Structural Estimation and Solution of International Trade Models with Heterogeneous Firms*

Edward J. Balistreri[†]
Colorado School of Mines

Russell H. Hillberry
University of Melbourne

Thomas F. Rutherford
The Swiss Federal Institute of Technology (ETH-Zürich)

August 2009

Abstract

We present an empirical implementation of a general-equilibrium model of international trade with heterogeneous manufacturing firms. The theory underlying our model is consistent with Melitz (2003). A nonlinear structural estimation procedure identifies a set of core parameters and unobserved firm-level trade frictions that best fit the geographic pattern of trade. Our estimation model is consistent with the specified general equilibrium model, and we conduct general equilibrium counterfactual analyses to illustrate model responses. We first assess the economic effects of reductions in measured tariffs. Taking the simple-average welfare change across regions the Melitz structure indicates welfare gains from liberalization that are nearly four times larger than in a standard trade policy simulation. Furthermore, when we compare the economic impact of tariff reductions with reductions in estimated fixed trade costs we find that policy measures affecting the fixed costs are of greater importance than tariff barriers. (JEL C68,F12)

---

[*] We received helpful comments on earlier drafts from David Hummels, Tom Hertel and Phil McCalman, from seminar participants at the Colorado School of Mines, the U.S. International Trade Commission, the Universities of Melbourne, Iowa and Memphis, and the Australian National University, and from conference participants at the 10th Annual Conference on Global Economic Analysis (GTAP Conference), the 82nd Annual Conference of the Western Economic Association, the 14th Empirical Investigations in International Trade conference, the fall 2006 Midwest International Economics Group meetings, the 12th annual Dynamics, Economic Growth and International Trade (DEGIT) conference and the Asia-Pacific Trade Seminar 2007. Any errors remain the sole responsibility of the authors.

[†] Corresponding author: Engineering Hall 311, Division of Economics and Business, Colorado School of Mines, Golden, CO 80401-1887; email: `ebalistr@mines.edu`.

1 Introduction

In canonical models of international trade, trade-policy changes induce factors of production to move between industries. Recent empirical and theoretical developments suggest that *within*-industry factor movements may also be an important channel through which policy affects economic outcomes. More productive plants are more likely to engage in exporting, so trade liberalization allows a reallocation of factors from less- to more-productive plants. Melitz (2003) derives a model that makes qualitative predictions about these phenomena. We evaluate the quantitative significance of these insights within the Melitz framework.[1]

Recent papers in the geography-of-trade literature use structural assumptions and the bilateral trade pattern to make inferences about the size of unobserved trade costs. Anderson and van Wincoop (2003) develop a structural estimation technique that can be used to infer trade costs.[2] Balistreri and Hillberry (2007) develop an extensive-form estimation strategy for the same model, and argue that it can be extended to most general equilibrium models of trade.[3] Balistreri and Hillberry (2008) link structural estimation techniques to established methods for calibrating general equilibrium models. We adapt these methods to calibrate a multi-sector, multi-region general equilibrium model in which the manufacturing sector has a Melitz-style market structure.

Our structural estimation procedure uses the equations defined by the Melitz model as a series of side constraints on the econometric objective. The model is underidentified, but assumptions about a few key structural parameters allow the econometric procedure to complete the identification of both the production technology and average bilateral trade

---

[1] Alvarez and Lucas (2007) evaluate the consequences of tariff reductions in the Eaton and Kortum (2003) model. Our exercise is motivated along similar lines, but focuses on the Melitz (2003) model.

[2] Eaton and Kortum (2003) and Behrens et al. (2008) engage in similar exercises, using structural assumptions to infer trade costs from the bilateral trade pattern.

[3] Su and Judd (2008) suggest a similar, extensive-form, technique as a general strategy for structural estimation of economic models.

costs. Leaning heavily on our structural model, we attribute differences between observed and fitted bilateral trade in manufactured goods to unobserved fixed costs of trade. This interpretation allows for an exact calibration of the model, without any role for idiosyncratic preferences in manufacturing trade.

The key structural parameters of the model are the distance elasticity of ad valorem trade costs and the parameter defining the shape of the Pareto distribution of firm level productivities. We estimate these under three different sets of identifying assumptions. Our preferred specification ties down the distance elasticity of trade costs, using unexplained variation in the trade pattern to fit the implied shape of the productivity distribution. Our estimate of this parameter is largely consistent with estimates from the confidential plant-level data.

An important strength of our estimation approach is that it facilitates direct, consistent, general equilibrium counterfactual analysis. Solving a multidimensional representation of the Melitz structure is, however, computationally demanding. The model requires joint solution of the non-linear trade equilibrium and a host of region specific non-linear entry and intra-industry production allocation conditions. We solve this problem with a computational algorithm that decomposes the general equilibrium into an industry-specific module, which determines the industrial organization, and a general equilibrium module which evaluates relative prices, comparative advantage, and the terms of trade. The full general equilibrium solution is achieved through a convergent iterative procedure.[4]

With our general equilibrium system parameterized and computable, we proceed to a quantitative assessment of the effect of trade policy changes. Using a standard Armington structure as the benchmark, we consider a 50 percent reduction in manufacturing tariffs. Endogenous productivity changes and changes in the number of varieties lead to larger welfare changes in the Melitz model than in the Armington baseline. Taking the simple-average welfare change across regions, the Melitz structure indicates welfare gains from liberaliza-

---

[4] Our discussion of the procedure is limited in the context of this paper, but details on the method and the computer programs are available from the authors.

tion that are nearly four times larger than those from the baseline model. We also consider reductions in the inferred bilateral fixed costs, and find that the welfare gains from these changes are substantially larger.[5] Joint reductions of tariffs and the inferred fixed costs of trade generate even larger welfare gains. When we reduce both tariffs and fixed border costs by 50% the average welfare gain is 30 times larger than in the case of tariff cuts in the Armington structure.

Our results may seem at odds with the theoretic work of Arkolakis et al. (2008), who show that iceberg trade cost reductions generate welfare results that are independent across economic models, conditional on fitting the observed trade pattern. That is, they contend that the Melitz (2003) structure does not offer new, or at least observationally unique, gains from trade. In related work, Feenstra (2009) shows that, in the same simplified structure, there are no net gains from new import varieties, because these are exactly offset by lost domestic varieties. Our framework differs from these authors' in at least three respects: a) we examine experiments that remove observed revenue-generating tariffs rather than iceberg trade cost reductions, b) we consider a multisector model, with resources moving between Melitz-style manufacturing and sectors with constant returns to scale, and c)our model has multiple factors and intermediate inputs. These empirically relevant features of our model allow net entry into the Melitz sector, so welfare gains or losses can be linked to changes in varieties consumed.

In section 2 we provide a review of the relevant literature. Section 3 outlines the full set of equilibrium conditions and a discussion of our solution method. In Section 4 we outline the nonlinear estimation procedure, and estimate the structural parameters that allow the model to best fit the data. These parameters are used to specify an operational general equilibrium model, which we employ to conduct counterfactual analysis. The results of this

---

[5] As in Anderson and van Wincoop (2003), Eaton and Kortum (2003), and Behrens et al. (2008), these exercises should be seen as illustrative calculations about the value of a significant move toward full integration, rather than as an assessment of particular trade policy changes.

analysis appear in section 5. In section 6 we discuss implications of our work and directions for further research.

2 Literature Review

Two broad areas of the empirical trade literature motivate heterogeneous-firms models like that of Melitz (2003). First, is the evidence that firm heterogeneity and intra-industry reallocation of resources is important. Second, there is considerable evidence of trade growth in new varieties. That is, trade growth along the *extensive* margin. The following provides an overview of the motivating evidence.

The literature now documents heterogeneity across establishments in productivity, export behavior, and responses to trade shocks. Important findings from this literature include a) there is wide variation in productivity levels among coexisting plants;[6] b) only a small fraction of establishments engage in exporting, and exporters tend to be larger and more productive than non-exporters;[7] c) there is considerable heterogeneity *among* exporters in the number of markets served per firm;[8] d) within-industry reallocation of market share from less-productive to more-productive establishments is an important component of aggregate productivity growth;[9] and e) productivity growth via shifting market shares (including the exit of the lowest productivity plants) is an important channel through which trade cost reductions induce aggregate productivity growth.[10]

A recent literature has focussed attention on the *extensive* margin of trade. Several authors have linked variation in aggregate trade flows to variation in the number of a) firms

---

[6] This is a robust feature of the data, as discussed in Bartelsman and Doms (2000).

[7] See, for example, Bernard and Jensen (1999), Roberts and Tybout (1997), and Bernard et al. (2003).

[8] Eaton et al. (2004) document this using French data.

[9] See Foster et al. (2001) and Aw et al. (2001), among others. An important component of the contribution of shifting market share is the exit of less-productive establishments.

[10] See, for example, Bernard et al. (2006) and Pavcnik (2002).

trading, b) commodities traded, and c) trading partners.[11] Of particular interest in this literature is explaining trade growth via this *extensive* margin.[12] Several authors have argued that new import varieties are a source of sizable welfare gains from trade liberalization.[13]

In a critique of the performance of applied general equilibrium models commonly used in trade policy analysis, Kehoe (2005) argues that the models typically fail along two dimensions: they do not allow trade policy to affect aggregate productivity, and they do not allow trade policy to induce trade growth along the extensive margin. While some policy estimates include *ad hoc* productivity adjustments, these attempts do not typically specify the mechanism by which trade policy is meant to induce productivity growth.[14] Analysis of the extensive margin of trade remains outside the scope of most policy-oriented trade models.

While the empirical literature has demonstrated the relevance of within-industry productivity heterogeneity and trade growth via the extensive margin, what has been lacking until recently is a sound theoretical structure that formalizes the insights from the empirical literature. Melitz-type models with firm heterogeneity and fixed trade costs offer a useful framework for addressing Kehoe's critique. Trade policy changes affect industry productivity by shifting market share away from low-productivity non-exporters, and toward high-productivity exporters. The model also allows for trade growth along the extensive margin, and provides a mechanism by which such trade growth can be linked to policy changes.

---

[11] Eaton et al. (2004) and Hillberry and Hummels (2008) show that variation in the number of firms serving a market explains variation in exports to that market. Hummels and Klenow (2005) and Broda and Weinstein (2006) identify the extensive margin in terms of the role of added commodities/trading partners.

[12] See Kehoe and Ruhl (2002) and Evenett and Venables (2002).

[13] Romer (1994) argues that the gains from such phenomena could be large. Several authors (including Feenstra (1994), Kehoe and Ruhl (2002), Klenow and Rodriguez-Clare (1997), Broda and Weinstein (2006)) have undertaken efforts to estimate these gains.

[14] See Anderson et al. (2005) for an example.

3 Theory

We assume Cobb-Douglas preferences over sectoral commodity bundles that are defined as constant-elasticity-of-substitution (CES) aggregates of differentiated products. Heterogeneous firms engage in monopolistic competition in the manufacturing sector. Other sectors of the economies are formulated with the standard assumptions in trade policy models: constant returns to scale and regional (Armington) differentiation.

In manufacturing, the heterogeneous-firms theory follows that of Melitz (2003). Firms pay a fixed cost of entry. Entrants receive a random productivity draw. Firms with sufficiently high productivity draws produce with a technology exhibiting increasing returns to scale, whereas firms with sufficiently low productivity draws choose not to produce. Trade costs include *ad valorem* iceberg costs, revenue-generating tariffs, and a fixed cost of entering each market. Firms with higher levels of productivity will be able to profitably serve more markets. The model is simplified by isolating the characteristics and behavior of the average firm participating in each bilateral market. Melitz (2003) develops the critical links between the average and marginal firms.

3.1 Demand

Consumers in region $s \in R$ are assumed to have Cobb-Douglas preferences over composites from different sectors, A_{is}, where the sector is indexed by $i \in I$ and α_{is} is the expenditure share;

$$U_s = \prod_i (A_{is})^{\alpha_{is}}. \tag{1}$$

The corresponding unit expenditure function is

$$e_s = \prod_i \left(\frac{P_{is}}{\alpha_{is}} \right)^{\alpha_{is}}, \tag{2}$$

where the P_{is} are the prices of the composites. Let E_s be each region's total expenditures. Demand for each sectoral composite is given by expenditures on that sector scaled by the price:

$$A_{is} = \frac{\alpha_{is} E_s}{P_{is}}. \tag{3}$$

The form of A_{is} depends on the nature of the subaggregates. In the heterogeneous firms sector, they are composed of firm-level varieties, while in the constant returns sector they aggregate varieties differentiated by region of origin. Let us index those sectors characterized by heterogeneous-firms industrial organization by $k \in K \subset I$, where in our application K only includes the manufacturing sector. The Dixit-Stiglitz composite of differentiated firm-level goods consumed in region s is given by

$$A_{ks} = \left[\sum_r \int_{\omega_{rs} \in \Omega_r} q_s(\omega_{rs})^{\rho} \, d\omega_{rs} \right]^{\frac{1}{\rho}}, \tag{4}$$

where ω_{rs} indexes the differentiated firm products sourced from region $r \in R$ (and Ω_r is the set of goods produced in r by the manufacturing sector). Substitution across the products is indicated by $\rho = 1 - 1/\sigma$, where σ is the constant elasticity of substitution. Notice that the arguments and parameters in equation (4) do not carry a k index because they are unique to heterogeneous-firms sectors, of which there is only one in our application. We therefore suppress the index in our discussion of the heterogeneous-firms sector.

It is more convenient to represent the aggregation [in (4)] in its dual form. The dual price index on the Dixit-Stiglitz composite consumed in region s is given by

$$P_{ks} = \left[\sum_r \int_{\omega_{rs} \in \Omega_r} p_s(\omega_{rs})^{1-\sigma} \, d\omega_{rs} \right]^{\frac{1}{1-\sigma}}. \tag{5}$$

Defining this in terms of the price of the *average* variety sourced from each region, $\tilde{p}_{rs}$, we have

$$P_{ks} = \left[\sum_r N_{rs} \tilde{p}_{rs}^{1-\sigma} \right]^{1/(1-\sigma)} \tag{6}$$

where N_{rs} is the number of varieties shipped from r to s. Melitz (2003) obtains this simplification by noting that $\tilde{p}_{rs}$ is the price set by a small firm with the CES weighted average productivity $\tilde{\varphi}_{rs}$.[15] Demand for the average variety to be shipped from r to s at a gross of trade and tax price of $\tilde{p}_{rs}$ is

$$\tilde{q}_{rs} = \frac{\alpha_{ks} E_s}{P_{ks}} \left(\frac{P_{ks}}{\tilde{p}_{rs}} \right)^{\sigma} \tag{7}$$

where E_s is, again, the value of gross expenditures in region s. Normally regional income would enter equation (7), but our data include expenditures on intermediates. E_s is income plus all intermediate purchases in region s. The model is simplified such that we do not have commodity-specific intermediate purchases, rather firms purchase the composite, which has a unit cost of e_s.

For the constant-returns goods, $i \notin K$, the aggregation of regional varieties is given by the standard Armington unit-cost function. Denoting marginal cost c_{ir} and the tariff rate t_{irs}, we have

$$P_{is} = \left[\sum_r \xi_{irs} [(1 + t_{irs}) c_{ir}]^{1-\sigma} \right]^{1/(1-\sigma)} \qquad \forall i \notin K, \tag{8}$$

where ξ_{irs} is the bilateral distribution parameter. Let Q_{ir} be the total output of sector i in region r. Market clearance for output of the constant-returns sectors is then given by the following:

$$Q_{ir} = \sum_s \xi_{irs} A_{is} \left(\frac{P_{is}}{(1 + t_{irs}) c_{ir}} \right)^{\sigma} \qquad \forall i \notin K. \tag{9}$$

---

[15] The weighted average productivity is given by

$$\tilde{\varphi}_{rs} = \left[\int_{\varphi_{rs}^*}^{\infty} \varphi_{rs}^{\sigma-1} \mu_{rs}(\varphi_{rs}) d\varphi_{rs} \right]^{\frac{1}{\sigma-1}},$$

where $\mu_{rs}(\varphi_{rs})$ is the distribution of productivities of each of the N_{rs} firms, and φ_{rs}^* is the productivity of the firm earning zero profits along the *r-s* bilateral trade route.

3.2 Firm-level environment

We assume a constant-returns Cobb-Douglas technology for the composite-input in each sector. The optimized composite-input price, c_{ir}, is a function of the price of intermediates and the prices of the primary-factors. Denote the factor prices w_{jr}, where $j \in J$ indexes the factor. Intermediates are purchased out of the gross expenditure composite, so their price is e_r. The unit-cost function is given by

$$c_{ir} = e_r^{\beta_{ir}} \prod_j (w_{jr})^{\lambda_{ijr}}, \tag{10}$$

where $\beta_{ir} + \sum_j \lambda_{ijr} = 1$.[16] For $i \notin K$ we do not need to go beyond (10), but for the heterogeneous-firms sectors we need to consider that the composite inputs are used to cover entry and fixed operating costs.

Operating firms in a given market use the composite input to cover both fixed-operating and marginal costs, but firms also face an *entry* cost. The entry cost entitles the firm to a productivity draw. If the productivity draw is sufficiently high the firm will operate profitably. Let f_r^e indicate the entry cost (in composite-input units), and let M_r denote the number of entered firms in region r. Then each of the M_r firms incur the nominal entry payment $c_{kr} f_r^e$, although this payment is spread across time (as there is a nonzero probability that the firm will survive beyond the current period).

Now consider the input technology for a firm from region r that finds it profitable to sell into market s. Let f_{rs} indicate the recurring fixed cost of operating on the r–s link, and let φ_{rs} represent the firm-specific measure of productivity. A firm supplying q units to s uses

$$f_{rs} + \frac{q}{\varphi_{rs}}$$

---

[16] We use GTAP input-output data [Dimaranan (2006)] to calculate the industry and country specific β_{ir} and λ_{ijr}.

units of inputs. Higher productivity (higher φ_{rs}) indicates lower marginal cost.

Once a firm incurs the *entry* cost, f_r^e, it is sunk and has no bearing on the firm's decision to operate in a given bilateral market. The profits earned by infra-marginal firms in the bilateral markets do, however, give firms the incentive to incur the entry cost in the first place. There is no restriction on the markets that can be served by a given member of M_r. If a firm's productivity is high enough such that it is profitable to operate in multiple markets it can costlessly replicate its technology, maintaining the same marginal cost but incurring the fixed operating cost, f_{rs}, for each of the s markets it serves.

The small firms, facing constant-elasticity demand for their differentiated products, follow the usual optimal markup rule. Let τ_{rs} indicate the *iceberg* transport-cost factor, and let t_{krs} indicate the tariff. Focusing on the average firm (with productivity draw $\tilde{\varphi}_{rs}$) shipping from r to s, optimal pricing is given by

$$\tilde{p}_{rs} = \frac{c_{kr}\tau_{rs}(1+t_{krs})}{\rho\tilde{\varphi}_{rs}}. \tag{11}$$

Consistent with our representation of (6), the $\tilde{p}_{rs}$ is the delivered price, gross of both tariffs and iceberg trade costs.

3.3 Operation, Entry, and the Average Firm

We assume that each of the M_r firms choosing to incur the entry cost receive their firm-specific productivity draw φ from a Pareto distribution with probability density

$$g(\varphi) = \frac{a}{\varphi}\left(\frac{b}{\varphi}\right)^a; \tag{12}$$

and cumulative distribution

$$G(\varphi) = 1 - \left(\frac{b}{\varphi}\right)^a, \tag{13}$$

where a is the shape parameter and b is the minimum productivity.

Considering the fixed cost of operating, f_{rs}, on the r–s link there will be some level of productivity, φ^*_{rs}, at which operating profits are zero. All firms drawing a φ above φ^*_{rs} will serve the s market, and firms drawing a φ below φ^*_{rs} will not. A firm drawing φ^*_{rs} is the marginal firm from r supplying region s. This leads us to the fundamental condition which determines the number of operating firms in a given market, N_{rs}. Let $r(\varphi) = p(\varphi)q(\varphi)$ indicate the gross-of-tariff firm revenues as a function of the draw φ. Zero profits for the marginal firm requires

$$c_{kr} f_{rs} = \frac{r(\varphi^*_{rs})}{\sigma(1 + t_{krs})}. \tag{14}$$

The $1 + t_{krs}$ term in the denominator reflects our definition of prices as gross of tariffs.

Equation (14) is intuitively appealing, but we would like to define this condition in terms of the average operating firm rather than the marginal firm. Following Melitz (2003) we define $\tilde{\varphi}_{rs}$ as the productivity of a firm pricing at $\tilde{p}_{rs}$, such that our simplification in equation (6) is consistent. The probability that a firm will operate is $1 - G(\varphi^*_{rs})$, so we find the CES weighted average productivity,

$$\tilde{\varphi}_{rs} = \left[\frac{1}{1 - G(\varphi^*_{rs})} \int_{\varphi^*_{rs}}^{\infty} \varphi^{\sigma-1} g(\varphi) d\varphi \right]^{\frac{1}{\sigma-1}}. \tag{15}$$

Assuming the Pareto distribution this gives us the relationship between the productivities of the average and marginal firms:

$$\tilde{\varphi}_{rs} = \left[\frac{a}{a + 1 - \sigma} \right]^{\frac{1}{\sigma-1}} \varphi^*_{rs}. \tag{16}$$

Following Melitz (2003) once again we establish the relationship between the revenues of firms with different productivity draws, using optimal firm pricing and the input technology

$(f + q/\varphi)$:

$$\frac{r(\varphi_1)}{r(\varphi_2)} = \left(\frac{\varphi_1}{\varphi_2}\right)^{\sigma-1}. \tag{17}$$

Using (25) and (17) to simplify (14) we derive the zero cutoff profit condition in terms of average-firm's revenues and the parameters:

$$c_{kr} f_{rs} = \frac{\tilde{p}_{rs}\tilde{q}_{rs}}{(1+t_{krs})} \frac{(a+1-\sigma)}{a\sigma}. \tag{18}$$

Next we turn to the entry condition which determines the mass of firms, M_r. Firm entry requires a one-time payment of f_r^e, and entered firms face a probability δ in each future period of a *bad* shock, which forces exit. In a steady-state equilibrium δM_r firms are lost in a given period so total entry payments in that period must be $c_{kr}\delta M_r f_r^e$. From an individual firm's perspective the *annualized* flow of entry payments is $c_{kr}\delta f_r^e$.

Assuming risk neutrality and no discounting, firms enter to the point that expected operating profits equal the entry payment. A firm from r operating in market s can expect to earn the average profit in that market:

$$\tilde{\pi}_{rs} = \frac{\tilde{p}_{rs}\tilde{q}_{rs}}{\sigma(1+t_{krs})} - c_{kr} f_{rs}. \tag{19}$$

Using the zero cutoff profit condition to substitute out the operating fixed cost this reduces to

$$\tilde{\pi}_{rs} = \frac{\tilde{p}_{rs}\tilde{q}_{rs}}{(1+t_{krs})} \frac{(\sigma-1)}{a\sigma}. \tag{20}$$

The probability that a firm in r will service the s market is simply given by the ratio of N_{rs}/M_r.[17] Setting the firm-level entry-payment flow equal to the expected profits from each

---

[17] In Melitz (2003) the probability that a firm will operate, which equals the fraction of operating firms in equilibrium, is presented as $1 - G(\varphi_{rs}^*)$.

potential market gives us the free entry condition

$$c_{kr}\delta f_r^e = \sum_s \frac{N_{rs}}{M_r}\frac{\tilde{p}_{rs}\tilde{q}_{rs}}{(1+t_{krs})}\frac{(\sigma-1)}{a\sigma} \tag{21}$$

which determines the mass of firms, M_r.

With entry established the nominal value of output from the heterogeneous-firms sector is given by

$$c_{kr}Q_{kr} = c_{kr}\delta f_r^e M_r + c_{kr}\sum_s N_{rs}\left(f_{rs} + \frac{\tilde{q}_{rs}\tau_{rs}}{\tilde{\varphi}_{rs}}\right), \tag{22}$$

where Q_{kr} can be interpreted as a quantity index on composite-input supply [analogous to Q_{ir} in equation (9)]. Notice the left-hand side of (22) is the minimized cost of supplying the composite input to manufacturing (because the value of c_{kr} is given by the unit-cost function). Canceling the c_{kr} terms, this reduces to the resource constraint on composite units for the heterogeneous-firms sector:

$$Q_{kr} = \delta f_r^e M_r + \sum_s N_{rs}\left(f_{rs} + \frac{\tilde{q}_{rs}\tau_{rs}}{\tilde{\varphi}_{rs}}\right). \tag{23}$$

We can recover the marginal productivity in each market as a function of the fraction of operating firms, $N_{rs}/M_r = 1 - G(\varphi_{rs}^*)$. Applying the Pareto distribution and inverting we have

$$\varphi_{rs}^* = \frac{b}{\left(\frac{N_{rs}}{M_r}\right)^{1/a}}. \tag{24}$$

Now using (16) we can substitute φ_{rs}^* out of the system;

$$\tilde{\varphi}_{rs} = b\left[\frac{a}{a+1-\sigma}\right]^{\frac{1}{\sigma-1}}\left(\frac{N_{rs}}{M_r}\right)^{-1/a}. \tag{25}$$

3.4 General Equilibrium

To complete the model we need to close the general equilibrium with a specification of endowments, expenditures, and income. Let $\bar{L}_{jr}$ be the endowment of factor j in region r, which equals total demand:

$$\bar{L}_{jr} = \sum_i \frac{\lambda_{ijr} c_{ir} Q_{ir}}{w_{jr}}. \tag{26}$$

Let W_r be the welfare level, and let Y_r equal nominal income. Then welfare is simply the level of income scaled by the true-cost-of-living index (which is the unit expenditure index):

$$W_r = \frac{Y_r}{e_r}. \tag{27}$$

Income in a region equals the value of factor endowments plus tariff revenues:

$$Y_s = \sum_j w_{js} \bar{L}_{js} + \sum_r \left[\frac{t_{krs}}{(1+t_{krs})} N_{rs} \tilde{p}_{rs} \tilde{q}_{rs} + \sum_{i \notin K} t_{irs} \xi_{irs} A_{is} \left(\frac{P_{is}}{(1+t_{irs}) c_{ir}} \right)^{\sigma} \right]. \tag{28}$$

Finally, we complete the general equilibrium by reconciling gross expenditures with nominal income. Expenditures equal income plus the nominal value of intermediate purchases:

$$E_r = Y_r + \sum_i \beta_{ir} c_{ir} Q_{ir}. \tag{29}$$

The full set of equilibrium conditions are summarized in Table 1. We use the convention of associating specific variables with equilibrium conditions. So, for example, a price is generally associated with a market clearance condition, and an activity level is generally associated with an (optimized) value function. The table also indicates the dimensionality of the system.

Table 1: General equilibrium conditions

Equilibrium Condition	**(Equation)**		**Associated Variable**	**Dimensions**
General conditions:				
Expenditure function	(2)		E_r : Total Expenditures	R
Demand by sector	(3)		A_{ir} : Composite activity level	$I \times R$
Unit-cost function	(10)		Q_{ir} : Sectoral quantity index	$I \times R$
Input market clearance	(26)		w_{jr} : Factor price by type	$J \times R$
Aggregate supply-demand balance	(29)		e_r : Aggregate price index	R
Final demand	(27)		W_r : Hicks welfare index	R
Regional budget	(28)		Y_r : Nominal income	R
Conditions dependent on the type of sector:				
Sectoral aggregation	(8)	$\forall i \notin K$	P_{ir} : Sectoral price index	$I \times R$
	(6)	$\forall i \in K$		
Sectoral resource constraint	(9)	$\forall i \notin K$	c_{ir} : Sectoral composite-input price	$I \times R$
	(23)	$\forall i \in K$		
Heterogeneous-firms industrial organization:				
Zero cutoff profits (ZCP)	(18)		N_{rs} : Number of operating firms	$R \times R$
Free entry (FE)	(21)		M_r : Mass of firms taking a draw	R
Firm-level demand	(7)		$\tilde{q}_{rs}$: Average-firm quantity	$R \times R$
Firm-level pricing	(11)		$\tilde{p}_{rs}$: Average-firm price	$R \times R$
CES wtd. Average φ	(25)		$\tilde{\varphi}_{rs}$: Average-firm productivity	$R \times R$
Total Dimensions:				$5R+(J \times R)+4(I \times R)+4(R \times R)$

3.5 Numeric solution method

In most policy modeling exercises, applied economists prefer to work with integrated equilibrium models formulated as systems of equations in which all variables are determined simultaneously. The system of non-linear equations in Table 1, however, is difficult to solve computationally once we consider multiple regions, factors, and commodities. Dimensionality and out-of-equilibrium nonconvexities limit our ability to solve the model robustly as an integrated equilibrium.

Our solution to this problem employs a *decomposition* algorithm that subdivides the system into two related equilibrium problems. A partial equilibrium (PE) model captures the heterogeneous-firms industrial organization in manufacturing, and the associated impact on productivity and prices. The PE model takes aggregate income levels and supply schedules as given. The second module is a *constant-returns* general equilibrium (GE) of global trade in composite input bundles. The GE model takes industrial structure (and thus, productivity, price indices, and the bilateral trade pattern) as given, and determines relative prices, comparative advantage and the terms of trade.

We iterate between the two models in policy simulations, letting the PE module determine the industrial structure and the GE model establish regional incomes and relative costs. The industrial structure (number of firms operating within and across borders and average productivity levels in the Melitz sector) is passed from the PE to the GE module, whereas the structure of aggregate demand (income levels and supply prices of inputs) are passed back from the GE to the PE module. Once the models are mutually consistent all of the conditions in Table 1 are satisfied and we have a solution to the multi-region general equilibrium with heterogeneous manufacturing firms.[18]

Our decomposition approach is successful because it confronts dimensionality and po-

---

[18] Contact the authors for the computer code, as well as additional information and documentation on our solution method.

tential out-of-equilibrium nonconvexities with a *divide-and-conquer* strategy. When we solve the industrial organization model in isolation from aggregate income changes, we avoid dealing with the excessively high dimensionalities that would otherwise arise when there are a large number of activities, markets, and agents. Similar decomposition strategies have been used in the past to compute standard constant-returns multiregion trade models, before the widespread availability of advanced nonlinear solvers. See, for example, Mansur and Whalley (1982). Our approach falls within this tradition, although we are solving a much more complex structure.

4 Nonlinear Least-squares Estimation

4.1 Data

We utilize data that is commonly employed in gravity estimations. The economic data includes gross manufacturing output by region, bilateral trade flows, and measured tariffs. Because we are interested in fitting a complete general equilibrium (including various non-manufacturing sectors), we take these data from the Global Trade Analysis Project (GTAP) [Dimaranan (2006)], a data set commonly employed in general equilibrium simulations of trade policy changes. The data track factor endowments, industry production, input-output relationships and bilateral trade for a broad set of countries. The version of the data we employ tracks these phenomena for the year 2001. The GTAP data has been balanced for use in general equilibrium studies (household income equals household expenditure, for example).[19] We aggregate the data into twelve regions,

---

[19] A key advantage of our data and approach, relative to other studies in the geography of trade literature, is that we are able to capture aspects of observable reality that other studies often miss. For example, we are able to account for a sizable role for the services sector in the economy, for observable cross-country differences in input bundles, and for geographical variation in intermediate demands. The cost of incorporating such richness is a reduced ability to incorporate a large number of regions in the analysis.

`CHN` China	`JPN` Japan	`CAN` Canada	`USA` United States
`MEX` Mexico	`ANZ` Australia and N. Zel.	`KTW` Korea and Taiwan	`ROA` Rest of Asia
`LAM` Latin America	`EUR` Europe	`EER` Eastern Europe and FSU	`ROW` Rest of World ;

seven aggregate sectors,

`AGR` Agriculture	`MTL` Mtls-related industry	`EIS` Other Energy Intensive
`MFR` Manufacturing	`SER` Services	`ENG` Energy
`CGD` Savings good .		

and five factors of production

`LND` Land	`SKL` Skilled labor	`LAB` Unskilled labor
`CAP` Capital	`RES` Natural Resources.	

We focus our estimation of the heterogeneous-firms model on the aggregate manufacturing sector. The other sectors in the general equilibrium are assumed competitive and calibrated via the usual techniques.[20] The regional aggregation attempts to strike a balance between our desire to capture the relevant geographic trade patterns and provide useful reports given the data and computational challenges.[21]

In addition to the economic data we utilize distances between regions to inform transportation costs.[22] Consistent with the gravity literature we assume that iceberg trade costs take on the following form:

$$\tau_{rs} = d_{rs}^{\theta};$$

where θ is an estimated distance elasticity.[23] In addition, we include rent generating ad val-

---

[20] We assume an Armington trade structure, with constant returns and perfect competition, in the sectors other than manufacturing.

[21] Despite the algorithm, high levels of dimensionality limit our ability to substantially expand the sample beyond what we report here. Earlier versions of this paper reported results from a nine-region aggregation (which provided no country-level reports). In addition, we have explored variations on other regional aggregations of the GTAP data. The estimated distance elasticities, Pareto shape parameters, and overall welfare analysis of integration where not substantially altered by our choice of geographic aggregation. The region specific fixed costs of trade are not estimated with great precision, and are therefore more sensitive to the particular aggregation.

[22] Our baseline distance data are described in Mayer and Zignago (2006). In our cross-country aggregation we must aggregate distances, and we apply Mayer and Zignago (2006)'s *distw* method for aggregating across bilateral city pairs to aggregate over bilateral country pairs.

[23] A normalization of distance is required to pin down the absolute scale of distance-related costs. We scale distance such that $\tau_{rs} = 1$ on the shortest link.

orem tariffs as measured in the GTAP data, t_{irs}. The c.i.f. import prices of manufacturing goods, thus, include both the tariff and the iceberg cost factor (i.e., $(1+t_{krs})\tau_{rs}$).

4.2 Estimation strategy

Consider a B dimensional subset of the general-equilibrium conditions that characterizes trade and industrial organization in the heterogeneous-firms sector. The relevant conditions are given by equations (6), (23), (18), (21), (7), (11), and (25). These form a $B = 3R + 4R^2$ dimensional non-linear system that might be written as $F(\boldsymbol{x},\boldsymbol{\gamma}) = 0$, which implicitly maps a set of exogenous parameters , $\boldsymbol{\gamma} \in \mathbb{R}^A$, to a vector of endogenous variables $\boldsymbol{x} \in \mathbb{R}^B$.[24] Let $\hat{\boldsymbol{\gamma}} \in \{\mathbb{R}^{\hat{A}} : \hat{A} \leq A\}$ denote a vector of core parameters to be estimated, and let $\hat{\boldsymbol{x}} \in \{\mathbb{R}^{\hat{B}} : \hat{B} \leq B\}$ denote a key endogenous series (e.g., bilateral trade flows). Our estimation strategy is to find the $\hat{\boldsymbol{\gamma}}$ that minimize the sum of the squared differences between the $\log \hat{\boldsymbol{x}}$ and observed $\log \boldsymbol{x}^0$ subject to $F(\boldsymbol{x},\boldsymbol{\gamma}) = 0$ and an additional $A - \hat{A}$ direct assumptions about the values of the remaining parameters:

$$\begin{array}{cl} \min_{\{\hat{\boldsymbol{\gamma}},\hat{\boldsymbol{x}}\}} & ||\log \hat{\boldsymbol{x}} - \log \boldsymbol{x}^0||^2 \\ \text{subject to:} & F(\boldsymbol{x},\boldsymbol{\gamma}) = 0, \\ \text{and} & \bar{\boldsymbol{\gamma}} = \mathbf{k}, \end{array}$$

where $\bar{\boldsymbol{\gamma}}$ are the assumed parameters and $\mathbf{k}$ is a vector of constants.[25] The norm is defined in logs to be consistent with the empirical trade literature, which often assumes a log-linear form of the trade equation.

Taking the GTAP data as given, and given our assumed structure of iceberg trade costs,

---

[24] We included in set of parameters a few observable variables from the broader general equilibrium. These are held at their observed levels for the purpose of estimation.

[25] In addition to the assumed parameters, the general equilibrium values of expenditures, expenditure shares, and manufacturing output are given in the benchmark data; E_r, α_{kr} and $c_{kr}Q_{kr}$ are observed. We estimate conditional on matching these values. The endogenous variable c_{kr} is determined (and fixed in estimation) by selecting units such that $c_{kr} = 1$, which uniquely determines Q_{kr} as it enters equation (23).

we have the following candidates for inclusion in $\hat{\boldsymbol{\gamma}}$:

σ : the inter-variety elasticity of substitution,
δ : the probability of firm death,
a : shape parameter for the Pareto distribution,
b : minimum productivity parameter for Pareto distribution,
θ : distance elasticity of iceberg trade costs,
f_r^e : fixed entry cost,
f_{rs} : bilateral fixed cost of shipping from region r to region s.

Informing these parameters off observed bilateral flows is not meaningful unless we are willing to significantly reduce the parameter space (beyond our implicit assumptions that the core distribution, substitution, and transport cost parameters are identical across regions). With R regions there are potentially $\hat{B} = R^2$ observable flows, but there are at least $R^2 + R + 5$ parameters. Following Anderson and van Wincoop (2003), we might eliminate σ from the list of parameters; noting that we are interested in identifying trade costs conditional on an assumed substitution elasticity. We assume that $\sigma = 3.8$ throughout our analysis following the plant-level empirical analysis of Bernard et al. (2003). In addition, we directly assume the values $\delta = 0.025$, $f_r^e = 2$, and $b = 0.2$ following Bernard et al. (2007).[26]

The primary assumption that we employ to reduce the parameter space is to impose structure on the fixed costs. Let f_r^p be a fixed cost that is specific to goods produced in region r, and f_s^x be a fixed cost that is specific to goods exported to region s. Now consider decomposing the bilateral fixed costs as follows:

$$f_{rs} = \begin{cases} f_r^p + f_s^x + f_{rs}^r, & \text{for } r \neq s; \\ f_r^p + f_{rs}^r, & \text{for } r = s. \end{cases}$$

When $r = s$ the f_s^x term drops out reflecting the idea that f_s^x is an outward trade barrier. The

---

[26] Bernard et al. (2007) explain that changes in f^e rescales the mass of firms where as changes in δ rescale the mass of entrants relative to the mass of firms. For consistency we simply adopt their values.

f_{rs}^{r} are idiosyncratic residual bilateral costs. In the initial estimation the f_{rs}^{r} are constrained to be zero. The number of parameters to be estimated is thus reduced to $\hat{A} = 2R + 2$. Once the core parameters are estimated, and locked down, the system can be used to calculate the matrix of residual f_{rs}^{r} which generate an exact fit on trade flows.

4.3 Estimation Results

The primary purpose of our nonlinear estimation is to complete a calibration of the numerical model. This entails a complete enumeration of the structural parameters necessary to reconcile the structural model with observed data. The primary parameters of interest are those that are taken to be common across the world: the shape of the implied Pareto distribution of productivity draws, a, and the distance elasticity of trade costs, θ. The model links both these parameters to the geographic pattern of trade. θ governs the degree to which delivered prices rise over distance. Chaney (2008) shows that a exerts a substantial influence on the geography of bilateral trade, via the extensive margin.

We conduct three econometric calibrations of the model. In the first, we allow both θ and a to be free parameters; they take the values that minimize the econometric objective, subject to the constraints defined by the model and our choices of the parameters in $\bar{\boldsymbol{\gamma}}$. Very good estimates of θ appear in the literature, and our second set of estimates constrains the estimation procedure to replicate a commonly accepted value, $\bar{\theta} = 0.27$. As a sensitivity check, our third set of estimates imposes the constraint $\bar{\theta} = 0.46$.[27] Our estimates of key structural parameters appear in Table 2. In order to characterize the degree to which our procedure fits the observed trade pattern, we conduct a log linear regression of observed flows on fitted flows. This regression returns an R^2 of 0.939, 0.897, and 0.827, respectively.

---

[27] These latter two estimates are taken from Hummels (2001), who estimates θ directly off observed transportation cost margins. 0.27 is Hummels' central estimate, using data from seven countries that report transport margins in their international trade statistics. 0.46 is the elasticity of air freight charges with respect to distance in U.S. data, and Hummels uses this as a plausible upper bound on θ. Our unconstrained estimate of θ lies well below Hummels' central estimate, and we treat this as a lower bound.

Table 2: Nonlinear estimation results (dependent variable is log bilateral flows; core fixed parameters are $\sigma = 3.8$, $\delta f^e = 0.05$, and $b = 0.2$)

			Specification	
		θ =free	$\bar{\theta} = 0.27$	$\bar{\theta} = 0.46$
Pareto shape parameter: a		5.171	4.582	3.924
		(0.336)	(0.227)	(0.019)
Distance elasticity: θ		0.155	0.27	0.46
		(0.028)		
Source-specific fixed cost: f_r^p				
`CHN`	China	0.036	0.026	0.430
		(0.385)	(0.396)	(0.220)
`JPN`	Japan	0.300	0.340	0.519
		(0.239)	(0.212)	(0.237)
`CAN`	Canada	0.347	0.303	0.228
		(0.305)	(0.210)	(0.103)
`USA`	United States	0.299	0.172	0.328
		(0.337)	(0.369)	(0.065)
`MEX`	Mexico	0.024	0.063	0.316
		(0.185)	(0.153)	(0.088)
`ANZ`	Australia and N. Zel.	0.039	0.030	0.019
		(0.014)	(0.004)	(0.000)
`KTW`	Korea, and Taiwan	0.047	0.027	0.023
		(0.088)	(0.070)	(0.008)
`ROA`	Rest of Asia	0.047	0.026	0.016
		(0.074)	(0.020)	(0.003)
`LAM`	Latin America	0.218	0.255	0.120
		(0.262)	(0.170)	(0.008)
`EUR`	Europe	0.503	0.330	0.753
		(0.322)	(0.312)	(0.157)
`EER`	Eastern Europe and FSU	1.300	0.562	0.222
		(0.518)	(0.338)	(0.022)
`ROW`	Rest of World	3.666	1.280	0.237
		(1.449)	(0.738)	(0.015)
Destination-specific fixed cost: f_s^x				
`CHN`	China	0.258	0.147	20.647
		(4.402)	(6.379)	(15.872)
`JPN`	Japan	0.872	0.493	0.035
		(0.986)	(0.461)	(0.102)
`CAN`	Canada	15.153	13.453	0.0 **bound**
		(11.352)	(7.838)	(0.006)
`USA`	United States	0.714	0.149	0.0 **bound**
		(1.021)	(0.707)	(0.000)
`MEX`	Mexico	0.0 **bound**	0.0 **bound**	0.0 **bound**
		(0.300)	(0.052)	(0.000)
`ANZ`	Australia and N. Zel.	6.291	12.004	37.512
		(3.184)	(6.714)	(1.324)
`KTW`	Korea, and Taiwan	0.913	0.346	0.0 **bound**
		(1.003)	(0.673)	(0.001)
`ROA`	Rest of Asia	1.602	2.067	2.708
		(1.618)	(1.497)	(1.723)
`LAM`	Latin America	8.254	23.532	33.984
		(9.412)	(20.009)	(0.992)
`EUR`	Europe	1.339	0.293	0.0 **bound**
		(1.293)	(0.403)	(0.000)
`EER`	Eastern Europe and FSU	55.005	35.110	18.628
		(18.955)	(3.000)	(0.505)
`ROW`	Rest of World	142.158	98.561	59.633
		(47.798)	(6.241)	(1.294)

Table 3: Heterogeneity in the productivity distribution: $\frac{\varphi}{b}$ for selected values of a

	Percentiles			
	50th	75th	90th	95th
Fitted values of a				
$\hat{a}$=5.171	1.143	1.307	1.561	1.785
$\hat{a}$=4.582	1.163	1.353	1.653	1.923
$\hat{a}$=3.924	1.193	1.424	1.798	2.146
Implied estimate from Eaton et al. (2004)				
a=4.2	1.179	1.391	1.730	2.041
Value used in Bernard et al. (2007)				
a=3.4	1.226	1.503	1.968	2.414

The interaction between θ and a is a key point of interest. Conditional on the bilateral trade pattern, our procedure must assign responsibility for trade reductions to these two parameters (along with the fixed costs). Our unconstrained estimate of θ is $\hat{\theta} = 0.155$, while $\hat{a} = 5.171$. This is a relatively low estimated distance elasticity, and a somewhat high estimate for the Pareto distribution parameter.[28] As we constrain $\bar{\theta}$ to higher values, the estimated value of a falls. For $\bar{\theta} = 0.27$, $\hat{a} = 4.582$, and for $\bar{\theta} = 0.46$, $\hat{a} = 3.924$.

The lower values of a that occur in our restricted estimates imply greater heterogeneity in firm productivities. Table 3 illustrates some features of the productivity distributions implied by different values of a. For our unconstrained estimate of a, a firm with a productivity draw at the median of the distribution would be 1.143 times as productive as a firm with the minimum draw. As a falls, the productivity distribution flattens out. In our subsequent counterfactual scenarios, we will be employing the constrained estimate $a = 4.582$, the estimate corresponding to $\bar{\theta} = 0.27$. In this case, the median productivity draw is 1.163 times the size of the minimum draw.

The structural estimation results in Table 2 also contain estimated values of source- and destination-specific fixed costs (f_r^p and f_s^x, respectively). Just as other structural procedures impute values for model-consistent P_{ks}, our fitting procedure imputes model-consistent

---

[28] As noted earlier, Hummels' central estimate of θ is 0.27. Estimates of a—which are taken from distributions of plant/firm level market shares—vary, and are conditional on a choice of σ. Bernard et al. (2007) choose $a = 3.4$, and the estimates in Eaton et al. (2004) imply $a = 4.2$ under our maintained assumption that $\sigma = 3.8$.

fixed costs (as well as P_{ks}). Model consistency includes two requirements. First, by (18), the size of $f_r^p + f_s^x$ should be such that the average firm pays $\frac{a+1-\sigma}{a\sigma}$ of revenue, net of tariffs, in fixed costs. Second, as Chaney (2008) shows, the elasticity of fitted trade with respect to $f_r^p + f_s^x$ should equal $1 - \frac{a}{\sigma-1}$. Conditional on other trade resistances, the model attributes *home bias* in fitted trade flows to f_s^x. The source-specific fixed costs, f_r^p, are, in effect, source-specific fixed effects that help the procedure to best fit the data under these constraints.

The results in Table 2 indicate that both relative and absolute estimates of fixed costs vary with θ and a.[29] This is unsurprising, as both the adding up conditions of the model and the marginal conditions derived by Chaney (2008) link fixed costs to the values of these parameters. Some destinations (i.e. China) have substantial variation across columns in the implied fixed costs of importing, f_s^x. Substantial differences across columns are understandable. Observed levels of import penetration are being replicated across successive calibrations, even as changes in θ are substantially shifting the burden of trade resistance from distance related ad valorem costs to fixed costs of trade.

Our non-linear estimating system fits the data in a manner quite similar to a conventional gravity equation. An OLS regression of log($\hat{\boldsymbol{x}}$) on (logged) d_{rs}, $(1 + t_{krs})$, $(f_r^p + f_s^x)$, $c_{kr}Q_{kr}$, $\alpha_{ks}E_s$, and P_{ks} returns a perfect fit. This intuitive OLS specification relies, of course, on our non-linear estimates of f_r^p, f_s^x, and P_{ks}.[30] Following Hillberry and Hummels (2008) we decompose trade flows into extensive and intensive margins, regressing log N_{rs} and $\tilde{p}_{rs}\tilde{q}_{rs}$ on the determinants of bilateral trade. We report these results in Table 4. The results show

---

[29] Standard errors for the estimated fixed costs are generally tighter when the distance elasticity is fixed. Fixing θ allows a to be more precisely estimated (exploiting any unexplained variation in the distance elasticity of trade, for example). More precise estimates of a allow more precise estimates of the fixed costs in the model.

[30] The fitted values of trade in the regression are consistent with optimizing behavior of the agents in the model. As Helpman et al. (2008) show, a Pareto distribution for firm heterogeneity coupled with an assumption that fixed costs are exporter- and importer-specific (which is true in the case of our fitted values) generates a multiplicative gravity relationship (and thus a perfect fit in a properly specified log-linear regression of structurally fitted flows on the appropriate dependent variables). The estimates here link the flows to variables that appear in the general equilibrium trade model; there is no need for a summary measure like *outward multilateral resistance* as described in Anderson and van Wincoop (2004).

Table 4: Intensive and Extensive Margins in Fitted Bilateral Trade Flows

	$\theta = free$		$\theta = 0.27$		$\theta = 0.46$	
Regressand	N_{rs}	$\tilde{p}_{rs}\tilde{q}_{rs}$	N_{rs}	$\tilde{p}_{rs}\tilde{q}_{rs}$	N_{rs}	$\tilde{p}_{rs}\tilde{q}_{rs}$
Regressor						
d_{rs}	-0.799	0	-1.236	0	-1.804	0
	(0.001)	(0)	(0.001)	(0)	(0.001)	(0)
$(1+t_{krs})$	-7.037	1	-6.240	1	-5.349	1
	(0.006)	(0)	(0.005)	(0)	(0.005)	(0)
$(f_r^p + f_s^x)$	-1.846	1	-1.636	1	-1.401	1
	(0.000)	(0)	(0.000)	(0)	(0.000)	(0)
$c_{kr}Q_{kr}$	1.000	0	1.000	0	1.001	0
	(0.000)	(0)	(0.000)	(0)	(0.000)	(0)
$\alpha_{ks}E_s$	1.846	0	1.635	0	1.401	0
	(0.000)	(0)	(0.001)	(0)	(0.001)	(0)
P_{ks}	5.165	0	4.578	0	3.921	0
	(0.003)	(0)	(0.002)	(0)	(0.002)	(0)
constant	-11.328	2.114	-9.800	2.281	-8.074	2.585
	(0.004)	(0)	(0.004)	(0)	(0.004)	(0)

Note: All variables are in natural logarithms. Standard errors in parentheses.
All regressions have 144 observations and $R^2 = 1$.

that average firm revenues (c.i.f.) rise in tandem with tariffs and fixed trade costs.[31] Most trade responses to the gravity-type variables are observed as changes in the number of firms serving a particular market. The regression coefficients in Table 4 can be decomposed in each case into theory-consistent functions of the structural parameters a, σ, and θ.

Once the core parameters reported in Table 2 are established we can freeze these at their point estimates and find a set of residual bilateral costs, f_{rs}^r, that provide consistency with observed trade flows.[32] From the perspective of the nonlinear estimation these are residuals—they allow the structure to fit the data exactly. From the perspective of performing theory consistent counterfactual analysis they are idiosyncratic calibration parameters.[33]

---

[31] Note that the implications for the intensive margin follow directly from equation (18).

[32] There are a number of potential matrices of residual fixed costs that are consistent with observed trade and the estimated parameters. We choose the one that minimizes the squared residual bilateral costs.

[33] Hillberry et al. (2005) show the usefulness of framing standard general-equilibrium calibration exercises

Table 5: Total Bilateral Fixed Trade Costs (f_{rs})

						Destination						
	`CHN`	`JPN`	`CAN`	`USA`	`MEX`	`ANZ`	`KTW`	`ROA`	`LAM`	`EUR`	`EER`	`ROW`
Source												
`CHN`	0.028	1.005	1.950	0.080	0.010	8.399	2.781	1.813	2.511	0.293	8.962	1.933
`JPN`	1.986	0.313	2.826	0.111	0.581	4.243	1.539	0.380	1.304	0.227	23.574	2.518
`CAN`	0.323	0.103	0.048	0.184	4.319	4.739	0.279	2.093	4.894	0.611	35.851	8.298
`USA`	0.654	0.341	5.180	0.043	3.403	5.077	0.098	0.384	4.893	0.478	64.550	3.970
`MEX`	1.668	1.146	2.571	0.133	0.138	19.710	0.706	4.606	4.018	1.153	30.763	38.193
`ANZ`	0.195	0.010	0.183	0.056	0.094	0.026	0.017	0.087	1.788	0.016	3.002	0.397
`KTW`	0.703	3.016	0.734	0.034	0.110	2.436	0.021	0.244	0.296	0.091	8.074	0.508
`ROA`	0.449	0.123	0.621	0.017	0.178	2.187	0.080	0.022	2.389	0.068	18.781	1.478
`LAM`	0.258	0.185	5.982	0.251	1.006	20.952	0.071	0.160	0.036	0.127	1.550	1.670
`EUR`	0.343	0.394	5.147	0.324	0.895	1.738	0.182	0.592	2.738	0.145	11.974	2.172
`EER`	0.678	0.583	8.884	1.217	2.050	14.418	0.859	3.406	12.735	0.316	0.195	6.599
`ROW`	3.132	1.289	25.687	0.378	18.280	23.717	1.494	3.667	40.663	0.510	23.840	0.030

Table 5 shows the full matrix of total bilateral fixed costs including the residual plus the source and destination charges. So, for example, we might consider that import penetration into `EER` is difficult, but it is particularly difficult for American firms. On this particular link the total fixed cost of 64.6 is nearly twice as large as the base destination charge to get into `EER`, 35.1.[34]

5 Counterfactual Simulations

We analyze four scenarios that compare the impacts of tariff and fixed cost reductions:

- **(A)** Armington constant-returns formulation with a 50% reduction in manufacturing tariffs;
- **(B)** Heterogeneous-firms model with a 50% reduction in manufacturing tariffs;
- **(C)** Heterogeneous-firms model with a 50% reduction in fixed trade costs; and
- **(D)** Heterogeneous-firms model with both the tariff and fixed cost cuts.

Scenario (**A**) is a reference case where we assume a standard Armington trade structure and constant-returns production for manufacturing. The model structure is the same but

---

as the systematic identification of idiosyncratic residual parameters. As in any standard econometric exercise these residual parameters are useful indicators of model fit. In our preferred specification, variation in the estimated f_{rs} explains 14 percent of the variation in bilateral trade flows.

[34] Relatively large fixed costs of entry into aggregated regions like `EER` may be partially attributed to geographical aggregation. Our aggregation implicitly treats the collection of smaller countries in `EER` as a large integrated market, so large fixed costs are needed to explain a relatively low volume of trade.

Table 6: Counterfactual Welfare Impacts (% Equivalent Variation)

	Scenario			
	A	B	C	D
	CRTS-Tariff	Tariff	Fix Cost	Both
Region				
`CHN`	0.3	1.3	2.9	4.7
`JPN`	0.1	0.3	1.5	1.7
`CAN`	-0.0	0.1	4.2	4.2
`USA`	0.0	0.0	0.8	0.9
`MEX`	0.2	0.7	4.3	5.3
`ANZ`	0.3	1.4	2.0	4.0
`KTW`	0.4	1.1	5.3	6.4
`ROA`	0.3	1.0	4.8	6.0
`LAM`	0.1	0.4	1.4	2.1
`EUR`	0.1	0.2	1.2	1.5
`EER`	-0.1	-0.3	3.7	4.0
`ROW`	-0.2	-0.7	2.1	1.6

$k \in K = \{\ \}$ in scenario (**A**), which gives us a consistent reference point to judge the performance of the new theory.[35] Table 6 shows welfare changes induced by the tariff cut. Although most regions gain from the tariff cuts, three regions suffer welfare losses (`CAN`, `EER`, and `ROW`). Reported welfare losses from tariff removal are not uncommon in Armington models, because the structure implies relatively high optimal tariffs [see Brown (1987) or Balistreri and Markusen (2009)].

Examining the same tariff cuts in the heterogeneous-firms model, Scenario (**B**), indicates substantially greater gains. Taking the simple-average welfare change across regions the heterogeneous-firms structure indicates welfare gains from liberalization that are nearly four times larger than in the baseline case. The simple-average welfare gain of 0.5% in Sce-

---

[35] One might argue, based on the theoretic work of Arkolakis et al. (2008), that an appropriate comparison would increase the manufacturing Armington elasticity of substitution in the reference scenario to $a+1$ to generate the same trade responses across models. Given that we have departed from Arkolakis et al. (2008) by modeling many sectors and factors (and we consider rent-generating tariffs), there is no Armington elasticity that can give us bilateral trade responses that are the same as those generated from the Melitz structure. In our reported results we maintain $\sigma = 3.8$ in the Armington model—a value generally consistent with the policy-simulation literature critiqued by Kehoe (2005). In a sensitivity run we set the Armington elasticity for manufacturing at $a+1 = 5.6$. With the higher elasticity in the reference model, the simple-average welfare gains in the Melitz model are on the order of 2.5 times larger than the Armington model.

nario (**B**) may not seem particularly impressive, but consider the following statistics from our aggregation of the GTAP data: gross manufacturing output is only 26% of world gross output, only 17% of manufacturing output is traded to another region, and the simple average benchmark tariff on these flows is only 9.7%. So the typical tariff cut is less than 5%, and applies to less than 5% of gross output. In this context, an average welfare gain of 0.5% seems quite large. With the exception of the `ROW` and `EER` regions, tariff cuts in the heterogeneous-firms structure produce larger net welfare gains than the constant returns benchmark.

In Scenario (**C**) we examine a 50% cut in the fixed costs associated with non-domestic trade links. This generates important gains across the board. The results are consistent with a recent trade literature focussing on the relative importance of unobserved (non-tariff) barriers and tariffs.[36] In Scenario (**D**) both the tariff and fixed cost reductions are combined. There are considerable increases in welfare under Scenario (**D**) considering that Manufacturing is the only sector being liberalized. The simple-average welfare gain under Scenario (**D**) is 30 times larger than in the Armington reference case. Notice also that fixed-cost (or non-tariff barrier) reductions often complement tariff cuts. Absolute welfare increases of 2% to 6% are considerably larger than most computational estimates of the value of trade liberalization.[37]

As noted above, one of the key critiques of current policy simulation models is that they fail to account for the productivity growth associated with trade liberalization. Table 7 indicates the simulated gains in average productivity across firms active in their respective domestic markets. Increased exposure to external markets, whether induced by a reduction in tariff or non-tariff barriers, induces productivity growth.

The other key component of the model is that trade policy affects the extensive margin. The number of foreign varieties increases when trade costs fall. The threshold for import penetration falls and more foreign firms find it profitable to enter a given market. In contrast,

---

[36] Anderson and van Wincoop (2004)

[37] See Rutherford and Tarr (2002).

Table 7: Domestic-firm Productivity Growth (% Change)

	Scenario		
	B	C	D
	Tariff	Fix Cost	Both
Region			
CHN	0.8	1.4	2.5
JPN	0.5	1.4	2.2
CAN	0.7	4.7	5.5
USA	0.3	1.5	1.9
MEX	1.1	4.2	5.7
ANZ	2.1	2.7	6.1
KTW	1.1	3.4	5.0
ROA	1.3	3.2	5.1
LAM	1.1	1.4	3.0
EUR	0.4	1.5	2.1
EER	0.8	2.8	4.0
ROW	0.6	1.5	2.3

the effect of changes in trade costs on the number of domestic varieties is not clear. The number of exporting firms increases and the profits of all exporting firms increase, which induces entry of new varieties. The increased activity of these firms, however, bids up the input price. This effect, together with substitution toward imported varieties, induces exit of domestic varieties with low productivity realizations.[38] All of the variety and price effects can be summarized in the solution price index on manufactured goods, P_{kr}. Table 8 presents the percentage change in the price index across the scenarios. Further, we break out the variety effects in Table 9. Although the number of overall varieties falls for many regions, trade growth on the extensive margin combined with lower domestic prices result in lower overall price indexes and welfare gains.

---

[38] Feenstra (2009) shows, in a one-sector one-factor heterogeneous-firms model where iceberg trade costs are removed, that the welfare gains of new import varieties is just offset by the lost domestic varieties. So although the number of varieties falls (a result shown by Baldwin and Forslid (forthcoming)), there are no gains or losses from changes in varieties consumed. Feenstra (2009) does show, however, gains due to the reallocation of resources toward more productive firms. The strong analytical results found by Feenstra do no apply in our model, because our model has multiple sectors and factors resulting in elastic factor supply to the heterogeneous firms sector. With elastic factor supply, welfare is impacted by changes in varieties.

Table 8: Manufacturing Price Index, P_{kr} (% Change)

	Scenario		
	B	C	D
	Tariff	Fix Cost	Both
Region			
CHN	0.8	-0.9	-0.9
JPN	0.1	0.3	-0.2
CAN	-0.6	-4.3	-4.9
USA	-0.2	-2.2	-2.6
MEX	-2.4	-3.8	-6.3
ANZ	0.3	-2.4	-3.3
KTW	-0.1	-1.7	-2.5
ROA	-0.6	-2.1	-3.3
LAM	-1.0	-2.4	-3.7
EUR	0.1	-1.4	-1.7
EER	-1.0	-3.6	-5.0
ROW	-2.3	-3.5	-5.7

6 Conclusion

A broad body of empirical literature documents persistent differences in plant-level productivity. This literature has also shown that the reallocation of production activities, from less- to more-productive plants, is an important part of aggregate productivity growth. These basic characteristics of industrial organization have important implications for international trade and commercial policy. The unifying theory proposed by Melitz (2003) offers insights into these implications. Our contribution is to present a quantitative assessment of the effects of trade cost changes within this structure.

In the case of a 50% reduction in tariffs on traded manufactured goods the simple-average welfare gains are on the order four times greater in the Melitz structure than in a conventional trade policy model. These gains are complemented, and compounded, when reductions in the inferred fixed costs are also considered. When we add a 50% reduction in cross-border fixed costs to the tariff cuts the welfare gains grow to roughly 30 times what is measured in the constant-returns reference liberalization.

Table 9: Changes in the Number of Operating Firms, N_{rs} (% Change)

	Scenario		
	B	C	D
	Tariff	Fix Cost	Both
Imported Varieties (extensive margin):			
`CHN`	28.9	196.6	275.7
`JPN`	54.6	212.7	377.8
`CAN`	11.0	158.0	190.0
`USA`	8.8	155.0	177.9
`MEX`	207.1	158.6	715.9
`ANZ`	26.1	174.6	230.7
`KTW`	22.1	206.3	266.0
`ROA`	20.8	190.4	246.2
`LAM`	27.2	155.0	225.1
`EUR`	17.4	177.8	225.1
`EER`	20.3	168.0	224.9
`ROW`	12.1	147.8	185.9
Domestic Varieties:			
`CHN`	-2.1	-4.7	-7.3
`JPN`	-2.1	-5.8	-8.9
`CAN`	-2.3	-16.2	-19.6
`USA`	-1.0	-6.6	-7.8
`MEX`	-4.5	-14.7	-20.4
`ANZ`	-6.9	-13.0	-20.3
`KTW`	-3.8	-11.7	-17.0
`ROA`	-4.4	-11.0	-17.0
`LAM`	-3.9	-6.0	-11.1
`EUR`	-1.5	-6.4	-8.2
`EER`	-3.1	-9.2	-12.6
`ROW`	-2.7	-5.2	-8.8
Total Varieties:			
`CHN`	-1.9	-3.4	-5.4
`JPN`	8.3	34.0	61.6
`CAN`	-2.2	-14.5	-17.6
`USA`	0.0	11.1	12.5
`MEX`	26.6	10.7	87.6
`ANZ`	-6.8	-12.4	-19.5
`KTW`	-2.6	-1.8	-4.2
`ROA`	-3.9	-7.4	-12.2
`LAM`	-3.8	-5.5	-10.3
`EUR`	0.7	15.3	19.2
`EER`	-3.0	-8.1	-11.1
`ROW`	-2.6	-4.6	-8.1

We offer a few caveats to consider, given our application and methods. First, we employ a novel method for measuring unobserved fixed costs. We depart from the econometric literature by employing a nonlinear estimation that includes *extensive-form* conditions as side constraints. Our focus is on arriving at fitted values, while maintaining consistency between the econometric and simulation models. Our estimation method is a departure from traditional calibration methods used to fit simulation models; we do not allow preference-bias parameters to drive trade. The onus of explaining the observed pattern of trade is on the theory and the standard parameters that appear in the theory, not on added preference-bias parameters. Our estimates of fixed costs play a large role in determining the trade pattern. It is generally accepted by economists that unobserved trade costs are an important component of the world trade equilibrium. We follow one of the few paths available, which is to adopt the proposed structure and use it to inform unobservables from the observables.

The second major caveat that we place on our results involves the data. We accept the GTAP data as given and further aggregate it. This is useful in terms of reducing computational complexity and in allowing us to efficiently summarize reports. The GTAP data are balanced; they have already been fitted to a set of fundamental accounting identities. The data are consistent with general-equilibrium adding-up restrictions, but the original fitting procedure weakens the validity of any statistical inference that one might draw from our estimation.

The usual aggregation biases abound in our data, and we have additional concerns given the theory's focus on firm-level behavior. Our aggregate manufacturing sector is not a satisfying definition of an industry or product. Regional aggregation is also problematic. The aggregate rest-of-world region is actually numerous small disjoint markets rather than a large integrated market. We probably overstate the fixed costs of entering the aggregate regions because large fixed costs are necessary for explaining the relative lack of trade with artificially large regions.

We thus present our estimates conditional on the particular aggregation of the data, the assumed structure, and our maintained hypotheses about key structural parameters. We see important extensions in the area of regional and industry disaggregation. We are somewhat unique in our development of an econometric method that facilitates direct, and consistent, welfare analysis of policy. Others may find this departure from standard regression analysis useful and relevant. We are firmly within the empirical-trade tradition which places theory, not established statistical methods, as the foundation for analysis.

References

Alvarez, Fernando, and Robert E. Lucas (2007) 'General equilibrium analysis of the Eaton-Kortum model of international trade.' *Journal of Monetary Economics* 54, 1726–1768

Anderson, James, and Eric van Wincoop (2003) 'Gravity with gravitas: A solution to the border puzzle.' *American Economic Review* 93(1), 170–192

_ (2004) 'Trade costs.' *Journal of Economic Literature* 42(3), 691–751

Anderson, Kym, William J. Martin, and Dominique van der Mensbrugghe (2005) 'Global impacts of the Doha scenarios on poverty.' *World Bank Policy Research Working Paper Series 3735*

Arkolakis, Costas, Svetlana Demidova, Peter J. Klenow, and Andrés Rodríguez-Clare (2008) 'Endogenous variety and the gains from trade.' *American Economic Review: Papers & Proceedings* 98(2), 444–450

Armington, Paul S. (1969) 'A theory of demand for products distinguished by place of production.' *Staff Papers - International Monetary Fund* 16(1), 159–178

Aw, Bee Yan, Xiaomin Chen, and Mark J. Roberts (2001) 'Firm-level evidence on productivity differentials and turnover in Taiwanese manufacturing.' *Journal of Development Economics* 66, 51–86

Baldwin, Richard, and Rikard Forslid (forthcoming) 'Trade liberalization with heterogeneous firms.' *Review of Development Economics*

Balistreri, Edward J., and James R. Markusen (2009) 'Sub-national differentiation and the role of the firm in optimal international pricing.' *Economic Modelling* 26(1), 47 – 62

Balistreri, Edward J., and Russell H. Hillberry (2007) 'Structural estimation and the border puzzle.' *Journal of International Economics* 72(2), 451–463

— (2008) 'The gravity model: An illustration of structural estimation as calibration.' *Economic Inquiry* 46(4), 511–527

Bartelsman, Eric J., and Mark Doms (2000) 'Understanding productivity: Lessons from longitudinal microdata.' *Journal of Economic Literature* 38(3), 569–594

Behrens, Kristian, Giordano Mion, Yasusada Murata, and Jens Sudekum (2008) 'Trade, wages, and productivity.' *IZA Discussion Paper No. 3682*

Bernard, Andrew, and J. Bradford Jensen (1999) 'Exceptional exporter performance: Cause, effect, or both?' *Journal of International Economics* 47(1), 1–25

Bernard, Andrew B., J. Bradford Jensen, and Peter K. Schott (2006) 'Trade costs, firms and productivity.' *Journal of Monetary Economics* 53(5), 917–937

Bernard, Andrew B., Jonathan Eaton, J. Bradford Jensen, and Samuel Kortum (2003) 'Plants and productivity in international trade.' *American Economic Review* 93, 1268–1290

Bernard, Andrew B., Stephen Redding, and Peter K. Schott (2007) 'Comparative advantage and heterogeneous firms.' *Review of Economic Studies* 74, 31–66

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

Brown, Drusilla K. (1987) 'Tariffs, the terms of trade, and national product differentiation.' *Journal of Policy Modeling* 9(3), 503–526

Chaney, Thomas (2008) 'Distorted gravity: The intensive and extensive margins of international trade.' *American Economic Review* 98(4), 1707–1721

Dimaranan, Betina V. (Editor) (2006) *Global Trade, Assistance, and Production: The GTAP 6 Data Base* (West Lafayette: Center for Global Trade Analysis, Purdue University)

Eaton, Jonathan, and Samuel Kortum (2003) 'Technology, geography and trade.' *Econometrica* 70(5), 1741–1780

Eaton, Jonathan, Samuel Kortum, and Francis Kramarz (2004) 'Dissecting trade: Firms, industries and export destinations.' *American Economic Review* 94(2), 150–154

Evenett, Simon, and Anthony Venables (2002) 'Export growth in developing countries: Market entry and bilateral trade flows.' *mimeo, University of St. Gallen*

Feenstra, Robert C. (1994) 'New product varieties and the measurement of international prices.' *The American Economic Review* 84(1), 157–177

— (2009) 'Measuring the gains from trade under monopolistic competition.' *mimeo*, University of California, Davis

Foster, Lucia, John Haltiwanger, and C.J. Krizan (2001) 'Aggregate productivity growth: Lessons from the microeconomic evidence.' In *New Directions in Productivity Analysis,* ed. Edward Dean, Michael Harper, and Charles Hulten (University of Chicago Press)

Helpman, Elhanan, Marc Melitz, and Yona Rubinstein (2008) 'Estimating trade flows: Trading partners and trading volumes.' *The Quarterly Journal of Economics* 123(2), 441–487

Hillberry, Russell, and David Hummels (2008) 'Trade responses to geographic frictions: A decomposition using micro-data.' *European Economic Review* 52(3), 527–550

Hillberry, Russell H., Michael A. Anderson, Edward J. Balistreri, and Alan K. Fox (2005) 'Taste parameters as model residuals: Assessing the 'fit' of an Armington trade model.' *Review of International Economics* 13(5), 973–984

Hummels, David (2001) 'Toward a geography of trade costs.' *mimeo*, Purdue University

Hummels, David, and Peter Klenow (2005) 'The variety and quality of a nation's exports.' *American Economic Review* 95(3), 704–723

Kehoe, Timothy J. (2005) 'An evaluation of the performance of applied general equilibrium models of the impact of NAFTA.' In *Frontiers in Applied General Equilibrium Modeling: Essays in Honor of Herbert Scarf,* ed. Timothy J. Kehoe, T.N. Srinivasan, and John Whalley (Cambridge University Press) pp. 341–377

Kehoe, Timothy J., and Kim J. Ruhl (2002) 'How important is the new goods margin in international trade?' *mimeo, University of Minnesota*

Klenow, Peter, and Andres Rodriguez-Clare (1997) 'Quantifying variety gains from trade liberalization.' *mimeo*, University of Chicago

Mansur, Ahsan, and John Whalley (1982) 'A decomposition algorithm for general equilibrium computation with applications to international trade models.' *Econometrica* 50(6), 1547–1557

Mayer, Thierry, and Soledad Zignago (2006) 'Notes on CEPII's distance measures.' *CEPII explanatory note.* http://www.cepii.fr/distance/noticedist_en.pdf

Melitz, Marc J. (2003) 'The impact of trade on intra-industry reallocations and aggregate industry productivity.' *Econometrica* 71(6), 1695–1725

Pavcnik, Nina (2002) 'Trade liberalization, exit, and productivity improvement: Evidence from Chilean plants.' *Review of Economic Studies* 69(1), 245–276

Roberts, Mark J., and James Tybout (1997) 'The decision to export in Columbia: An empirical model of entry with sunk costs.' *American Economic Review* 87(4), 545–564

Romer, Paul M. (1994) 'New goods, old theory, and the welfare costs of trade restrictions.' *Journal of Development Economics* 43(1), 5–38

Rutherford, Thomas F., and David G. Tarr (2002) 'Trade liberalization, product variety and growth in a small open economy: a quantitative assessment.' *Journal of International Economics* 56(2), 247–272

Su, Che-Lin, and Kenneth L. Judd (2008) 'Constrained optimization approaches to estimation of structural models.' *SSRN eLibrary*

A Appendix: Industrial organization and the welfare gains from trade liberalization (not for publication)

Arkolakis et al. (2008) identify a simple constant-elasticity relationship between iceberg trade costs and trade flows across a family of models. This family of models include Melitz (2003) models and Armington (1969) based models. Furthermore, the particular structure assumed does not give discernibly different welfare impacts. Arkolakis et al. (2008) show that, given appropriate parameterizations of the trade response, the welfare gains from iceberg trade cost reductions are equivalent across models. They conclude that "(c)ontrary to what many have claimed, new trade models do not really offer new gains from trade, given observed trade levels." These conclusions would seem to be at odds with our results, which indicate important differences between the Armington and Melitz models.

In this appendix we demonstrate that the Melitz (2003) and Armington (1969) structure have equivalent impacts on trade and welfare under the particular assumptions of Arkolakis et al. (2008). We also demonstrate, however, that the Arkolakis et al. (2008) results depend critically on the assumption that the supply of labor to the Melitz sector is perfectly inelastic. We introduce a second constant returns to scale sector, and allow labor to move between the two sectors. We show that the equivalence of the two models occurs in a single case, when the supply of labor to the Melitz sector is perfectly inelastic. Elastic factor supplies generate room for entry into the Melitz sector, and this breaks the equivalence. The equivalence found by Arkolakis et al. (2008) is not particularly relevant to our present application because we have multiple factor and multiple sectors, each of which use factors in different proportions and compete for factor services. Unlike Arkolakis et al. (2008), this indicates that a liberalization of the heterogeneous-firms sector will induce entry and net variety effects.

A.1 Models

We compare the results of trade cost removal across a Melitz (2003) based model and an Armington (1969) based model with equivalent treatments of factor supply. In our simulations we include three regions (indexed by r or s) each with an endowment of labor that can be used in a traded sector or directly consumed in utility (which is the same as having a second constant-returns sector that just uses labor). The trade theories are well developed in the literature, so we simply present our notation and the equilibrium conditions for each model. The theory presented by Arkolakis et al. (2008) is a special case of the Melitz model when we parameterize it such that the implied factor-supply elasticity to the traded sector is zero.[39]

A.1.1 Notation

Variables:

---

[39] Either the share of the non-traded sector can be set to zero, or the top-level elasticity of substitution between the traded and non-traded goods can be set to one, in order to generate perfectly inelastic labor supply to the traded sector.

W_r	Welfare
e_r	Unit expenditure index
P_r	Price index on traded goods composite
V_r	Value of demand for traded composite
M_r	Total number of entered heterogeneous firms
N_{rs}	Number of operating firms from r in market s
$\tilde{r}_{rs}$	Revenue of the average firm gross of trade costs
$\tilde{p}_{rs}$	Price of the average firm's product gross of trade costs
$\tilde{\varphi}_{rs}$	Productivity of the average firm
w_r	Wage
Y_r	Nominal income

Exogenous instruments:

τ_{rs}	Iceberg trade cost factor
α	Top-level elasticity or substitution between traded and non-traded goods

Fixed Parameters:

$a = 3.4$	Pareto shape parameter
$b = 0.2$	Lower bound on Pareto distribution
$\sigma_M = 3.8$	Melitz model elasticity of substitution between traded varieties
$\sigma_A = 4.4$	Armington elasticity of substitution (set to $a+1$)
$\delta = 0.05$	Probability of firm death
$\gamma = 0.5$	Value share of traded sector in consumption
$\bar{L}_r = 2/3$	Labor endowment
$\tau^0_{rr} = 1$	Assumed benchmark home-market trade cost factor
$\tau^0_{rs} = 2$	Assumed benchmark external-market trade cost factor
$m^0_r = 10$	Assumed benchmark number of entered firms (draws)
$n^0_{rr} = 9.5$	Assumed benchmark number of operating home firms
$n^0_{rs} = 0.6$	Assumed benchmark number of exporting firms

Calibrated Parameters:

f_{rs}	Fixed operating cost on the r to s link
f^e_r	Fixed cost of a productivity draw
ψ_T	Top-level CES parameter on traded sector
ψ_L	Top-level CES parameter on non-traded sector
ξ_{rs}	Armington CES distribution parameters

A.1.2 Melitz General Equilibrium

The CES top-level unit expenditure function:

$$e_r = \left(\psi_T P_r^{1-\alpha} + \psi_L w_r^{1-\alpha}\right)^{1/(1-\alpha)}. \tag{A1}$$

Of course, if $\alpha = 1$ this reverts to the familiar Cobb-Douglas form of the unit expenditure function. The CES price index on traded goods is given by the following:

$$P_r = \left(\sum_s N_{sr} \tilde{p}_{sr}^{1-\sigma_M}\right)^{1/(1-\sigma_M)}. \tag{A2}$$

Demand for the traded composite in nominal units:

$$V_r = \psi_T Y_r \left(\frac{e_r}{P_r}\right)^{\alpha-1}. \tag{A3}$$

Firm-level demand in nominal units:

$$\tilde{r}_{rs} = V_s \left(\frac{P_s}{\tilde{p}_{rs}}\right)^{\sigma_M-1}. \tag{A4}$$

Firm-level optimal pricing:

$$\tilde{p}_{rs} = \frac{w_r \tau_{rs}}{\tilde{\varphi}_{rs}(1-1/\sigma_M)}. \tag{A5}$$

Equilibrium average productivity based on the Pareto distribution:

$$\tilde{\varphi}_{rs} = b\left(\frac{a}{a+1-\sigma_M}\right)^{1/(\sigma_M-1)} \left(\frac{N_{rs}}{M_r}\right)^{-1/a} \tag{A6}$$

Free entry:

$$w_r \delta f_r^e = \sum_s \frac{N_{rs}}{M_r} \frac{\tilde{r}_{rs}(\sigma_M - 1)}{a\sigma_M}. \tag{A7}$$

Zero cutoff profits:

$$w_r f_{rs} = \frac{\tilde{r}_{rs}(a+1-\sigma_M)}{a\sigma_M}. \tag{A8}$$

Market clearance for labor:

$$\bar{L}_r = \psi_L \frac{Y_r}{e_r}\left(\frac{e_r}{w_r}\right)^{\alpha} + \delta f_r^e M_r + \sum_s N_{rs}\left(f_{rs} + \frac{\tau_{rs}\tilde{r}_{rs}}{\tilde{\varphi}_{rs}\tilde{p}_{rs}}\right) \tag{A9}$$

Nominal Income:

$$Y_r = w_r \bar{L} \tag{A10}$$

Welfare:

$$W_r = Y_r / e_r \tag{A11}$$

Equations (A1) through (A11) form a complete (and relatively transparent) general equilibrium based on the Melitz (2003) theory for the traded sector.

A.1.3 Armington General Equilibrium

Under the Armington assumption of traded regional goods, we have the same unit expenditure function [equation (A1)], but the price index on the imported composite is a direct function of the foreign wages (given one factor and our assumption of constant returns):

$$P_r = \left(\sum_s \xi_{sr} (\tau_{sr} w_s)^{1-\sigma_A} \right)^{1/(1-\sigma_A)}. \tag{A12}$$

The calibration parameter, ξ_{sr}, is added so we can match the initial trade equilibrium implied by the Melitz structure under an arbitrary choice of labor units. The market clearance condition for labor reflects the direct exports:

$$\bar{L}_r = \psi_L \frac{Y_r}{e_r} \left(\frac{e_r}{w_r} \right)^{\alpha} + \sum_s \frac{\xi_{rs} \tau_{rs} V_s}{P_s} \left(\frac{P_s}{\tau_{rs} w_r} \right)^{\sigma_A} \tag{A13}$$

The full Armington general equilibrium includes the common equations (A1), (A3), (A10), and (A11); in addition to the reformulated equations for the price index (A12), and market clearance (A13).

A.2 Calibration

Given the values of the fixed parameters, the calibration parameters for the Melitz model are determined by inverting the equilibrium conditions. The remaining Armington parameters (the ξ_{rs}) are then set such that the Armington and Melitz models have identical benchmark trade flows. Table A1 presents the nominal bilateral trade flows. We normalize on the value of region-1 consumption of Melitz goods so that reported nominal trade flows into region 1 are directly comparable to the shares that are the focus of Arkolakis et al. (2008).[40]

In the calibration we choose labor and welfare units such that the initial wages and true-cost-of-living indexes are one; $w_r^0 = e_r^0 = 1$ (where the superscript indicates the benchmark equilibrium). This is a convenient choice because it simplifies our calculation of the elasticity of labor supply available to the traded sector of the economy. The relevant residual labor supply function is given by

$$g(w_r) = \bar{L}_r - \psi_L \frac{Y_r}{e_r(w_r)} \left(\frac{e_r(w_r)}{w_r} \right)^{\alpha}, \tag{A14}$$

---

[40] We maintain this price normalization in the scenarios that follow. Reported nominal imports into region 1 can be directly interpreted as consumption shares of Melitz goods.

Table A1: Benchmark nominal trade Melitz = Armington models

		Destination Region		
		1	2	3
Origin Region	1	0.850	0.075	0.075
	2	0.075	0.850	0.075
	3	0.075	0.075	0.850

which is derived from equation (A9). Substituting in the unit expenditure function and $Y_r = w_r \bar{L}_r$, and then calculating the elasticity evaluated at the benchmark ($w = e(w) = 1$), yields

$$\eta = (1-\gamma)(\alpha - 1) = \frac{\alpha - 1}{2}. \tag{A15}$$

So, we use the instrument, α, to control the implied labor supply elasticity. If we set $\alpha = 1$ then the elasticity is zero and we have a model that is consistent with Arkolakis et al. (2008).[41]

A.3 Experiment and Results

In order to compare the Armington versus Melitz models we compute a simple experiment where we remove iceberg trade costs between regions one and two.[42] Using the instrument α, we control the implied labor-supply elasticity (η) faced by the trade sectors. Setting the Armington elasticity as suggested ($\sigma_A = a + 1$), we find that the welfare impacts of removing the iceberg costs are different across the models, except in the special case that the implied labor-supply elasticity is exactly zero.

Figure 1 plots the region-1 welfare impact of the trade-cost removal as a function of the implied labor-supply elasticity. Notice that the welfare impacts are only equivalent at $\eta = 0$. The results for region 2 are identical to region 1, because of the symmetry built into our illustrative model. Figure 2 shows the welfare impacts on the third region for the same set of experiments. Although the curves in figure 2 intersect twice, it is only at $\eta = 0$ that we have equivalence in the models across the multiregion equilibrium.

One key feature of the environment set up by Arkolakis et al. (2008) is that there is never any entry. Labor supply is perfectly inelastic so all of the adjustments in firm revenues and number of operating firms shows up in the wage. Changes in nominal entry costs are mirrored by changes in expected profits, so equation (A7) is satisfied with no changes in M_r. The number of entered firms is unaffected by changes in iceberg costs, as long as $\eta = 0$. At $\eta \neq 0$ the wage only partially absorbs the adjustments in the industrial organization and M_r

---

[41] We also ran experiments where we fixed $\alpha = 3$ and calibrated γ to the assumed labor-supply elasticity. Again, at values of η above zero the models did not generated the same results. It is only in the special case that $\gamma = 1$, as in the Arkolakis et al. (2008) environment, that equivalence between the Armington and Melitz models is obtained.

[42] The simulations are computed using GAMS software. All programs are available from the authors.

Figure 1: Region-1 welfare comparison ($\sigma_A = a + 1$)

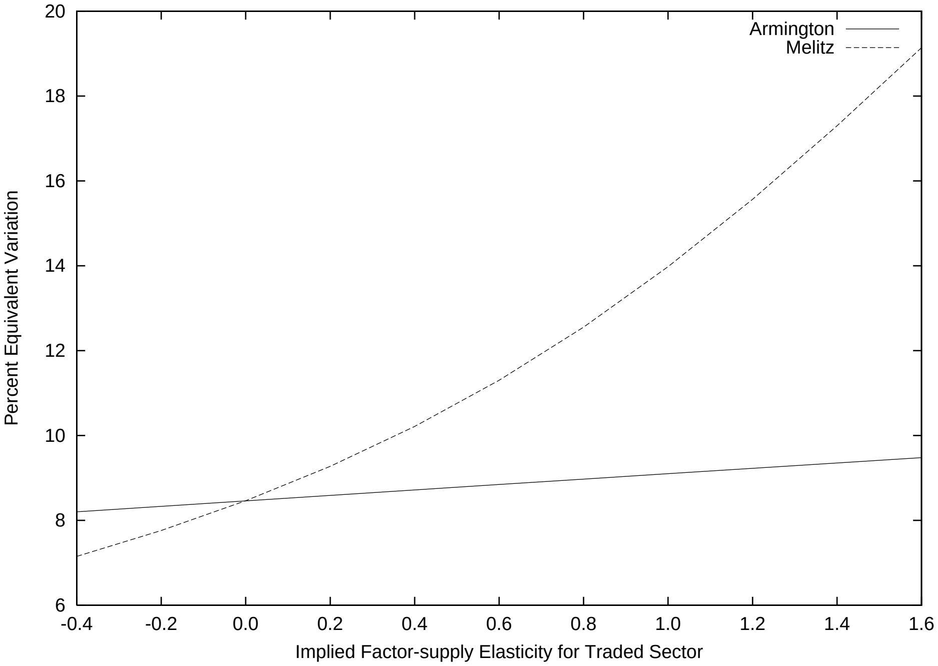

Figure 2: Region-3 welfare comparison ($\sigma_A = a + 1$)

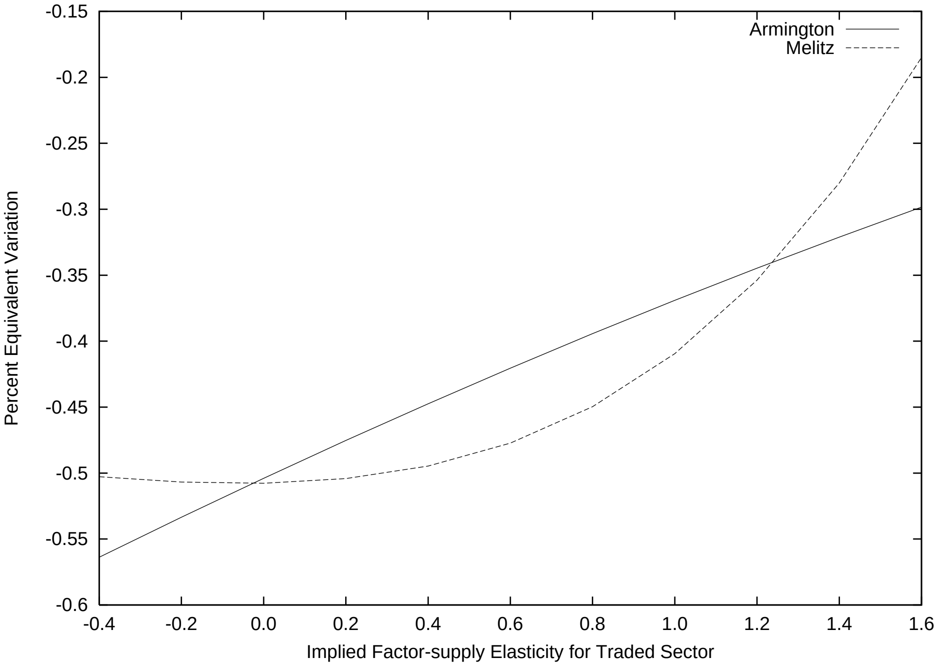

Table A2: Heterogeneous-firms model region-1 entry and consumption of varieties

		Benchmark	Scenario $\eta = 0$	Scenario $\eta = 1$
Entered Firms:	M_1	10.00	10.00	12.30
Varieties Consumed:	$N_{1,1}$	9.50	5.47	6.77
	$N_{2,1}$	0.59	3.61	4.47
	$N_{3,1}$	0.59	0.48	0.53
Total Varieties:	$\sum_r N_{r,1}$	10.69	9.55	11.77
Feenstra Ratio:	$(\lambda_1^1/\lambda_1^0)^{-1/(\sigma_M-1)}$		1.00	1.08

changes. In table A2 we present the basic industrial organization in the Melitz model in the benchmark and scenarios with different labor supply responses. At $\eta = 1$ we have entry as labor is drawn into the Melitz sector.

At $\eta = 0$ table A2 shows the "anti-variety effect" emphasized by Baldwin and Forslid (forthcoming) and Arkolakis et al. (2008), where the new import varieties generated by trade liberalization are more than offset by lost domestic varieties. Notice, however, that the total number of varieties consumed in region 1 goes from 10.69 in the benchmark to 11.77 in the scenario, when $\eta = 1$. The anti-variety effect is dominated when there is enough response in factor supplies. Feenstra (2009) emphasizes, however, that because these varieties enter the expenditure system at different prices we cannot simply count up varieties and infer variety gains or losses. Feenstra shows that variety gains, when comparing equilibriums t versus $t-1$, are given by deviations in the ratio $(\lambda_r^t/\lambda_r^{t-1})^{-1/(\sigma_M-1)}$ from unity, where λ_r^τ represents region-r's share of expenditures at equilibrium τ on goods available in both equilibria to the total expenditures at τ. We confirm the Feenstra (2009) analytical result that there are no *import*-variety gains in the Melitz structure (for the case that $\eta = 0$), but we find that the variety gains reemerge when we allow resources to be drawn into the Melitz sector.

To emphasize the fragility of the equivalence between the Armington and Melitz models we look at trade flows. In the case that $\eta = 0$ the trade patterns before and after the removal of trade costs are identical, but we can only match the multiregion trade flows in this special case. Table A3 shows the bilateral nominal-trade shares after the removal of trade costs between regions 1 and 2 in the case that $\eta = 0$. In table A4 we show that the scenario trade shares are different if there is a labor supply response.

One might ask if the σ_A parameter can be set to match the trade reactions in the Melitz model when $\eta \neq 0$? The answer is no. In Table A5 we decrease σ_A to match the Melitz-model domestic consumption share for region 1 in the scenario. It is possible to match some of the trade responses, but the errors on other flows in the bilateral matrix become larger. In fact, we found that $\sigma_A = a + 1$ is arguably the best setting, because it minimizes the sum of squared errors between the Armington and Melitz trade flow responses, but this sum of squared errors is only zero at $\eta = 0$.

Table A3: Scenario nominal trade Melitz = Armington with no labor-supply response ($\eta = 0$)

		Destination Region		
		1	2	3
Origin Region	1	0.489	0.455	0.056
	2	0.455	0.489	0.056
	3	0.056	0.056	0.820

Table A4: Scenario nominal trade with labor-supply response ($\eta = 1$)

		Melitz			**Armington**		
		Destination Region			Destination Region		
		1	2	3	1	2	3
Origin Region	1	0.493	0.458	0.049	0.491	0.457	0.052
	2	0.458	0.493	0.049	0.457	0.491	0.052
	3	0.049	0.049	0.635	0.052	0.052	0.707

Table A5: Armington-model scenario nominal trade with σ_A adjusted

		Destination Region					
		1		2		3	
Origin Region	1	**0.493**	=**Melitz**	0.455	$\neq$ Melitz	0.052	$\neq$ Melitz
	2	0.455	$\neq$ Melitz	**0.493**	=**Melitz**	0.052	$\neq$ Melitz
	3	0.052	$\neq$ Melitz	0.052	$\neq$ Melitz	0.708	$\neq$ Melitz

A.4 Conclusion

We agree with Arkolakis et al. (2008) that for the purposes of comparison, models should be parameterized so as to produce similar (and where possible equivalent) responses to trade cost reductions. In an important theoretic contribution Arkolakis et al. (2008) show, in a one sector model with heterogeneous-firms, that the new theories “do not really offer new gains from trade, given observed trade levels.” In our simple comparison of Armington and Melitz formulations with iceberg trade costs, we set the Armington elasticity of substitution equal to the Melitz Pareto-shape parameter plus one, because this setting minimizes the sum of the squared differences in trade flow reactions. If the labor-supply elasticity faced by the traded sector is zero, the trade flow reactions and welfare impacts of trade-cost removal are identical across the Armington and Melitz models. We show, however, that this result is fragile. If the labor-supply elasticity is different than zero the industrial organization starts to matter. Firm entry and import variety effects becomes important at labor-supply elasticities above zero, and the Armington elasticity of substitution cannot be adjusted to compensate for the industrial organization effects in a multiregion world.