ITI
7/30/07
11:00 am

Gravity, Productivity and the Pattern of Production and Trade*

James E. Anderson
Boston College and NBER

July 9, 2007

Abstract

Trade frictions reduce productivity. This paper gathers and develops the implications by embedding gravity in a wide class of general equilibrium production models. Clean implications for the global equilibrium pattern of production and trade are extracted for the specific factors model of production.

JEL Classification: F10, D24

*For presentation at the NBER ITI Summer Institute. I thank participants in seminars at Drexel University and Brandeis University for comments on an earlier version.

Trade costs should intuitively have a big impact on productivity and the pattern of production and trade. Differential access to markets should alter the pattern of returns to countries' factors of production. Differential trade costs across products should tilt the pattern of production. The impacts should be big because trade costs are big. Several key problems have prevented development of these intuitions about the implications of trade costs. This paper offers a promising solution.

Productivity measurement in a world economy with trade frictions is problematic for three reasons. First, distribution costs affect prices on both the supply and demand sides of the market, while productivity reflects only the supply side incidence of the costs. Second, standard productivity measures in distribution sectors fail to capture the effects of globalization because quality changes produce savings that are hidden in the books of the ultimate buyers and sellers. Inference from gravity models of trade suggests that these effects are big. Third, outsourcing implies changes on the extensive margin that require different accounting than the intensive margin changes that standard methods measure.

All three problems are addressed in this paper by building on recent progress in understanding, interpreting and using the gravity model. The incidence problem is solved by extending the economic theory of gravity (Anderson, 1979; Anderson and van Wincoop, 2003, 2004). Outward and inward multilateral resistance give respectively the supply side and demand side incidence in general equilibrium. At the same time they consistently aggregate bilateral trade costs. Thus outward multilateral resistance indexes are equivalent to a set of productivity penalties, as if each producing sector in each economy traded with a single world market at varying incidence of trade costs. Selection into trade on the extensive margin is incorporated in the gravity model by Helpman, Melitz and Rubinstein (2007), with implications for productivity drawn out here and related to outsourcing.

Productivity differences have important implications for the equilibrium pattern of production and trade. These implications depend on the specification of technology and preferences.[1] Sharp implications are drawn out here for the specific gravity model — gravity embedded in the specific factors model of production. The equilibrium pattern of production is explained by specific factor endowments (a supply shifter), taste parameters (a demand

[1]As is well known, even with convex technology it is not generally possible to derive a perfect negative correlation of productivity penalties with output changes.

shifter) and the productivity penalty imposed by trade costs (outward multilateral resistance).

As context, the previous literature contains only one case in which the equilibrium pattern of production is characterized in a world of costly trade — the Ricardian continuum model of Eaton and Kortum (2002). The Eaton-Kortum model features endogenous generation of the range of goods produced and traded bilaterally, driven by the process of technology generation. Labor productivities are drawn from a Frechet probability distribution. Sufficiently good draws survive to produce in general equilibrium, trading with those destinations sufficiently inexpensive to reach. Since each good is homogeneous in demand, the Eaton-Kortum model is exclusively focused on the extensive margin. In contrast, the specific gravity model has both intensive and extensive margins active. Labor productivity is endogenous, with the equilibrium productivities determined as a function of factor endowments and trade costs, along with parameters of technology and preferences. The Heckscher-Ohlin-Vanek factor content model also features the role of factor endowments, but only in a frictionless world.[2]

Section 1 sets the stage by describing the incidence and aggregation problems in partial equilibrium. Assuming that a solution can be found, the supply side incidence of trade frictions can be treated as equivalent to sectoral productivity penalties in the standard abstract model of production and trade. Multifactor productivity is the ratio of the usual Hicks neutral productivity parameter to the supply side incidence measure. Sectoral measures aggregate to a productivity measure for each country in the world economy for given equilibrium prices.

Section 2 provides the solution to incidence and aggregation in general equilibrium. Given the sectoral supply and expenditure shares determined by the model of Section 1, the bilateral allocation of trade determines the multilateral resistance indexes that determine incidence and deliver the sectoral allocation of production and expenditure. Full general equilibrium is achieved with mutual consistency of the two modules. Section 2 characterizes the relationship of multilateral resistances to the pattern of production and expenditure at world equilibrium prices and draws some lessons for productivity measurement. Section 3 sets out the specific factors model of pro-

[2]Davis and Weinstein (2001) integrate trade costs inferred from gravity with the two factor Heckscher-Ohlin continuum model, but only on the demand side. In parallel work I plan to draw out the implications of gravity in the HOV setting.

duction in a special case. The world equilibrium reduced form pattern of production and trade that results is set out and characterized in Section 4. Section 4 goes on to characterize productivity in terms of its reduced form drivers. Section 5 extends the discussion to treat intermediate products trade and selection into exporting. Section 6 concludes.

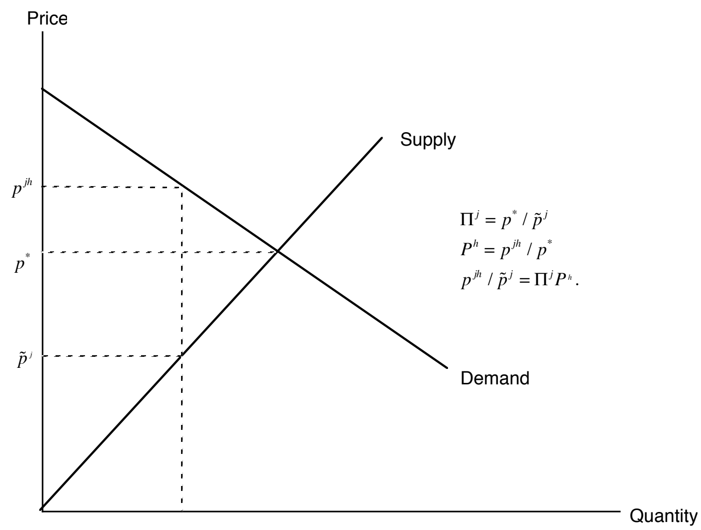
1 Trade Frictions and Productivity

Each country produces and distributes goods to its trading partners subject to trade frictions. Suppose that the aggregate incidence of these frictions on the supply side can be represented by an index Π_k^j for each product category k in each country j. With unit production cost $\widetilde{p}_k^j$ in country j, it is as if there was an average ('world') destination price for goods k delivered from j, $p_k^j = \widetilde{p}_k^j \Pi_k^j$. The incidence of the trade frictions will be solved for in general equilibrium in the next section, but intuition is aided with a review of the incidence problem in partial equilibrium.

1.1 Incidence

The incidence of a trade cost in partial equilibrium is presented in Figure 1.

Figure 1. Incidence of Trade Cost

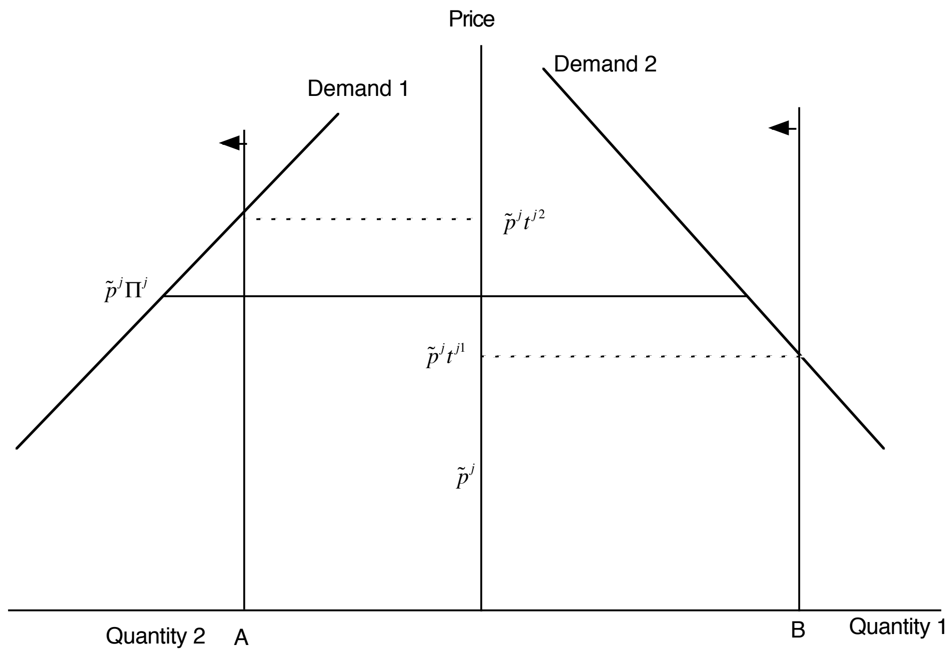


The price of the representative good from origin j at location h is p^{jh} while its unit cost is $\widetilde{p}^j$. The full trade cost friction marks up the unit production cost by $\Pi^j P^h$ to yield the buyer's price p^{jh}. The standard incidence analysis uses the hypothetical frictionless equilibrium with price p^* to split the cost into two components, of which Π^j falls on the supply side of the market and P^h falls on the demand side. The distribution cost component Π

is conceptually identical to a productivity penalty that shifts the cost of production and distribution upward to where the hypothetical supply schedule intersects the horizontal line at p^* at the equilibrium quantity.

Figure 2 portrays the aggregation of supply side incidence in partial equilibrium for the case of two markets. The equilibrium factory gate price $\widetilde{p}^j$ is preserved by maintaining the total quantity shipped while replacing the nonuniform trade costs with the uniform trade cost Π^j. An analogous diagram (not shown) illustrates the aggregation of demand side incidence.

Figure 2. Quantity-Preserving Aggregation



Total shipments AB are preserved by moving the goalposts left such that a uniform markup is applied to each shipment.

1.2 Productivity

Accounting for multilateral resistance as a productivity component uses familiar principles. The key building block is the gross domestic product (GDP) function. It is written as $g(\widetilde{p}^j, v^j)$ where v^j is the vector of factor endowments. g is convex and homogeneous of degree one in prices, by its maximum value properties.

The supply vector to *final demand* is given by $g_{\widetilde{p}}$, by Hotelling's lemma, using the convention that subscripts with variable labels denote partial differentation with respect to the variable. The production share of good k in country j is given by

$$s_k^j = g^j_{\widetilde{p}^j_k} \widetilde{p}^j_k / g^j.$$

Now consider productivity accounting with trade frictions. Take p^j as a given vector of 'world' prices. Then $\widetilde{p}^j_k = p^j_k / \Pi^j_k, \forall k$. The aggregate productivity penalty due to trade frictions is given by

$$\bar{\Pi}^j \equiv g^j(p^j, v^j)/g^j(\widetilde{p}^j, v^j).$$

The logic uses the distance function. The reference GDP is that for a frictionless economy, $\Pi_k = 1, \forall k$. The uniform productivity penalty that is equivalent to the vector of productivity penalties satisfies

$$g(p/\bar{\Pi}, v) = g(\widetilde{p}, v).$$

GDP is homogeneous of degree one in the prices, hence $\bar{\Pi}$ has the explicit solution given. $\bar{\Pi}$ is equal to the cost of delivered goods relative to the cost of production. The GDP price deflator (i.e., the price index for delivered goods) is given by

$$g(p, \cdot) = \bar{\Pi} g(\widetilde{p}, \cdot).$$

$\bar{\Pi}$ is homogeneous of degree one in $\{\Pi_k\}$. In rates of change the aggregate penalty to productivity imposed by trade frictions is given by $\sum_k s_k \hat{\Pi}_k$. Various approximations to the rate of change of $\bar{\Pi}$ can be used, such as a Laspeyres index. Section 4 offers an exact index based on a special case of production technology, the specific factors/Cobb-Douglas model.

The analysis readily extends to encompass the effects of Hicks neutral technological differences across goods and countries. Let $1/a_k$ denote the productivity parameter (relative to some benchmark) in sector k. The $\widetilde{p}$'s are then reinterpreted as 'efficiency' unit costs, $\widetilde{p}_k = p_k/a_k\Pi_k$. Supply is given by $g_{\widetilde{p}_k} = g_{p_k}/a_k\Pi_k$. Total factor productivity is measured by

$$T = \frac{g(\widetilde{p}, v)}{g(p, v)}.$$

T is decomposable into $1/\bar{a}\bar{\Pi}$ based on $1/\bar{a} = g(\widetilde{p}, v)/g(\{p_k/\Pi_k\}, v)$ and the analogous operation for $\bar{\Pi}$.

The next section will show that the multilateral resistance variables $\{\Pi_k^j, P_k^h\}$ properly decompose the supply and demand side incidence of the aggregated trade frictions in general equilibrium. In contrast, productivity analysis based on a trade-weighted index number of the full bilateral trade costs would overstate their impact unless the incidence fell entirely on the supply side, illustrated by the case where demand is infinitely elastic at price p^* and $P^j = 1$ in Figure 1.

The explanation of differing levels of productivity over time or space involves incorporating differences in equilibrium prices. Figure 1 again illustrates. Imagine shifts in supply or demand, or in the full trade cost. The new solutions for Π and P yield differences to be explained in the reduced form model by the shifts in the exogenous variables, incorporating the shifts in the equilibrium price. Dealing with incidence properly in a reduced form global general equilibrium setting requires building a specific production structure, dealt with in Section 4.

2 Incidence and Aggregation

The incidence of trade costs on productivity is determined in general equilibrium. Insight and the prospect of operationality through aggregation are available with the specializing assumption of *trade separability* — the composition of expenditure or production within a product group is independent of prices outside the product group.

On the supply side, separability is imposed by the assumption the goods from j in class k shipped to each destination are perfect substitutes in supply. On the demand side, separability is imposed by assuming that expenditure on goods class k forms a separable group containing shipments from all origins. Goods are differentiated by place of origin, an assumption that has a deeper rationale in monopolistic competition, as developed in Section 5. This setup enables two stage budgeting analysis. A further specialization to

CES structure for the separable groups yields yields operational multilateral resistance indexes.

Subsection 2.1 sets out the upper level allocation of expenditure and production. Subsection 2.2 derives multilateral resistance from the lower level allocation of goods across trading partners.

2.1 General Equilibrium Allocation

Within class k at each destination h the consumers face prices $p_k^{jh} = \widetilde{p}_k^j t_k^{jh}$ where parametric iceberg trade frictions $t_k^{jh} \geq 1$ margin up the factory gate prices $\widetilde{p}_k^j$.[3]

Expenditure is driven by homothetic preferences that are separable with respect to the partition between goods classes. Then exact price aggregators P_k^h are defined, eventually specified as CES aggregators of the prices of goods from all origins to destination h in class k. The domestic price vector for goods classes at location h is given by $q^h = \{P_k^h\}, \forall k$. The expenditure function is given by $e(q^h)u^h$, imposing identical preferences across countries. The quantity demanded of good k from origin j in destination h is given by $e_{P_k^h}\partial P_k^h/\partial p_k^{jh}$, using Shephard's Lemma.

Market clearance in the world economy requires that for each good k from source j the quantity produced is equal to the quantity demanded. Using the budget constraint for each economy assuming no foreign owned factors or international transfers, $u^h = g^h(\widetilde{p}^h, v^h)/e(q^h)$. [4] The market clearance condition is then expressed as

$$g_k^j(\widetilde{p}^j, v^j) = \sum_h e_{P_k^h}(q^h)\frac{\partial P_k^h}{\partial p_k^{jh}}\frac{g^h(\widetilde{p}^h, v^h)}{e(q^h)}, \forall k, j. \tag{1}$$

With N countries and M goods classes, there are MN factory gate prices ($\widetilde{p}$'s) to be determined by the MN equations, obtaining the user prices from

[3]The analysis abstracts from tariffs for simplicity. Tariffs impose an additional markup factor over the origin price, the difference being that the revenue is collected by the imposing government instead of the original shipper.

[4]When there are final good tariffs, this expression is multiplied by the foreign exchange multiplier. The foreign exchange multiplier under homothetic preferences is equal to $1/(1-\mu T^a)$ where $T^a \in [0, 1)$ is the trade weighted average final goods tariff on the domestic price base and $\mu \in (0, 1)$ is the share of total expenditure falling on tariff-ridden final goods. With intermediate goods tariffs, g in the preceding expression is added to the tariff revenue from the intermediate goods.

the factory gate prices, the arbitrage conditions and the price aggregator definitions. Due to homogeneity, only $MN-1$ relative prices can be determined.

2.2 Multilateral Resistance

Impose CES preferences on the sub-expenditure functions, implying that the within-class expenditure share is given by

$$\frac{\partial P_k^h p_k^{jh}}{\partial p_k^{jh} P_k^h} = \left\{\frac{\beta_k^j \widetilde{p}_k^j t_k^{jh}}{P_k^h}\right\}^{1-\sigma_k},$$

where on the right hand side p_k^{jh} is replaced by $\widetilde{p}_k^j t_k^{jh}$, σ_k is the elasticity of substitution parameter for goods class k and $(\beta_k^j)^{1-\sigma_k}$ is a quality parameter for goods from j in class k, and the true cost of living index P is defined by

$$P_k^h \equiv \sum_j [(\beta_k^j \widetilde{p}_k^j t_k^{jh})^{1-\sigma_k}]^{1/(1-\sigma_k)}.$$

Let the expenditure in destination h on product class k be denoted by E_k^h. The share of expenditure on k from all origins at destination h is given by

$$\theta_k^h = e_{P_k^h} \frac{P_k^h}{e(q^h)}$$

and thus $E_k^h = \theta_k^h g^h$. Let the value of shipments at *delivered* prices from origin h in product class k be denoted by Y_k^h.

Market clearance requires:

$$Y_k^j = \sum_h E_k^h \left\{\frac{\beta_k^j \widetilde{p}_k^j t_k^{jh}}{P_k^h}\right\}^{1-\sigma_k}. \tag{2}$$

Now solve (2) for the quality adjusted factory gate prices $\{\beta_k^j \widetilde{p}_k^j\}$:

$$(\beta_k^j \widetilde{p}_k^j)^{1-\sigma_k} = \frac{Y_k^j}{\sum_h (t_k^{jh}/P_k^h)^{1-\sigma_k} E_k^h}. \tag{3}$$

Define

$$(\Pi_k^j)^{1-\sigma_k} \equiv \sum_h \left\{\frac{t_k^{jh}}{P_k^h}\right\}^{1-\sigma_k} \frac{E_k^h}{\sum_h E_k^h}.$$

Divide numerator and denominator of the right hand side of (3) by total shipments of k and use the definition of Π, yielding:

$$(\beta_k^j \widetilde{p}_k^j \Pi_k^j)^{1-\sigma_k} = Y_k^j / \sum_j Y_k^j. \tag{4}$$

The right hand side is the global expenditure share for class k goods from country j. The left hand side is a 'global behavioral expenditure share', effectively generated by the common global CES preferences over varieties faced by the globally uniform quality adjusted prices $\beta_k^j \widetilde{p}_k^j \Pi_k^j$, understanding that the CES price index is equal to one due to the normalization implied by summing (4):

$$\sum_j (\beta_k^j \widetilde{p}_k^j \Pi_k^j)^{1-\sigma_k} = 1. \tag{5}$$

Now substitute for quality adjusted factory gate prices from (3) in the definition of the true cost of living index, using the definition of the Π's:

$$(P_k^h)^{1-\sigma_k} = \sum_j \left\{ \frac{t_k^{hj}}{\Pi_k^j} \right\}^{1-\sigma_k} \frac{Y_k^j}{\sum_j Y_k^j}. \tag{6}$$

Collect this with the definition of the Π's:

$$(\Pi_k^j)^{1-\sigma_k} = \sum_h \left\{ \frac{t_k^{hj}}{P_k^h} \right\}^{1-\sigma_k} \frac{E_k^h}{\sum_h E_k^h}. \tag{7}$$

These two sets of equations jointly determine the inward multilateral resistances, the P's and the outward multilateral resistances, the Π's, given the expenditure and supply shares and the bilateral trade costs and subject to a normalization such as (5). A normalization of the Π's is needed to determine the P's and Π's because (6)-(7) determine them only up to a scalar.[5]

The multilateral resistances play a key role in determining the bilateral trade flows. The CES specification of within-class expenditure shares implies, after substitution from (3),

$$X_k^{jh} = \left\{ \frac{t_k^{jh}}{\Pi_k^j P_k^h} \right\}^{1-\sigma_k} \frac{Y_k^j E_k^h}{\sum_j Y_k^j}.$$

[5]If $\{P_k^0, \Pi_k^0\}$ is a solution to (6)-(7), then so is $\{\lambda P_k^0, \Pi_k^0/\lambda\}$ for any positive scalar λ; where P_k denotes the vector of P's and the superscript 0 denotes a particular value of this vector, and similarly for Π_k.

Anderson and van Wincoop (2004) show that the multilateral resistance indexes are ideal indexes of trade frictions in the following sense. Replace all the bilateral trade frictions with the hypothetical frictions $\widetilde{t}_k^{jh} = \Pi_k^j P_k^h$. The budget constraint (6) and market clearance (7) equations continue to hold at the same prices, even though individual bilateral trade volumes change.

The implication is that the incidence of bilateral trade costs is decomposed *on average* by $\widetilde{t}$. Thus the gravity model extends the incidence analysis of Figures 1 and 2 to general equilibrium. The demands and supplies of the upper level allocation remain constant, as they do in Figure 1. The aggregation of bilateral t's into 'Π's at constant $\widetilde{p}$ is analogous to the aggregation shown in Figure 2. On average the outward incidence is given by the Π's and the inward incidence is given by the P's.

Thus for each good k in each country j, bilateral distribution costs aggregate to an ideal average outward multilateral resistance Π_k^j, with market clearance at the same producer price and volume. It is as if a single shipment was made to the 'world market' at the average cost. This justifies the treatment of multilateral resistance as a productivity penalty on the activity of production and delivery. On the demand side, similarly, inward multilateral resistance consistently aggregates the demand side incidence of inward trade costs, as if a single shipment was made from the 'world market' at the average trade cost.

Computing the multilateral resistances is readily operational, given estimates of gravity models that yield the inferred t's. In conditional general equilibrium using the data on the global shares, $\{E_k^h / \sum_h E_k^h, Y_k^j / \sum_j Y_k^j\}$ that accompany the estimation, these shares are given. Take quality-adjusted good 1 in country 1 as the numeraire, implying $\beta_1^1 \widetilde{p}_1^1 = 1$. Then

$$(\Pi_1^1)^{1-\sigma_1} = Y_1^j / \sum_j Y_1^j.$$

This restriction plus (6)-(7) suffices to compute the multilateral resistances for goods class 1.[6] For the remaining goods classes there are two alternatives. First, suppose that the upper level general equilibrium model gives the quality adjusted factory gate prices $\{\beta_k^j \widetilde{p}_k^j\}$. Alternatively, in the absence of such information, by appropriate choice of units for all the goods produced by country 1, a *particular* equilibrium can be characterized by $\beta_k^1 \widetilde{p}_k^1 = 1, \forall k$.

[6]The numeraire assignment plays a role analogous to the assignment of the frictionless equilibrium price p^* in the partial equilibrium analysis of Figure 1.

In either case, for the remaining goods classes $k > 1$, utilize $\beta_k^1 \widetilde{p}_k^1$ to calculate

$$(\Pi_k^1)^{1-\sigma_k} = \frac{Y_k^j}{\sum_j Y_k^j} \frac{1}{(\beta_k^1 \widetilde{p}_k^1)^{1-\sigma_k}}, \forall k > 1.$$

Combine this restriction with (6)-(7) to compute the full set of multilateral resistances for all goods classes $k > 1$. Computation of the multilateral resistances is quite fast for modern computers. That is because the system is essentially quadratic in the power transforms of the multilateral resistances.

For full general equilibrium computations that simulate equilibria away from the initial equilibrium above, the upper level general equilibrium model yields the global shares $\{E_k^h / \sum_h E_k^h, Y_k^j / \sum_j Y_k^j\}$ and the normalized quality adjusted factory gate prices $\{\beta_k^j \widetilde{p}_k^j\}$ that are the inputs into the computation of the multilateral resistances from (6)-(7) subject to numeraire choice or other normalization.

There are important regularities in the cross section pattern of multilateral resistance. At the initial conditional general equilibrium, for each good:

Proposition 1 *Given $\sigma_k > 1$, if the trade costs are uniform border barriers, the multilateral resistances (inward and outward) are decreasing in the supply shares of economies and increasing in the expenditure shares of economies. For given expenditure shares, multilateral resistances are increasing in net import shares.*[7]

The proof is in the Appendix. The intuition is this. The larger the supply share, all else equal, the more trade will be domestic, not subject to the border barrier. This lowers outward multilateral resistance. Conversely, the larger the expenditure share, all else equal, the more trade is subject to the border barrier and thus the larger is inward multilateral resistance. General equilibrium links the outward and inward multilateral resistances together. While the uniform border barrier assumption is special, the intuition of Proposition 1 should apply more generally.

Another important implication of Proposition 1 is that productivity analysis that neglects the structure of trade frictions will tend to confound economies of scale with the effects of trade frictions. Large economies tend to have lower multilateral resistance and thus higher productivity even in the absence, as

[7] The proposition extends that of Anderson and van Wincoop (2003), which deals with a the introduction of a small uniform border barrier in a one good balanced trade economy for which $P^j = \Pi^j$.

here, of conventional scale economies. Conversely, neglecting scale economies will tend to overstate the effect of trade frictions.

The outward multilateral resistance variables $\{\Pi_k^j\}$ are endogenous with respect to $\{Y_k^j, E_k^j\}$ and the factory gate prices. A few simple benchmark cases yield useful analytic results that anchor intuition.

One important benchmark is invariance. Invariance occurs the case of uniform factor endowment growth everywhere in the world. Multilateral resistances are constant because all shares are constant. A second benchmark case gauges the significance of trade frictions for productivity. Imagine a pure globalization shock in which trade costs fall uniformly by 4 percent — the world gets literally smaller by 4 percent. Since (6)-(7) is homogeneous of degree 1/2 in the t's, all Π's fall by 2 percent and hence productivity rises by 2 percent everywhere in the world.[8] All relative prices remain constant, hence all shares are constant. Welfare rises everywhere by 4 percent because all P's fall by 2 percent while GDP rises by 2 percent due to the 2 percent fall in the Π's. A third benchmark is given by Anderson and van Wincoop (2003), who show that the introduction of a small uniform border barrier in a frictionless world with balanced trade will raise the multilateral resistance of small countries by more than that of large countries.

Beyond these cases, even a simple extension to the comparative statics of discrete uniform border barriers defeats analytics. In general, asymmetric declines in trade frictions and growth in factor endowments have asymmetric effects on multilateral resistance and productivity with even more complexity. The benchmark cases above do provide some insight to guide future simulations.

3 The Specific Factors Model

The causal links between multilateral resistance and the allocation of production and expenditure are clarified by considering the special case of the specific factors model. The implications for the equilibrium pattern of production and trade are very sharp in a useful special case.

Labor is intersectorally mobile, while there are sector specific factors in

[8] Another important implication of homogeneity is that gravity models alone can only provide information about *relative* trade costs. It is convenient to normalize by some *presumptively* small bilateral cost t_k^{jj} where region j is selected for its plausibly low internal distributive frictions.

fixed supply. The latter can be regarded as possibly mobile in the long run, an interpretation used below for reference. The sectorally fixed supply can be motivated by adjustment costs of various sorts. One form useful for future developments will be fixed costs of entry, providing a link to recent theories focused on firm heterogeneity and its implications for productivity and trade.

Supply understood as deliveries to final demand in product class k is given by

$$X_k = f^k(L_k, K_k)/a_k\Pi_k, \forall k \tag{8}$$

where the country index superscript j is omitted for notational ease. K_k is the specific factor endowment, possibly a bundle of such factors. f^k is a concave (usually homogeneous of degree one) production function. Labor L_k is mobile across sectors, with efficient allocation implied by the value of marginal product conditions

$$w = f^k_L p_k/a_k\Pi_k, \forall k \tag{9}$$

Labor market clearance implies

$$\sum_k L_k = L. \tag{10}$$

Gross domestic product $\sum_k \widetilde{p}_k X_k$ is given by the maximum value GDP function $g(\widetilde{p}, L, \{K_k\})$. As a reminder, $\widetilde{p} = \{p_k/a_k\Pi_k\}$, the vector of efficiency unit costs of production while p is the vector of 'world' prices. Hotelling's Lemma implies that the supply in general equilibrium is given by $X_k = g_{\widetilde{p}_k}/a_k\Pi_k$ while the equilibrium wage is given by g_L. This is the specific gravity model of production.[9]

A very useful closed form solution for g arises when f^k has the Cobb-Douglas form with identical share parameters across the sectors. While this is an extreme simplification, it is consistent with the stability of aggregate labor shares across periods of time when the composition of GDP has altered tremendously. Let $K = \sum_k K_k$ and let α be the parametric share parameter for labor.

[9]In physical science, specific gravity refers to the density of an object normalized by the density of water. In economics, stretching the reference, specific gravity refers to the opportunity cost of a good as its marginal labor requirement normalized by the labor requirement of a constant labor requirement frictionless good, in a setting where the latter is equivalent to opportunity cost.

For the special Cobb-Douglas case,

$$g = L^{\alpha} K^{1-\alpha} G \tag{11}$$

where G is given by

$$G = [\sum_k \lambda_k (p_k/a_k \Pi_k)^{1/(1-\alpha)}]^{1-\alpha}, \tag{12}$$

and $\lambda_k = K_k/K$, the proportionate allocation of specific capital to sector k.[10] GDP is the product of real activity in production and distribution $R = L^{\alpha} K^{1-\alpha}$ and the real activity deflator G. The effect of prices and productivity on the real activity deflator G decompose neatly: G is equal to the GDP price deflator divided by the aggregate technology penalty times the trade cost productivity penalty:

$$G = [\sum_k \lambda_k (p_k/a_k \Pi_k)^{1/(1-\alpha)}]^{1-\alpha} = [\sum_k \lambda_k p_k^{1/(1-\alpha}]^{1-\alpha}/\bar{a}\bar{\Pi}. \tag{13}$$

Notice that (13) implies a solution for the aggregate productivity penalty $\bar{\Pi}$ in terms of the initial Π's, the technology factors $\{1/a_k\}$ and the 'world' prices $\{p_k\}$. For the remainder of this section it is convenient to suppress explicit accounting for the technology factors, so the a's are set equal to one (or equivalently, the Π's can be understood as $\Pi_k a_k, \forall k$).

The special Cobb-Douglas specific factors model yields a constant elasticity of transformation (CET) GDP function. The elasticity of transformation is equal to $\alpha/(1-\alpha)$, the ratio of labor's share to capital's share.[11] The supply share for any good k is given by:

$$s_k = p_k X_k/g = \frac{\lambda_k \widetilde{p}_k^{1/(1-\alpha)}}{\sum_k \lambda_k \widetilde{p}_k^{1/(1-\alpha)}} \tag{14}$$

where $\widetilde{p}_k = p_k/\Pi_k$.

Because the Cobb-Douglas special case will be used extensively to obtain clean results, it is useful to consider how representative are its properties.

[10] Solve the labor market clearance condition for the equilibrium wage, then use the Cobb-Douglas property $wL/\alpha = g$.

[11] The CET form is commonly used in applied general equilibrium modeling. The microfoundations provided here for the CET structure may prove useful in this context, noting that the distribution parameters $\{\lambda_k\}$ are capital allocation shares.

For present purposes, these properties do seem to be representative. Let g_k denote the partial derivative of g with respect to $\widetilde{p}_k$. The specific factors model in general implies $g_{kj} < 0; k \neq j$, $g_{kk} > 0$ and $\sum_j g_{kj}\widetilde{p}_j = 0$. The supply share function (14) represents a wider class that cannot deviate very much from the properties of (14). The wider class of share functions must be homogeneous of degree zero in the prices and retain the derivative properties of g, hence:

$$\frac{\partial s_k}{\partial \widetilde{p}_j}\frac{\widetilde{p}_j}{s_k} = \delta_{kj} - s_j + \frac{g_{kj}\widetilde{p}_j}{g_k}.$$

In the Cobb-Douglas case, $g_{kj}\widetilde{p}_j/g_k = -s_j\alpha/(1-\alpha); k \neq j$ and $g_{kk}\widetilde{p}_k/g_k = (1-s_k)\alpha/(1-\alpha)$. In the general case the same sign properties obtain, as do the same adding up properties.

The characterization of prices and productivity presented in this section is useful in summarizing the implications of trade costs for productivity in partial equilibrium. But since prices and multilateral resistances are endogenous in the world general equilibrium, it is useful to build a general equilibrium 'reduced form' model that relates equilibrium prices and production shares to supply and demand parameters and to trade frictions.

4 World Trade Equilibrium

The specific gravity model yields very strong restrictions on the cross section pattern of production and trade. This section derives a world trade equilibrium 'reduced form' to characterize production and trade patterns, concluding with a world trade equilibrium reduced form structure for productivity.

Description of the expenditure side of the economy is completed with the upper level expenditure allocation that determines the θ's. For concreteness, CES preferences for final goods are assumed. (Any other homothetic form yields qualitatively similar results.) For each country j, its expenditure share on goods of class k is given by

$$\theta_k^j = \gamma_k \left\{ \frac{P_k^j}{P^j} \right\}^{1-\epsilon} \tag{15}$$

where P_k^j is the inward multilateral resistance for country j in goods class k and P^j is the CES index of inward multilateral resistances for country j.

The distribution parameters γ_k and the substitution parameter ϵ are common across countries.

Within each goods class, a CES sub-expenditure function allocates expenditure across trading partners. Within goods class k, the expenditure share on shipments from j delivered to h is given by

$$\Big\{\frac{\beta_{kj}\widetilde{p}_k^j t_k^{jh}}{P_k^h}\Big\}^{1-\sigma_k}.$$

Here, the unit cost of production $\widetilde{p}_k^j$ is augmented by iceberg trade cost factor t_k^{jh} to yield the destination h price of k from j.

Market clearance with balanced trade implies

$$s_k^j g_j - \sum_h \theta_k^h \Big\{\frac{\beta_{kj}\widetilde{p}_k^j t_k^{jh}}{P_k^h}\Big\}^{1-\sigma_k} g^h = 0, \tag{16}$$

$\forall k, j$. This system of equations determines the set of unit costs, $\widetilde{p}_k^j$, one for each k and j. The system is homogeneous of degree zero in the unit costs (understanding that the inward multilateral resistance indexes are homogeneous of degree one in the unit costs, being CES price indexes), hence relative unit costs only are determined.

4.1 Equilibrium Production and Trade Patterns

Using the market clearance equations and the definition of the supply shares, the unit costs are determined as an increasing power function of a shift 'parameter' in demand relative to supply:

$$\widetilde{p}_k^j = \Big\{\frac{D_k^j}{\omega^j \lambda_k^j (\Pi_k^j)^{\sigma_k - 1}}\Big\}^{\frac{1-\alpha}{\alpha+\sigma_k(1-\alpha)}} (G^j)^{1/(\alpha+\sigma_k(1-\alpha))} \tag{17}$$

where

$$D_k^j = \beta_{kj}^{1-\sigma_k}\Theta_k,$$

$$\Theta_k = \sum_h \theta_k^h \omega^h,$$

$$\omega^h = g^h / \sum_h g^h.$$

(17) is derived using (7) in (16). Here the demand shifter D_k^j is the product of a k specific component Θ_k reflecting tastes in the global economy for good k and a (j,k) specific 'quality' parameter $\beta_{kj}^{1-\sigma_k}$ reflecting tastes within goods class k for varieties from origin j. In the Cobb- Douglas case of (15), where $\epsilon \rightarrow 1$, Θ_k is a parameter, hence so is D_k^j.

The incidence of trade costs in varieties from j in class k, Π_k^j, drives the unit cost (suppliers' price) lower for the empirically relevant case $\sigma_k > 1$. Unit costs are increasing in the demand side drivers D_k^j, aggregate demand Θ_k and quality $\beta_{kj}^{1-\sigma_k}$. On the supply side, bigger country size ω^j and bigger sectoral allocations of specific factors λ_k^j both reduce unit costs. Finally, G^j represents the influence of the overall cost push on unit costs coming from all sectors in location j.

The GDP shares are may be expressed as 'reduced form' equations in the international equilibrium using (17). Let $\eta_k = \alpha + \sigma_k(1-\alpha)$. Then:

$$s_k^j = (\lambda_k^j)^{1-1/\eta_k}(\Pi_k^j)^{(1-\sigma_k)/\eta_k}(D_k^j)^{1/\eta_k}(\omega^j)^{-1/\eta_k}G_j^{1-\sigma_k}. \tag{18}$$

The G's can be solved for in terms of the λ's, the Π's and the D's using the adding up condition on the shares, $1 = \sum_k s_k^j$. First, define $R^j \equiv (L^j)^\alpha (K^j)^{1-\alpha}$, a parameter, and note that $\omega^j = R^j G^j / \sum_j R^j G^j$. The adding up condition is

$$1 = \sum_k (\lambda_k^j)^{1-1/\eta_k}(\Pi_k^j)^{(1-\sigma_k)/\eta_k}(D_k^j)^{1/\eta_k}(\omega^j)^{-1/\eta_k}G_j^{1-\sigma_k}, \forall j$$

Each G^j is uniquely solved in terms of the parameters and ω^j by the preceding equation. The solution for G^j can be substituted into the definition $\omega^j = R^j G^j / \sum_j R^j G^j$ to solve for the ω's.

A special case clarifies and is helpful in yielding tighter predictions of the model. Consider the preceding equation in the case where $\sigma_k = \sigma, \forall k$. Define $\Lambda^j \equiv \{\sum_k (\lambda_k^j)^{1-1/\eta}(\Pi_k^j)^{(1-\sigma)/\eta}(D_k^j)^{1/\eta}\}^\eta$. Then

$$G^j = (\Lambda^j/\omega^j)^{1/\eta(\sigma-1)} \tag{19}$$

where the equilibrium world GDP shares are given by

$$\omega^j = \frac{(\Lambda^j)^{\gamma/\eta(\sigma-1)}(R^j)^\gamma}{\sum_j (\Lambda^j)^{\gamma/(\eta(\sigma-1)}(R^j)^\gamma}, \tag{20}$$

where $\gamma \equiv \frac{(\sigma-1)\eta}{1+(1-\sigma)\eta}$.

In the case of $\sigma_k = \sigma$ the reduced form production share equations simplify to

$$s_k^j = \frac{(\lambda_k^j)^{1-1/\eta}(\Pi_k^j)^{(1-\sigma)/\eta}(D_k^j)^{1/\eta}}{\sum_k (\lambda_k^j)^{1-1/\eta}(\Pi_k^j)^{(1-\sigma)/\eta}(D_k^j)^{1/\eta}}. \tag{21}$$

As compared to (18), (21) eliminates the effect of country size on the equilibrium pattern of production. This simplification is not likely to distort the implications much, since the effect of country size tends to cancel out even in the more general case. Based on (21):

Proposition 2 *In the special case of equal elasticities of substitution in expenditure (with $\sigma > 1$) and uniform Cobb-Douglas production functions, the equilibrium production share is*

1. *increasing in the capital allocation share λ_k^j;*
2. *increasing in the demand 'parameter' D_k^j, the product of global market size Θ_k and national quality $\beta_{kj}^{1-\sigma_k}$;*
3. *increasing in the dispersion of D_k^j/λ_k^j and*
4. *decreasing in the incidence of trade costs Π_k^j.*

Proposition 2.3 follows because the deflator in (21) is concave in D/λ for $\eta \geq 1$. The economic intuition is that the mismatch of the sectoral allocation of capital with the pattern of demand lowers GDP, hence the relative size effect in the world economy improves the terms of trade. As a benchmark, consider the case where capital is allocated efficiently across sectors. With efficient allocation of the K's, it is readily shown that $D_k^j(\Pi_k^j)^{1-\sigma}/\lambda_k^j = 1, \forall k, j$.[12] Then better matches of demand and supply shifters act via reductions in dispersion that raise GDP.

If Proposition 1 applies, as is plausible despite non-uniform trade costs as argued above, then (21) implies that capital infusion reaps an externality via the resulting decline in the incidence of trade costs on supply. Larger

[12]The λ's must be chosen to equate the value of marginal product of capital in each sector. Using the GDP function this implies: $g_{\lambda_k}/K = g_K, \forall k$. For the special Cobb-Douglas case this implies that $s_k = \lambda_k$. Then in general equilibrium it must be true that $D_k \Pi_k^{1-\sigma}/\lambda_k = \bar{c}$, a constant. Moreover, by (14) and (21), $\bar{c} = 1$.

This also implies that $\widetilde{p}_k = 1, \forall k$. This property is no surprise in light of the Ricardian property of the specific factors model when the capital/labor ratios are all equal in the long run.

capital allocations λ_k^j gain market share through their direct effect in (21) and through their knock-on effect in lowering Π_k^j.

The reduced form unit cost equations simplify when $\sigma_k = \sigma$ to

$$\widetilde{p}_k^j = (\omega^j)^{-1/(\sigma-1)} \frac{((D_k^j/\lambda_k^j \Pi_k^j)^{\sigma-1})^{(1-\alpha)/\eta}}{\left\{ \sum_k (D_k^j)^{1/\eta} (\Pi_k^j)^{(1-\sigma)/\eta} (\lambda_k^j)^{1-1/\eta} \right\}^{1/(\sigma-1)}} \tag{22}$$

Compared to (17), the special case (22) implies that larger countries have uniformly lower unit production costs.

The implications of (22) for equilibrium 'competitiveness' are very intuitive and sharp:

Proposition 3 *In the special case model, all else equal:*

1. *larger specific endowments lower costs;*
2. *larger world demand for a good raises its cost;*
3. *higher quality costs more;*
4. *higher incidence of trade costs lowers unit costs;*
5. *bigger countries have lower costs.*
6. *higher dispersion of D_k^j/λ_k^j raises unit costs.*

That higher quality costs more is less obvious than it might seem. The CES model of preferences implies that some of each variety will be demanded, so it is not true that lower quality must have a lower price to be purchased by anyone.[13] Proposition 3.3 states that in general equilibrium, higher quality goods have higher unit costs, all else equal.

Proposition 3.6, like Proposition 2.3, reflects the concavity of the deflator in (21) and (22) in D/λ.

The model implies very strong restrictions on the equilibrium pattern of trade. The ratio of gross exports to GDP in the special case of equal elasticities of substitution is given by

$$s_k^j - \theta_k^j \left\{ \frac{\beta_{kj} t_k^{jj} \widetilde{p}_k^j}{P_k^j} \right\}^{1-\sigma}. \tag{23}$$

[13] The interpretation of $\beta_{kj}^{1-\sigma_k}$ as a quality parameter is natural from examining the sub-utility function that lies behind the CES expenditure function: starting from equal consumption of each variety, the consumer's willingness to pay is higher the larger is $\beta_{kj}^{1-\sigma_k}$.

Using (22) and the definition of θ_k^j, (15), and imposing $\epsilon = \sigma$ this reduces to

$$s_k^j - (s_k^j)^{(\eta-1)/\eta} \frac{\Delta_k^{jj}}{(D_k^j(\Pi_k^j)^{1-\sigma})^{1-1/\eta^2}} \tag{24}$$

where $\Delta_k^{jj} \equiv \gamma_k(\beta_{kj} t_k^{jj}/P^j)^{1-\sigma}$ and s_k^j is given by (21). The implications are that:

Proposition 4 *in the special case model the ratio of gross exports to GDP is*

1. *increasing in* s_k^j*, which moves according to Proposition 2;*
2. *increasing in* D_k^j*,*
3. *decreasing in the incidence of trade costs* Π_k^j*, and*
4. *decreasing in* Δ_k^{jj}*.*

Each item in the proposition is intuitive.

The levels of trade follow from scaling up the GDP shares by national GDP's. The implications of specific gravity model for the cross country pattern of aggregate production and wages are very strong.

For any pair of countries j and h, the ratio of their GDP's is given by

$$\frac{g^j}{g^h} = \frac{R^j}{R^h} \frac{G^j}{G^h}, \tag{25}$$

where $R^j \equiv (L^j)^{\alpha}(K^j)^{1-\alpha}, \forall j$. The influence of prices and productivity on the relative GDP's comes through the relative G's. These incorporate both the terms of trade and the relative productivity of the two economies due to trade costs.

In the Cobb-Douglas special case, the reduced form (19) applied to the G's given in (12) implies that the ratio of G's is given by:

$$\frac{G^j}{G^h} = \left\{ \frac{R^j}{R^h} \right\}^{-1/(1+\eta(\sigma-1))} \left\{ \frac{\sum_k \lambda_k^j (D_k^j/\lambda_k^j (\Pi_k^j)^{\sigma-1})^{1/\eta}}{\sum_k \lambda_k^h (D_k^h/\lambda_k^h (\Pi_k^h)^{\sigma-1})^{1/\eta}} \right\}^{1/(\sigma-1)(1+\eta(\sigma-1))}. \tag{26}$$

The first term on the right in (26) is a relative size effect, implying in light of (25) that bigger countries in real terms have lower wages and GDP price deflators, but not so much as to lower relative nominal GDP. (That is, size is not immiserizing.) The second term on the right in (26) is a composition effect reflecting the match of sector specific factor allocations to the pattern of demand in the global economy.

Intuition about the composition effect is provided by considering as a benchmark the efficient allocation of the K's. For that case, $\Lambda^j = 1$ in (19), hence $G^j = (\omega^j)^{-1/(\eta(\sigma-1)}$. Then for inefficient allocations, $\Lambda^j = \sum_k \lambda_k^j (D_k^j \Pi_k^{1-\sigma}/\lambda_k^j)^{1/\eta}$ measures how far short of efficient allocation of its specific factors economy j is. Relative GDP's are increasing in this inefficiency due to the relative size effect operating on terms of trade in the world economy.

The model has very strong implications for relative wages. The relative wage (using $w = g_L$) is given by

$$\frac{w^j}{w^h} = \left\{\frac{K^j/L^j}{K^h/L^h}\right\}^{1-\alpha} \frac{G^j}{G^h}$$

Using (26), relative wages are obviously increasing in the relative capital/labor ratio and decreasing in the relative inefficiency of the match of sectoral allocations to demand. One key influence on the latter is the average outward multilateral resistance: the relative wage is lower the higher is the relative outward multilateral resistance.

4.2 Productivity and Trade Frictions in General Equilibrium

The aggregate effect of trade frictions on productivity at given equilibrium prices is given by $\bar{\Pi} = g(p, \cdot)/g(\widetilde{p}, \cdot)$. The equilibrium prices in turn depend on depend on many exogenous variables of the world economy. Under the special assumptions and using the methods of this paper, the effect of exogenous variables on productivity can be conveniently and intuitively aggregated.

An important benchmark is the special case model when the specific factors are efficiently allocated. The identical Cobb-Douglas production function structure assumed here makes the production set effectively Ricardian in the long run, when capital allocation adjusts. In general equilibrium, due to the love of variety structure of preferences, prices adjust to permit diversification despite the Ricardian production set.

In general equilibrium, the GDP's are just functions of the real activity indexes $(L^j)^\alpha (K^j)^{1-\alpha}$. Compared to the pure Ricardian economy, differences in productivity are smoothed out by sectoral reallocations of capital.[14] In effect, supply is infinitely elastic for this case (the supply schedule in Figure 1 is horizontal) so all the incidence of distribution friction (or productivity differences) falls on consumers. All the Π's are equal to 1 in this full general equilibrium. The long run general equilibrium production shares reduce to

$$s_k^j = \lambda_k^j = D_k^j(\Pi_k^j)^{1-\sigma} = D_k^j.$$

Supply always adjusts to meet demand in this special case model. Also, globalization has *no* long run effect on productivity or the pattern of production; all gains are passed on to consumers in the form of consumer price index declines (as iceberg costs fall). The size of the gains in real income from globalization depend on the pattern of trade cost declines as they affect each consuming location, measured by inward multilateral resistance.

This general equilibrium reduced form Ricardian model is in polar opposition to the Ricardian solution of Eaton and Kortum (2002) that features supply side forces only. The mechanisms of the two models are quite different. In the Eaton-Kortum model, goods are homogeneous across countries (σ_k is very large in terms of the present model), leading to only one supplier in each destination for each good. The productivities are random, drawn from distributions that differ internationally by a 'location' parameter. The proportion of goods that each country will export in equilibrium is based on forces located on the supply side: trade costs, wages and the technology parameters. Demand side forces disappear into a constant term that cancels in trade shares. In contrast, the present model forces diversified production by assuming that goods are differentiated by place of origin (σ_k is finite). The long run equilibrium pattern of production is dictated by demand side forces only.

The long run equilibrium special case is analyzed only to develop intuition. More reasonable models in contrast imply shared incidence of trade frictions due to finite elasticity of supply in general equilibrium. For example, the special case in demand where $\sigma_k = \sigma$ along with Cobb-Douglas production, sector specific capital and identical labor shares yields the real activity deflator in equilibrium as (19):

$$G^j = (\Lambda^j/\omega^j)^{1/\eta(\sigma-1)}$$

[14] With $\Lambda^j = 1$, the activity deflators G are given by $G^j = (\omega^j)^{-1/(\eta(\sigma-1)}$. Also, $\omega^j = (R^j)^\gamma / \sum_j (R^j)^\gamma$. GDP is given by $g^j = (R^j)(\omega^j)^{-1/\eta(\sigma-1)}$.

where $\Lambda^j = \{\sum_k (\lambda_k^j)^{1-1/\eta}(\Pi_k^j)^{(1-\sigma)/\eta}(D_k^j)^{1/\eta}\}^\eta$. The effect of the incidence of trade frictions on G can be calculated and decomposed by sector with this formula (for the given equilibrium GDP shares and expenditure patterns Θ_k). The overall productivity effect is captured with the uniform Π that is equivalent to the actual set of $\{\Pi_k^j\}$'s (solved from $\Lambda(\Pi) = \Lambda(\bar{\Pi})$):

$$\bar{\Pi}^j = \Big\{ \sum_k (\Pi_k^j)^{-(\sigma-1)/\eta} \frac{\widetilde{\lambda}_k^j}{\sum_k \widetilde{\lambda}_k^j} \Big\}^{-\eta/(\sigma-1)} \tag{27}$$

where $\widetilde{\lambda}_k^j = (\lambda_k^j)^{1-1/\eta}\beta_{kj}^{(1-\sigma)/\eta}\Theta_k^{1/\eta}$ and D_k^j is replaced by$\beta_{kj}^{1-\sigma}\Theta_k$. (27) provides an explanation for cross section differences in productivity in world trading equilibrium that incorporates the effect of multilateral resistance on world prices.

The benchmark case in which the Π's are all equal to one occurs because the λ's are endogenous, driven by demand shares such that the incidence of trade costs falls entirely on the demand side. A rough intuition for Π exceeding one is that 'capital' immobility induces supply shares that prevent all incidence falling on the demand side; supply elasticities are finite in the general equilibrium. The greater the mismatch between specific factor allocations and the 'equilibrium' allocations, the greater the incidence on the supply side of the economy, loosely speaking.

5 Intermediate Inputs and Selection

Intermediate products trade comprises a large and growing share of world trade. Vertical disintegration is apparent — integrated production broken apart, with components produced in one location and assembled in another. Presumably the gains from vertical disintegration have a powerful impact on productivity. A simple extension of the specific factors model of production readily encompasses intermediate products trade.

The boundary between components still produced within one location and those produced elsewhere and traded is central to the phenomena to be explained. In the multi-country context, essentially the same boundary phenomenon arises because only a small portion of the potential bilateral trade links have any positive trade flows. Countries able to fill more of the links in intermediate products trade presumably reap a productivity benefit. The treatment of selection in this section also applies to the final goods trade of preceding sections, with a gain to consumers from variety when more links open up.

The data show that larger markets are served by more suppliers. The natural explanation is that fixed costs impose a barrier that selects only those markets large enough to be profitable to serve. Helpman, Melitz and Rubinstein (2007) treat selection in a gravity model of final goods trade. Their model is adapted here to intermediate products trade, and the gravity model is embedded in the specific factors model of production.

Firms are assumed to be competitive in factor markets (the common labor market and the sector specific factor markets, one for each sector) and monopolistic competitors in product markets. The marginal firms earn zero profits in production and distribution to each destination's market. Due to the CES specification, all firms mark up their unit costs by a common factor, inducing a distribution of prices that mirrors the distribution of productivities of extant firms.

The analysis is static for simplicity, so the fixed costs of entering production and trade are treated as sunk. Different countries will have different total factor productivities, due to their differing entry decisions in production, while the fixed costs of trade further alter total factor productivities through their effects on multilateral resistance.

The simplification to a static model shuts down an important general equilibrium factor market linkage between entry decisions and current production costs. See Melitz (2003) for a proper dynamic treatment in a one factor production model. The pragmatic reason for treating the static case here is to avoid the complexity of simultaneously treating entry and the reallocation of the specific factors.

5.1 Specific Factors Production with Intermediates

The production function for each industry k is comprised of the production functions of those firms that earn non- negative profits. At a prior stage, firms choose to enter production and then receive a Hicks-neutral productivity draw from a probability distribution. Those firms unlucky enough to receive draws too low to allow breaking even exit from production. The average productivity in industry k, $1/\bar{a}_k$, is determined by the cutoff productivity of the marginal firm in combination with the parameters of the productivity draw distribution. Average productivity is for present purposes taken as given. Since average productivity is Hicks-neutral, it enters the specific factors general equilibrium model multiplicatively with multilateral resistance.

The average productivity is associated with an average price, a constant markup over the the average unit cost of extant firms. See Melitz (2003) for details. This setup allows treatment of the heterogeneous firm model as a representative firm model easily linked to the general equilibrium production theory of preceding section. Profits are earned by inframarginal firms, and form part of the rents earned by the sector specific factors.

Intermediate products enter for simplicity as just a single intermediate product, potentially produced as a variety at each location.[15] The CES aggregate of the varieties is an input into production of all final goods and the intermediate good at each location. To ease notation, suppress country indexes. The production function for product k is given by

$$X_k = f^k(L_k, K_k, M_k), k = 1, ..., m;$$

where M_k is the quantity of the CES aggregate intermediate input used in sector k and sector m is the intermediate goods production sector. Specializing to the Cobb-Douglas case,

$$f^k = L_k^{\alpha} K_k^{1-\alpha-\nu} M_k^{\nu} / \bar{a}_k \Pi_k$$

Let P_m denote the price of the intermediate input used by the home country, a CES aggregate of the intermediate products purchased from all trading origins while p_m denotes the price of the intermediate input produced at home. Cost minimization combines with the labor market clearance condition to yield the GDP function $g(\widetilde{p}, P_m, L, K, \{\lambda_k\})$ with a closed form in the special case given by

$$L^{\alpha/(1-\nu)} K^{1-\alpha/(1-\nu)} [(\sum_{k=1}^{m} \lambda_k \widetilde{p}_k^{1/(1-\alpha-\nu)})^{1-\alpha-\nu} P_m^{-\nu}]^{1/(1-\nu)} c. \tag{28}$$

Here, c is a constant term combining the parameters, while $\widetilde{p}_k = p_k / \bar{a}_k \Pi_k$, the 'efficiency unit cost' in sector k.

5.2 Selection to Trade

Now allow for a fixed cost of exporting to each market, a cost that must be borne by each firm active in county i in goods class k. Helpman, Melitz

[15]The methods used here readily scale up to any number of intermediates.

and Rubinstein (2007) derive the gravity model with selection and show that there are two consequences that contrast with the preceding model: first, some bilateral trade flows may be zero due to no firm being able to pay the cost of entry, and second, where bilateral trade is positive it reflects both substitution on the intensive margin as in the preceding model and substitution on the extensive margin due to selection into exporting by marginal firms. The exposition below reviews their model, and reformulates it to highlight the role of multilateral resistance in both intensive and extensive margins.

The firm is assumed to be a monopolistic competitor facing a continuum of other firms selling to a market characterized by Dixit-Stiglitz love of variety preferences (and the analogous setup for intermediate products) represented by a CES expenditure function in each goods class. The mass N_k^i of firms that enter into production is given. Firms that enter receive a productivity draw from a Pareto distribution. In any market worth serving, the profit maximizing firms price to market with a constant markup over their costs $\mu_k = \sigma_k/(\sigma_k - 1)$.

The cost of a firm to serve its own market (assuming that $t_{ii}^k = 1$ for simplicity) is given by $\widetilde{p}_k^i$ times a_k^i, the inverse of the firm's productivity draw. Denote aggregate expenditure on product class k at destination j by E_k^j and use the CES expenditure system to allocate expenditure across origins. Sales by i to country $j \neq i$ are profitable only if $a_k^i \leq a_k^{ij}$ where a_k^{ij} is defined by the zero profit condition:

$$(1 - 1/\mu_k)(\frac{t_{ij}\widetilde{p}_k^i a_k^{ij}}{\mu_k P_k^j})^{1-\sigma_k} E_k^j = f_k^{ij}$$

Here, f_k^{ij} denotes the fixed cost.

It eases notational clutter in what follows to temporarily suppress the separate accounting for each goods class k, and to move the location indexes to the subscript position. Define the selection variable $V_{ij}(a_{ij})$ where

$$V_{ij} = \int_{a_L}^{a_{ij}} a^{1-\sigma_k} dF(a)$$

for $a_{ij} \geq a_L$ while

$$V_{ij} = 0$$

otherwise. Here, F is the cumulative density function. Denote the expenditure in location j on the generic good shipped from all origins as E_j while the value of shipments to all destinations from location i is denoted Y_i.

Now derive the gravity model. For simplicity, the quality adjuster β is uniformly equal to one. Then the bilateral import value of shipments is given by

$$X_{ij} = (\frac{\widetilde{p}_i t_{ij}}{P_j/\mu})^{1-\sigma} E_j N_i V_{ij}.$$

The total value of shipments is

$$Y_i = \sum_j X_{ij} = \widetilde{p}_i^{1-\sigma} N_i \sum_j (\frac{t_{ij}}{P_j/\mu})^{1-\sigma} V_{ij} E_j.$$

First, solve market clearance for $\widetilde{p}_i^{1-\sigma}$:

$$\widetilde{p}_i^{1-\sigma} = \frac{y_i/Y}{\Pi_i^{1-\sigma}}. \tag{29}$$

Here, y_i denotes the shipments of the average firm in country i, Y_i/N_i and $Y = \sum_i Y_i = \sum_j E_j$, while

$$\Pi_i^{1-\sigma} \equiv \sum_j (\frac{t_{ij}}{P_j/\mu})^{1-\sigma} V_{ij} E_j / Y \tag{30}$$

Substitution yields the bilateral flows as:

$$X_{ij} = (\frac{t_{ij}}{P_j \Pi_i/\mu})^{1-\sigma} V_{ij} Y_i E_j / Y,$$

where

$$P_j^{1-\sigma} = \sum_i (\frac{t_{ij}}{\Pi_i/\mu})^{1-\sigma} V_{ij} Y_i / Y. \tag{31}$$

The normalization condition for the Π's follows from manipulating ({SupplyPriceInt) and summing:

$$\sum_i N_i (\Pi_i \widetilde{p}_i)^{1-\sigma} = 1. \tag{32}$$

The selection equation can be restated to highlight the role of multilateral resistance. Selection is controlled by:

$$(1 - 1/\mu)(\frac{a_{ij} t_{ij}}{\mu P_j \Pi_i})^{1-\sigma} E_j y_i / Y = f_{ij}. \tag{33}$$

There are three implications. First, notice that the gravity model with selection combines the effects of trade costs on the intensive margin with their effects on the extensive margin acting through V_{ij}. Higher fixed costs reduce volume while larger markets draw more entrants. Second, μ plays a role in selection. Incorporating variation across goods class, higher markup (lower elasticity) goods classes will have more firms selected into exporting, all else equal. Third, most importantly, the multilateral resistance variables incorporate both the productivity penalty imposed by the incidence of trade costs and the productivity gain garnered by the incidence of selection into trade.

The formal model is completed by specifying a distribution function for G. With the Pareto distribution used by Helpman, Melitz and Rubinstein, let the Pareto parameter be κ. Then

$$V_{ij} = \frac{\kappa a_L^{\kappa-\sigma+1}}{(\kappa-\sigma+1)(a_H - a_L)} W_{ij}$$

$$W_{ij} = max[(a_{ij}/a_L)^{\kappa-\sigma+1} - 1, 0].$$

Helpman, Melitz and Rubinstein estimate selection with a Probit regression, then use these estimates to control for selection in the second stage gravity model regression with positive trade flows. Identification is achieved with an exclusion restriction that readers may find unconvincing (common religion affects fixed costs but not variable costs). The proposed research aims at more convincing exclusion restrictions by thinking of fixed costs as sunk (hence for example exchange rate variability and expropriation risk will affect selection but not variable cost) and by exploiting commodity class characteristics.

5.3 Selection, Productivity and Trade Patterns

The solution for the multilateral resistance terms in (30)-(31) subject to the normalization in (32) determines the productivity and comparative advantage implications of the incidence of trade costs in conditional general equilibrium.

The special case Cobb-Douglas model yields the GDP function (28). Higher incidence of trade costs in intermediate inputs penalizes GDP more heavily. Due to the separability of the GDP function, the reduced form production shares are independent of the incidence of trade costs on intermediate inputs P_m. This separability implies that all the production and trade pattern results of Section 4 apply.

The influence of selection on production and trade patterns is isolated in the outward multilateral resistance terms, the incidence of trade costs on productivity, while the effect of selection on aggregate productivity also enters through the inward multilateral resistance for intermediates. No analogue to Proposition 1 is available, even for the case of uniform border barriers, essentially because selection introduces destination specific asymmetries in the weights attaching to trade costs from any origin.

Selection is endogenous, with (33) permitting a characterization conditional on the expenditure and production shares. More firms will be selected from i to trade with j the larger is the market in j, the larger is i's share of world shipments, and the fewer (hence larger) are i's firms. Thus selection reinforces the productivity implications of trade costs: big market share shippers bear lower incidence of trade costs. In contrast, selection tends to offset the higher incidence induced by larger expenditure shares.

6 Conclusion

This paper provides a platform for consistent aggregation of the fine structure of trade costs into productivity measures that are suitable for the analysis of productivity differences across goods, countries and time. The implications of the productivity differences at a point in time for the pattern of production and trade are explored in detail for the special case of the specific factors model.

The paper points to future empirical work. First, it will be valuable to estimate multilateral resistance indexes for an appropriately disaggregated set of goods for a set countries and years. Second, the paper points to use of the multilateral resistance indexes as an explanatory variable for the pattern of production and trade.

The paper also points to future theoretical refinement. The extreme simplicity of the model buys strong results, while hinting that the results hold in less restrictive cases. How robust is the model?

Finally, the analysis reveals important channels through which technology shocks in production and in distribution in one country are transmitted to productivity in all trading partners. The specific factors structure suggests gradual adjustment to long run equilibrium. Future research might profitably explore these channels for their implications about inference of productivity and about the international transmission of shocks.

7 References

Anderson, James E. (1979), "A Theoretical Foundation for the Gravity Equation," *American Economic Review*, 69, 106-16.

Anderson, James E. and Eric van Wincoop (2003), "Gravity with Gravitas: A Solution to the Border Puzzle", *American Economic Review*, 93, 170-92.

Anderson, James E. and Eric van Wincoop (2004), "Trade Costs", *Journal of Economic Literature*, 42, 691-751.

Davis, Donald and David Weinstein (2001), "An Account of Global Factor Trade", *American Economic Review*, 91, 1423-53.

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

Helpman, Elhanan, Marc J. Melitz and Yonah Rubinstein (2007), "Trading Partners and Trading Volumes", NBER Working Paper No. 12927.

Melitz, Marc J. (2003), "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity", *Econometrica*, 71(6), 1695-1725.

8 Appendix: Multilateral Resistance and Size

Proof of Proposition 1

Proposition 1 states that with uniform border barriers, multilateral resistance is decreasing in the supply share of the countries, increasing in the expenditure share of countries and increasing in the net imports of countries.

The assumption of the Proposition imposes a *uniform* international trade cost t on all trades across borders, will internal trade (from j to j) assumed to be frictionless. A slightly different notational convention is used here than in the text. The analysis takes a representative good, so the subscript k is suppressed. It then eases notation slightly to move the location indexes from the superscript to the subscript position. Let s_j denote country j's share of world shipments (at delivered prices) of the generic good, while b_i denotes the expenditure share of country i on the generic good.

The system of equations that determine P_i, Π_i for all i, j is given by the simple case version of (6)-(7), which reduces under the changes in notation and the implications of the uniform international trade cost to:

$$P_j^{1-\sigma} = t^{1-\sigma}\bar{h} + (1 - t^{1-\sigma})s_j/\Pi_j^{1-\sigma} \tag{34}$$

$$\Pi_i^{1-\sigma} = t^{1-\sigma}\bar{h}' + (1 - t^{1-\sigma})b_i/P_i^{1-\sigma}. \tag{35}$$

Here, $\bar{h} = \sum_i s_i \Pi_i^{\sigma-1}$ and $\bar{h}' = \sum_j P_j^{\sigma-1} b_j$. Recognizing that $\bar{h} = \sum_j (\beta_j \widetilde{p}_j)^{1-\sigma}$, it is convenient to impose the normalization on the producer prices $\bar{h} = \bar{h}'$. $\bar{h}$ can be solved for eventually but the essential step is to provide the closed form solution for multilateral resistance given $\bar{h}$.

Multiply both sides of (34) by $\Pi_i^{1-\sigma}$ and multiply both sides of (35) by $P_i^{1-\sigma}$. Use the resulting equality to solve

$$\Pi_i^{1-\sigma} = P_i^{1-\sigma} + \frac{(1 - t^{1-\sigma})(b_i - s_i)}{\bar{h}t^{1-\sigma}}.$$

Then substitute into (34) and extract the positive root[16] of the resulting quadratic equation in the transform $P_i^{1-\sigma}$:

$$2P_i^{1-\sigma} = \gamma_i + [\gamma_i^2 + 4(1 - t^{1-\sigma})b_i]^{1/2} \tag{36}$$

where

$$\gamma_i = \bar{h}t^{1-\sigma} - \frac{(1 - t^{1-\sigma})(b_i - s_i)}{\bar{h}t^{1-\sigma}}.$$

[16] The positive root of the quadratic is necessary for P to be positive.

At this solution

$$2\Pi_i^{1-\sigma} = \bar{h}t^{1-\sigma} + [\gamma_i^2 + 4(1 - t^{1-\sigma})b_i]^{1/2}.$$

Multilateral resistance (inward and outward) is unambiguously decreasing in supply share s_i at equilibrium and unambiguously increasing in expenditure share b_i in equilibrium. It is unambiguously increasing in the net import share $b_i - s_i$ for given expenditure shares. $\|$

The solution for $\bar{h}$ is implicit in the next expression, obtained from using the definition of $\bar{h}$ and the preceding solution for Π_i,

$$\bar{h} = \sum_i s_i[\bar{h}t^{1-\sigma} + (\gamma_i^2 + 4(1 - t^{1-\sigma})b_i)^{1/2}]^{-1},$$

where γ_i is given as a function of $\bar{h}$ above.