In the aggregate model, the gains from trade relative to autarky are given by $\pi_{ii}^{-\frac{1-\gamma}{\beta\theta}}$, while they also depend on patterns of product-level comparative advantage, embodied in $\tilde{T}_{ii}$, in the product-level model. Thus, we can compare the two models' predictions for the gains from trade by examining these two quantities.

[15] This would occur if the world were characterized by case 2 of Proposition 2 and if $\sigma = 1$, which would imply that each country produced a unique set of products to which every country devoted a constant share of expenditure.

Figure 3: Gains From Trade: Aggregate vs. Product-Level Model

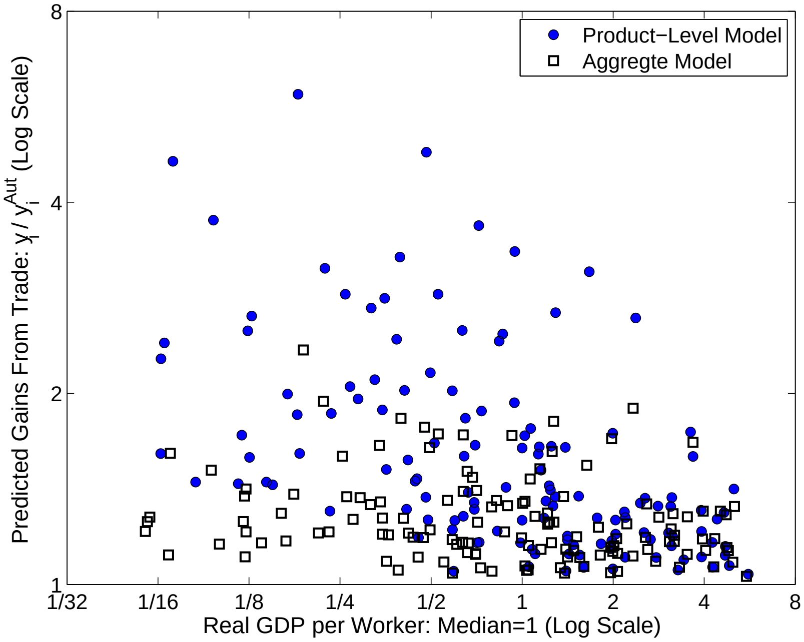


Figure 3 depicts the gains from trade relative to autarky for each model. Because both models match the data on aggregate domestic trade shares, this component of the predicted gains from trade is identical. The predictions differ for two reasons. First, the aggregate model assumes no across-product comparative advantage, i.e. $\tilde{T}_{ii} = 1$, so it will automatically predict no gains from trade arising from such. Second, 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.

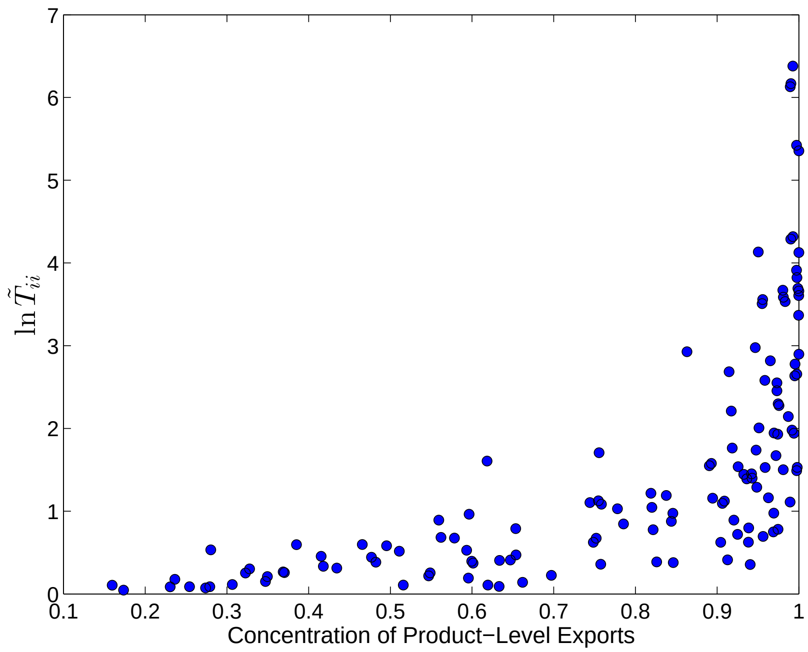
As is clear in Figure 3, the gains from trade predicted by the product-level model are generally larger and are more skewed in favor of low-income countries. For example, for high income OECD countries, the average gain in $\ln y_i$ compared to autarky is 0.21 in the product-level model compared to 0.19 in the aggregate model, while the average gains for all other countries are 0.52 and 0.25, respectively.[16] For only 13 mostly large, developed countries, including the US, Germany, UK, France, and Korea, are the gains from trade smaller in the product-level than in the aggregate model. 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 (14) allows for a simple decomposition of the gains from trade into across-product and within-product components, i.e.

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

[16]High-income is defined as in the World Bank's World Development Indicators database.

Figure 4: Concentration of Exports and $\tilde{T}_{ii}$



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 3 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, as trade costs rise, $\tilde{T}_{ii}$ decreases 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 a relatively small set of products are predicted to find autarky more costly.

To get a sense of the features of the data that give rise to the pattern evident in Figure 3, Figure 4 plots the values of $\tilde{T}_{ii}$ against a measure of the concentration of countries' product-level total exports, which is equal to the share of a country's exports made up of the set of products, which make up 5% of world trade flows, for which it has the highest market share.[17] Though ad hoc, this measure is highly correlated with countries' values of $\tilde{T}_{ii}$. This is consistent with the intuition that countries with relatively strong patterns of comparative advantage benefit relatively more from international trade because they have greater scope for concentrating their production in the products for which they are relatively most productive. In the product-level model, the strength of a country's comparative advantage is manifested in the concentration of its domestic trade flows. In addition, the pattern of domestic trade flows is related to the pattern of product-level exports because both depend on countries' relative values of T_i^k. Thus, the model infers that countries with relatively concentrated exports will have higher gains from trade. It turns out that this tends to be the case for low-income countries.

5.3 Trade Cost Elasticities

The product-level model deviates from the class of models delineated by Arkolakis et al. (2012), for which welfare depends only on aggregate variables, primarily in that it relaxes the macro-level restriction that the elasticity of bilateral trade flows with respect to trade costs is constant across bilateral country pairs. Thus, it is useful to examine the way in which the product-level trade data deviates from the special cases of Proposition 2, under which my model is consistent with the these aggregate trade models. To this end, I define the following Trade Elasticity Index, based on (12), which can be calculated directly from the data:

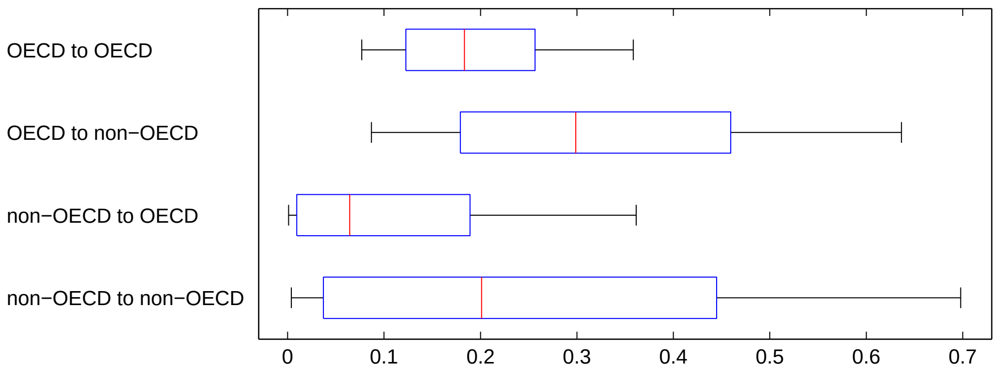
$$\mathrm{TEI}_{ni} = \frac{\frac{M_n}{X_{ni}} \sum_{k=1}^{K} \frac{M_n^k}{M_n} \left(\frac{X_{ni}^k}{M_n^k} - \frac{X_{ni}}{M_n} \right)^2}{1 - \frac{X_{ni}}{M_n}}$$

This index is equal to the weighted variance of i's product-level shares of n's imports, scaled to lie in the interval $[0, 1]$ and can be interpreted as a measure of the strength of i's comparative advantage in increasing i's exports to n by ameliorating the effect of the trade barriers between n and i. It differs somewhat from (12) due to the lack of product-level data on domestic trade flows and because of the normalization. TEI is equal to zero for every country pair if the world is characterized by case 1 of Proposition 2, in which $\varepsilon_{ni} = -\theta$ for every country pair, and is equal to one in case 2, in which $\varepsilon_{ni} = 1 - \sigma$. The model predicts that country pairs with a larger value of TEI will have a trade cost elasticity that is closer to zero.

Figure 5 depicts the distribution of TEI over all country pairs with positive trade flows, grouped by whether the source or destination is a high-income OECD country. The figure demonstrates 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 TEI value, while OECD imports tend to have a lower

[17] Formally, this measure is equal to $\sum_{k \in \Phi_i} E_i^k / E_i$, where $\Phi_i = \{k : \sum_{k \in \Phi_i} E^k / E = 0.05$ and $E_i^k / E^k > E_i^{k'} / E^{k'}, \forall k' \notin \Phi_i\}$. It is analogous to the concentration ratio commonly employed in the international organization literature, where in place of a fixed number of firms, a fixed share of the set of products, in terms of value, is used. Unlike other common measures of concentration, such as the Herfindahl or Theil index, this measure is robust to arbitrary merging and division of product categories.

Figure 5: The Trade Elasticity Index by Country Group (Data)



Note: The median (vertical line), interquartile rang (box width), and interdecile range (whiskers) are depicted for each group of bilateral country pairs.

TEI value. 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. In addition, the value of TEI tends to be asymmetric – higher when the source country is a developed country – which is consistent with the finding that trade barriers affect the aggregate trade flows, and hence the gains from trade, of developed and developing countries differently.

As a test of whether these patterns in the product-level data covary with aggregate bilateral trade flows in the way predicted by the model, I consider the following direct, if somewhat crude, evidence. Consider the expression, based on (7), for n's relative imports from i:

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

This value is a function of source-specific and destination-specific terms, bilateral trade costs, and $\tilde{T}_{ni}$. In the special cases of Proposition 2, only country-specific terms and trade costs affect relative imports.[18] However, in general, bilateral trade flows also depend on countries' patterns of comparative advantage, embodied in $\tilde{T}_{ni}$.

Because the form of TEI is derived from the expression for $\partial \ln \tilde{T}_{ni}/\partial \ln d_{ni}$, given in (12), this measure can be utilized as part of the following linear approximation of $\tilde{T}_{ni}$, which is a function only of observable, product-level trade flow data:

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

where $\widetilde{\mathrm{TEI}}_{ni} = (1 - X_{ni}/M_n) \times \mathrm{TEI}_{ni}$ and $\tilde{T}_{ii}^{\mathrm{FT}}$ is the value of $\tilde{T}_{ni}$ under frictionless trade, using the fact that, with frictionless trade, $\tilde{T}_{ni}$ does not vary by destination because prices are equal everywhere.

[18] In the first two cases, TEI_{ni} is constant across county pairs, so the relationship between TEI_{ni} and normalized import shares would be degenerate. In the third case, TEI_{ni} may vary across country pairs, but it would have no effect on aggregate trade flows.

Table 3: Correlation Between Trade Flows and Elasticity Index (Data)

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

Using this approximation of $\ln \tilde{T}_{ni}$, the expression for relative imports takes the following form:

$$\ln\left(X_{ni}/X_{nn}\right) = \xi_i + \xi_n - \theta \ln d_{ni} + [\theta - (\sigma - 1)] \ln(d_{ni}) \widetilde{\text{TEI}}_{ni} + \delta_{ni},$$

where δ_{ni} includes the approximation error for $\ln \tilde{T}_{ni}$. To evaluate the effect of $\ln \tilde{T}_{ni}$ on bilateral trade flows, 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 $\widetilde{\text{TEI}}_{ni}$. 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 of Proposition 2. Table 3 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, as the product-level model predicts.

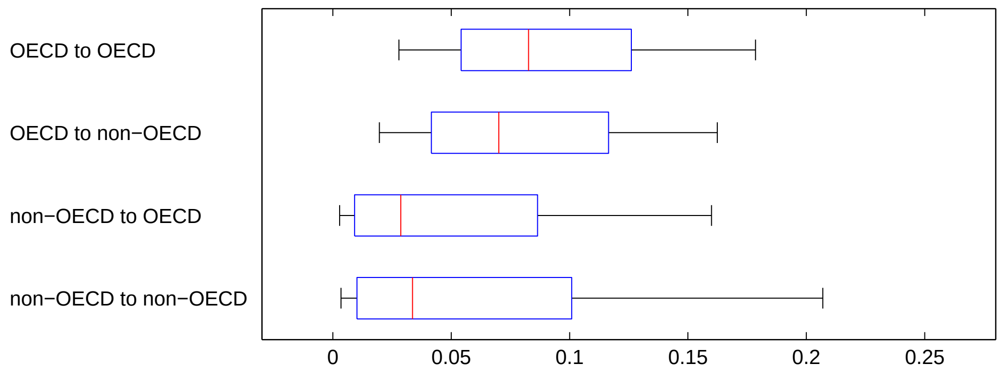
Because the approximation of $\ln \tilde{T}_{ni}$ ignores the contributions of trade costs between other pairs of countries, and given that non-distance-related bilateral trade barriers are ignored, this relationship could also be influenced by omitted variables.[19] However, there is no ex ante reason to suspect that these variables would be correlated with $\widetilde{\text{TEI}}_{ni}$. Therefore, the relationship evident in Table 3 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.[20]

It may seem surprising that TEI tends to be smaller for low-income exporters, while $\tilde{T}_{ii}$ tends to be larger for these countries. The former indicates that, in the cross-section, developing countries' patterns of comparative advantage are relatively weak in offsetting the effects of bilateral trade costs on their aggregate exports. The latter implies that these countries' product-level comparative advantage is relatively strong in that lowering trade barriers leads them to become relatively more specialized in producing the products for which they are relatively productive, causing them to experience greater gains from trade. The reason these two results hold simultaneously is that low-

[19] I use only distance to proxy for trade costs both to keep the exercise as simple as possible and so that it continues to rely only on observable data.

[20] There are other possible explanations for this relationship based on factors not modelled in either the product-level model or standard aggregate trade models. For example, if trade costs were allowed to vary by importer, exporter, and product, it would be possible to construct any pattern of product-level and aggregate trade flows, so it could be mere coincidence that the data fits the predictions of the product-level model. However, Appendix D shows that trade barriers of the form $d^k_{ni} = d_{ni} d^k_n$, consistent with tariffs conforming to the Most Favored Nation rule of the WTO, cannot generate the relationship evident in Table 3. Thus, it would take a very special set of product-specific trade costs to generate these patterns in the data in the absence the effects predicted by the product-level model.

Figure 6: The Trade Elasticity Index by Country Group (Model Prediction)



Note: The median (vertical line), interquartile rang (box width), and interdecile range (whiskers) are depicted for each group of bilateral country pairs.

income countries tend to have similar profiles of T_i^k to one another as well as relatively high variance in productivity across products. By contrast, high-income countries tend to have profiles that are relatively unique but also less extreme. 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, being the only country that does not face trade barriers in the domestic market, they face little competition from other low-income countries in their comparative advantage products, in which they specialize relatively heavily, while they benefit from access to their comparative disadvantage products, which tend to be imported from high-income countries.

5.4 Product-Level Trade Flows: Data vs. Model

The product-level model relaxes the assumptions regarding patterns of comparative advantage that are necessary for aggregate trade models to be able to safely ignore the information present in product-level trade flow data. However, it maintains many of the restrictions necessary for both models to remain highly tractable, notably identical CES demand systems for all countries and bilateral trade costs that are constant across products.[21] As a result, the product-level model does not perfectly match the full matrix of product-level 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 5 and Table 3 using the values of product-level trade flows predicted by the model. Figure 6 is the analogue of Figure 5, and Table 4 is the corresponding analogue of Table 3.

Overall, the product-level model does well in replicating these patterns. As in the data, the elasticity index tends to be smaller when the source is a non-OECD country, although the predicted values tend to be smaller than those in the data for all sets of countries.[22] Table 4 shows that the

[21]The extension in Section 7 allows trade costs to vary across sectors.

[22]The latter phenomenon appears to be mostly due to deviations in the data from expenditure patterns predicted by the model's CES demand system. For example, countries tend to import more of their comparative advantage products than the model predicts, perhaps due to a more complex input-output structure or non-homotheticities in demand which interact with patterns of comparative advantage.

model almost exactly matches the measured effect of comparative advantage, proxied by $\tilde{\text{EI}}_{ni}$, on the relationship between trade barriers and aggregate trade flows.[23] Keeping in mind that the aggregate model predicts no variation in TEI and thus zero correlation between TEI and aggregate trade flows, this indicates that the product-level model, despite maintaining most of the restrictions of the aggregate model, goes quite far in accurately predicting the patterns of product-level trade flows relevant for understanding the role of product-level comparative advantage in shaping aggregate trade flows and the gains from trade.

Table 4: Correlation Between Trade Flows and Elasticity Index (Model Prediction)

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	

6 Additional Counterfactual Experiments

In addition to comparing the gains from trade relative to autarky, it is possible to consider the effects of other counterfactual changes in trade barriers. In this section, I consider two counterfactual experiments involving global reductions in trade barriers, which allow for comparison with the results of similar experiments based on aggregate quantitative trade models in the literature. I also consider the effects of the growth of Chinese exports, which highlights the heterogeneous effects of trade barriers, which depend on countries' patterns of comparative advantage and which are ruled out by standard aggregate trade models.

6.1 Reductions in Trade Barriers

The first experiment is the elimination of all asymmetric trade barriers – i.e., setting $d_{ni} = \min\{d_{ni}, d_{in}\}$ – which Waugh (2010) argues likely reflects policy-related barriers, given that most natural impediments to trade, such as distance, are symmetric by nature. The second experiment is a move to entirely frictionless trade ($d_{ni} = 1$). The effects of these changes are also compared to the effects of moving to complete autarky, which is the reciprocal exercise to measuring the gains from trade, described above.

Table 5 reports the effects of the changes in trade barriers on the cross-country distribution of income per worker in both models. For each scenario, the table lists the average change in log income per worker across all countries as well as two measures of the dispersion of income across

[23] 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.

Table 5: Counterfactual Income per Worker: Aggregate vs. Product-Level Models

Variable	Model	Baseline	Autarky	$\min(d_{ni}, d_{in})$	$d_{ni}=1$
$\text{mean}(\Delta \ln y_i)$	Aggregate	–	-0.23	0.36	1.58
	Product-Level	–	-0.46	0.52	1.60
$\text{var}(\ln y_i)$	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 6: Decomposition of Changes in Income per Worker

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

countries, the variance of log income per worker and the ratio of the 90th to the 10th percentile. As with the gains from trade, I also decompose the changes into those driven by across-product and within-product comparative advantage – i.e., those related to changes in $\tilde{T}_{ii}$ and π_{ii}, respectively – which is reported in Table 6.[24]

The results of these experiments are all consistent with the basic finding 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), for 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 trade in the baseline case relative to autarky. The average change in log income is significantly higher 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.

[24]Specially, using (14), $\text{mean}(\Delta \ln y_i)$ is decomposed into $(1/N)\sum_i \kappa \Delta \ln \tilde{T}_{ii}$ and $-(1/N)\sum_i \kappa \Delta \ln \pi_{ii}$, while $\Delta\text{var}\,(\ln y_i)$ is decomposed into $[\kappa^2 \text{var}(\Delta \ln \tilde{T}_{ii}) + 2\kappa \text{cov}(\ln y_i, \Delta \ln \tilde{T}_{ii}) - \kappa^2 \text{cov}(\Delta \ln \tilde{T}_{ii}, \Delta \ln \tilde{\pi}_{ii})]$ and $[\kappa^2 \text{var}\,(\Delta \ln \tilde{\pi}_{ii}) - 2\kappa \text{cov}(\ln y_i, \Delta \ln \tilde{\pi}_{ii}) - \kappa^2 \text{cov}(\Delta \ln \tilde{T}_{ii}, \Delta \ln \tilde{\pi}_{ii})]$, where $\kappa = \frac{1-\gamma}{\beta\theta}$.

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.

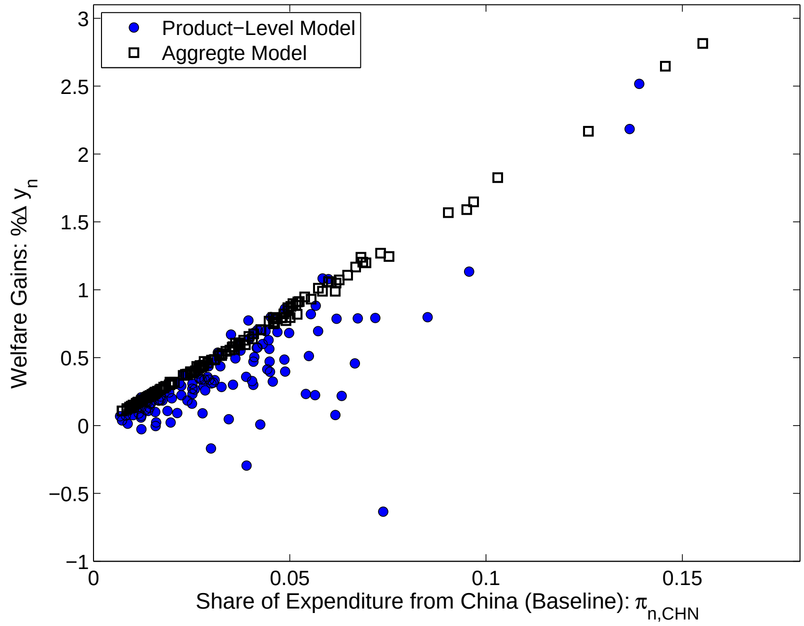
A similar line of reasoning is behind the differences between the two models in the gains from eliminating all trade barriers. Because the product-level model has generally lower estimated symmetric trade barriers, the additional gains from removing these barriers, above the gains from removing border costs, is predicted to be lower, while the additional gain to low-income countries is also relatively lower. Even though 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 6 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 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.

Another way to interpret this result is to consider the differences in the models' predictions for the total possible gains from trade and where the current state of the world lies within that range. Because the product-level model takes into account the gains from trade due to across-product comparative advantage, the possible gains from moving from autarky to frictionless trade are larger than in the aggregate model. Because the gains due to across-product comparative advantage 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 than the aggregate model would suggest, and the elimination of asymmetric trade barriers would move the world much closer to frictionless trade.[25]

Finally, it is worth noting that the summary statistics reported in Tables 5 and 6 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.

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

Figure 7: Welfare Gains from Fall in Chinese Export Costs



6.2 The Growth of Chinese Exports

The counterfactual experiments considered so far have involved simultaneous changes in the trade barriers for all countries. However, considering the effect of a change in the trade barriers faced by a single country highlights the differences between the aggregate and product-level models, as the latter allows for changes in worldwide relative prices according to the product-level comparative advantage of the affected country, leading to more heterogeneous effects across countries. As an example, I consider the welfare effects of a doubling of Chinese exports as a share of world trade flows, which is approximately what occurred between 1999 and 2005. Specifically, for each model, I compare the welfare gains of moving to the baseline model equilibrium from the counterfactual equilibrium in which China's export costs have been increased by the amount necessary to halve China's exports as a share of world trade flows.

Figure 7 plots the welfare gains from the growth of Chinese exports against China's share of domestic expenditure in each country, $\pi_{n,\mathrm{CHN}}$. A fall in China's export costs causes the prices of China's products to fall everywhere. In the aggregate model, this causes each country's price index, P_n, to simply fall in proportion to China's share of domestic expenditure, so the welfare gains are strongly associated with this measure. By contrast, in the product-level model, product-level price indexes, P_n^k, fall by more for China's comparative advantage goods. This implies that the welfare gains depend not just on the importance of China in the countries' domestic markets but also the similarity between their patterns of comparative advantage to China's both domestically and in foreign destination markets.

Figure 8 plots the welfare gains associated with the growth of Chinese exports on world maps. Figure 8(a) shows that the welfare gains in the aggregate model are driven by its gravity-like structure. China's share of domestic expenditure, and hence the welfare gains, are greatest for the countries that are relatively close, such as Southeast Asia, and relatively small (in terms of output), such as eastern Africa.

From Figure 8(b), it is clear that, in the product-level model, geography plays a smaller role in determining the welfare gains. As with the gains from trade relative to autarky, a large share (51%) of the average change in $\ln y_i$ is associated with changes in $\ln \tilde{T}_{ii}$ in the product-level model, and the share tends to by higher for developing countries. However, the differences in the welfare gains between the two models (depicted in Figure 8(c)) are largely driven by differences in the changes in π_{ii} in response to the growth of Chinese exports. Whereas in the aggregate model, the change in π_{ii} is driven primarily by China's effect on domestic prices, in the product-level model, China's effect on prices in other destination markets has a substantial effect on π_{ii}.

To better understand this effect, consider the elasticity of $\tilde{T}_{ni}$ with respect to trade costs between n and a third country, j:

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

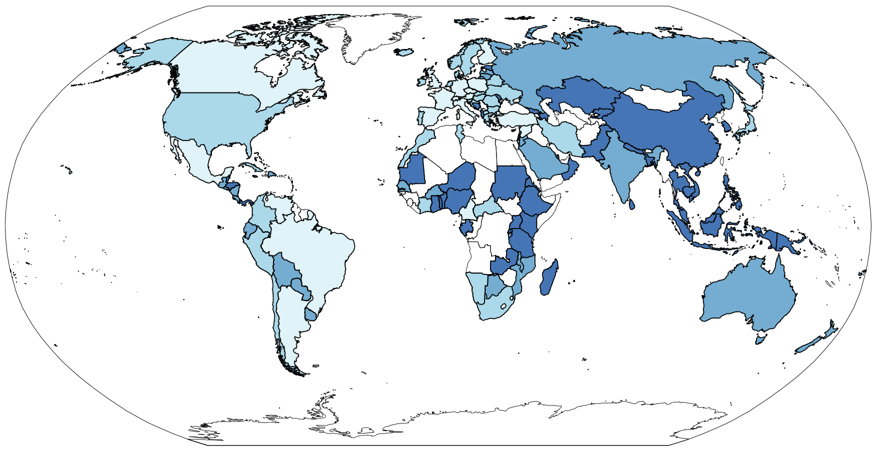
The summation term is the weighted covariance of the product-level trade shares of countries i and j in market n. If a country's product-level market shares are positively correlated with China's, it will see the strength of its comparative advantage in foreign markets erode as China's export costs fall. In terms of equation (14), this effect manifests itself in a smaller change in π_{ii} in the product-level model than in the aggregate model. This is the result of two forces. First, the fall in foreign demand for the country's products leads to a fall in domestic wages. Second, because China will have a relatively small market share in the country's comparative advantage products in the domestic market, compared with foreign markets, $\tilde{T}_{ii}$ will fall by relatively less. The result is that π_{ii} falls by less, or even rises, in the product-level model relative to the aggregate model.

Figure 8(c) highlights the role of product-level comparative advantage in determining the gains from the growth of Chinese exports, apart from the effect of the standard gravity variables that drive the welfare gains in the aggregate model. The countries whose welfare gains are smallest relative to the aggregate model are the countries of Southeast Asia, Central America, West Africa, southeastern Europe, and the Baltic states, which all tend to specialize in the production of similar products as China. Five of these countries (Bangladesh, Cape Verde, Cambodia, Morocco, and Tunisia) actually experience a decrease in welfare due to the growth of Chinese exports. The countries whose welfare gains are relatively large in the product-level model include highly developed countries, which experience an increase in foreign demand for their comparative advantage products, and eastern African countries that benefit from disproportionately large decreases in the domestic prices of their comparative disadvantage products.

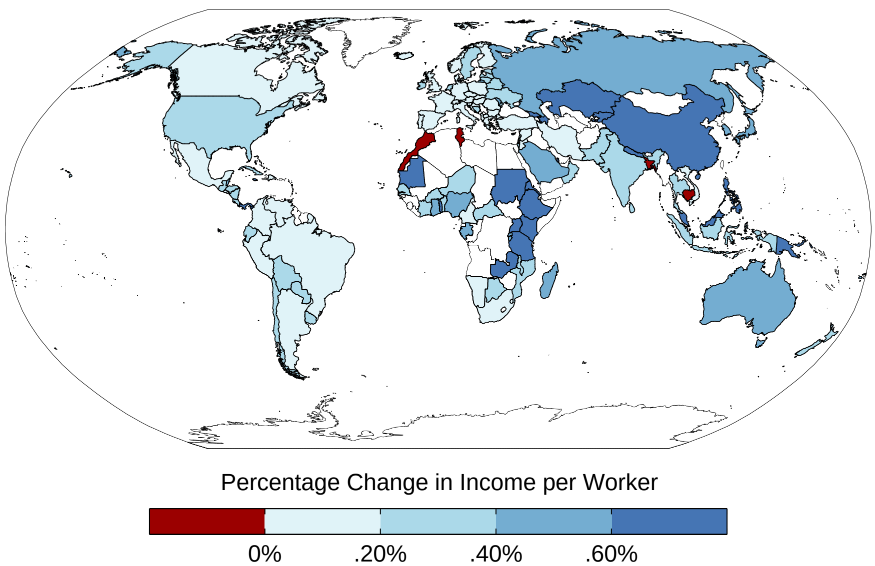
It is also interesting to note that China experiences greater welfare gains in the product-level model despite the fact that it necessarily experiences large declines in $\tilde{T}_{ni}$ for all foreign destinations, reflecting a worsening terms of trade for its comparative advantage products. However, this effect,

Figure 8: Welfare Gains from Fall in Chinese Export Costs

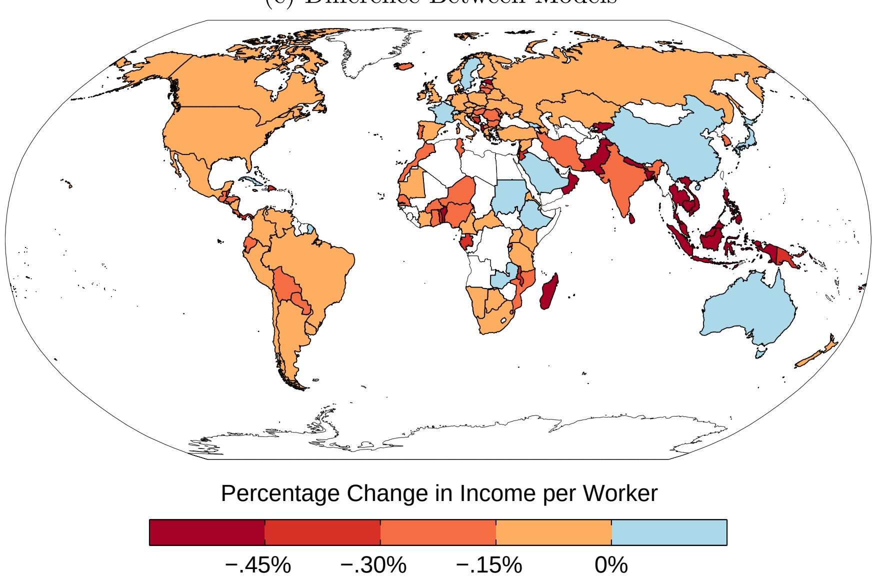
(a) Aggregate Model



(b) Product-Level Model



(c) Difference Between Models



which dampens China's gains from increased trade, is more than offset by the increase in $\tilde{T}_{ii}$, which reflects the fact that the increase in domestic wages induces China to specialize more heavily in the products for which it is relatively most productive.

7 Product-Level Comparative Advantage in a Multi-Sector Model

In Appendix C, I extend my model to allow for multiple manufacturing sectors, or "industries", in similar fashion to the models of Caliendo and Parro (2015) and Levchenko and Zhang (2015). 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 to examine the cross-country income distribution is compromised.

I consider the same set of counterfactual experiments reported in Table 5 using the multi-sector version of the model. Table 7 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 countries for which sectoral-level output data is available.

By and large, the differences in the welfare effects of changes in trade barriers between the product-level and aggregate models 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. By contrast, the differences between the single- and multi-sector versions of the aggregate model are only significant in regard to the gains from trade relative to autarky. This suggests that accounting for the pattern of product-level comparative advantage is at least as important as accounting for sectoral differences in trade barriers and average productivity in determining the effects of trade barriers on aggregate trade flows and welfare.

Still, 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.

Table 7: Counterfactual Real Income per Worker: Single vs. Multi-Sector Models

Variable	Sectors	Model	Baseline	Autarky	$\min(d_{ni}, d_{in})$	$d_{ni} = 1$
mean($\Delta \ln(y_i)$)	Single	Agg.	–	-0.17	0.19	1.24
		P.L.	–	-0.26	0.26	1.17
	Mult.	Agg.	–	-0.27	0.20	1.27
		P.L.	–	-0.35	0.44	1.46
var($\ln(y_i)$)	Single	Agg.	0.83	0.88	0.67	0.47
		P.L.	0.78	0.99	0.56	0.41
	Mult.	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
	Mult.	Agg.	11.59	17.19	8.32	5.73
		P.L.	11.10	21.22	6.44	4.34

The overall pattern that emerges from these counterfactual experiments is that countries' patterns of product-level comparative advantage are important in determining the effects of trade barriers on aggregate trade flows and welfare in both single- and multi-sector models. Further, once these patterns have been taken into account, allowing for heterogeneity in trade costs and average productivity across industries either makes little difference for the aggregate effects of trade barriers or amplifies the differences between the predictions of the product-level and aggregate single-sector models.

Finally, it is important to note that product-level trade data is generally available for a superset of the countries for which industry-level output data exists, and the product-level model does not require the imposition of any additional structure than multi-sector models such as Caliendo and Parro (2015) and Levchenko and Zhang (2015). Taken together, this suggests that using product-level trade data to account for patterns of comparative advantage at a much lower level of aggregation than the industry level is both important and relatively low-cost for studies interested in the welfare effects of trade barriers regardless of the sectoral structure of the model being employed.

8 Concluding Remarks

In this paper, I have 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. Using this framework to infer countries' patterns of comparative advantage from product-level trade flow data, I have shown that falling trade barriers lead to significantly larger increases in income per worker on average and relatively large increases for developing countries than a model that relies only on aggregate trade data would suggest. Thus, accounting for the way in which the effect of trade barriers on welfare depends on patterns of product-level comparative advantage is quantitatively important. This 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.

One advantage of my framework is that it maintains, to a large extent, the tractability, relatively low data requirements, and simple, intuitive expressions for the gains from trade of widely used aggregate quantitative trade models, while also taking into account the effects of the interactions among countries' patterns of comparative advantage across products. My framework nests models that belong to the class delineated by Arkolakis et al. (2012) and makes clear that, in general, gains from trade depend on a term, which is missing from these aggregate models, that summarizes the effect of product-level comparative advantage. 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 is available, but it imposes no theoretical restrictions or additional data requirements beyond those typical of this class of aggregate trade models. Thus, the insights and methods of my paper can be applied in a wide range of contexts in which the effects of trade barriers are inferred from aggregated trade flows, including industry-level studies.

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.

Arvis, Jean-François and Ben Shepherd, "The Poisson Quasi-Maximum Likelihood Estimator: A Solution to the 'Adding Up' Problem in Gravity Models," *Applied Economics Letters*, April 2013, *20* (6), 515–519.

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," *The Review of Economic Studies*, 2015, *82* (1), 1–44.

Caron, Justin, Thibault Fally, and James R. Markusen, "International Trade Puzzles: A Solution Linking Production and Preferences," *The Quarterly Journal of Economics*, August 2014, *129* (3), 1501–1552.

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 Andrés Rodríguez-Clare, "Trade Theory with Numbers: Quantifying the Consequences of Globalization," in Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff, ed., *Handbook of International Economics*, Vol. 4 of *Handbook of International Economics*, Elsevier, 2014, pp. 197–261.

— , **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.

Fally, Thibault, "Trade, Tastes and Nutrition in India," *Journal of International Economics*, 2015, *(Forthcoming).*

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.

Gourieroux, C., A. Monfort, and A. Trognon, "Pseudo Maximum Likelihood Methods: Theory," *Econometrica*, 1984, *52* (3), pp. 681–700.

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

— **and** — , "The Evolution of Sectoral Productivity: Measurement and Implications," 2015, *Working Paper.*

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., *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.

Santos Silva, J. M. C. 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: 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 A2: 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 A3: 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. $coeff = \sum_j \delta^j coeff^j$.

Table A4: 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 A2 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 A1.

C A Multi-Sector Model

C.1 Model Setup

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 as that detailed in Section 2. In the extended model, 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

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

aggregate of sectoral composite goods given by

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

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 (7), 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

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

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

The interpretation of T_i – the determinant of aggregate productivity in autarky – remains the same as in the single-sector model, 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 P_n^{jk} is relatively high, 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, the expression for welfare (14) remains the same. 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 (10).

C.2 Quantitative Implementation

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 before, 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 A3 reports the sector-weighted average coefficient estimates for the multi-sector estimations along with those from single-sector estimations, based on the sample of countries for which sectoral-level manufacturing output data is available. 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 A4. The average multi-sector estimates are generally 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.

For simplicity, I have maintained the assumption that θ is constant across sectors. The robustness exercises reported in Appendix E indicate that the welfare effects of changing border costs are not sensitive to the value of θ, suggesting that sectoral variation in θ would also be relatively unimportant. In addition, while there have been some attempts to estimate θ at the sectoral level, notably Caliendo and Parro (2015), due to data requirements, such estimates are likely to be less reliable than aggregate estimates such as that of Simonovska and Waugh (2013).[29] The results reported in Appendix E do indicate that a higher value of θ leads to somewhat smaller gains from the removal of geographic trade barriers. Thus, 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. However, such barriers are much less likely than border costs to be influenced by trade policy.

[29] 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.

D Mathematical Appendix

D.1 Proof of Proposition 2

D.1.1 Case 1

Given that $T_i^k = T_i T^k$, equation (8) 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 that

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

This further implies that $(P_n^k)^{-\theta} = T^k P_n^{-\theta}$. Applying these results to (9) gives that

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

Substituting these values into (7) and (14) yields the first and third results of Proposition 2. The second result follows by applying the following result, based on (3), to (13):

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

□

D.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

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

This implies that

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

where the last equality results from the definition of T_i in (8). This implies that

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

Substituting these values into (7) 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, (13) implies that $\varepsilon_{ni} = -(\sigma - 1)$. The third result follows from rearranging (7).

□

D.1.3 Case 3

If $\theta = \sigma - 1$, then, from (8), $T_i = \sum_k T_i^k$. This implies that $\tilde{T}_{ni} = 1$. Substituting this value into (7) and (14) yields the first and third results of Proposition 2. The second result holds trivially

□

D.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 $d_{nn}^k = 1$.[30] I will discuss the model's predictions for the patterns depicted in Figure 5 and Table 3 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 TEI_{ni} is unaffected by this form of variation in trade costs 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 $\text{TEI}_{ni} = 0$, 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 $\text{TEI}_{ni} = 1$, for all n and i.

In case 3, arbitrary patterns of comparative advantage may exist, but they do not affect aggregate trade flows. Thus, TEI_{ni} can potentially take on any value between zero and one, but, in the baseline setup, the coefficient on $\ln(\text{dist}_{ni}) \times \widetilde{\text{TEI}}_{ni}$ reported in Table 3 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 (7), except that

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

and $(P_n^k)^{-\theta} = \sum_i T_i^k (c_i d_{ni} d_n^k)^{-\theta}$. 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 3, the estimated coefficient on $\ln(\text{dist}_{ni}) \times \widetilde{\text{TEI}}_{ni}$ should be zero.

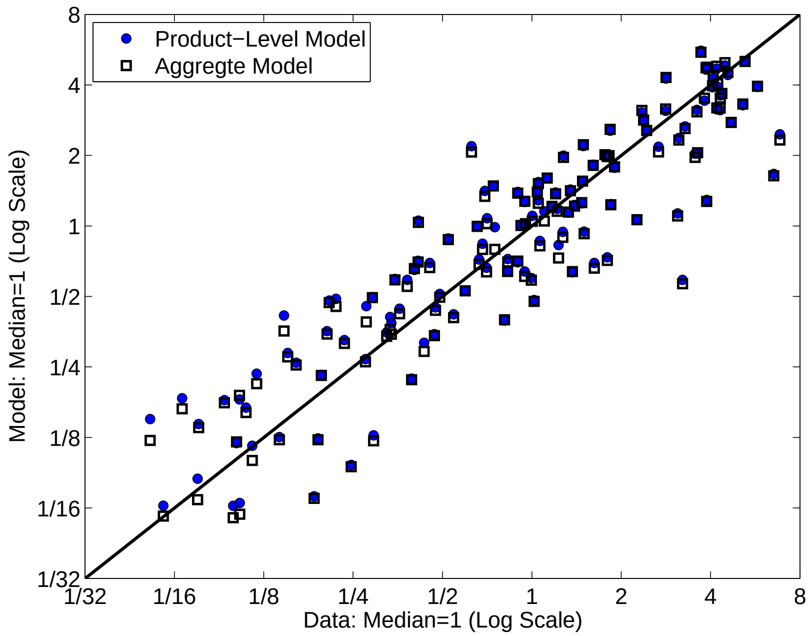
E Additional Quantitative Results

E.1 Income Per Worker

Figure 9 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)

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

Figure 9: Real GDP per Worker: Model vs. Data



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.

E.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. 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.

Table A5: Counterfactual Real Income per Worker: Alternative Parameters

Variable	θ	σ	Baseline	Autarky	$\min(d_{ni}, d_{in})$	$d_{ni} = 1$
mean($\Delta \ln(y_i)$)	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($\ln(y_i)$)	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

E.3 Robustness to Alternative Parameter Values

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 by conducting the same counterfactual experiments using alternative parameter values. 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.

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 A5 reports the same measures of income as Table 5 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 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.