From Structural Gravity to Welfare

Marc-Andreas Muendler*
UC San Diego, CESifo and NBER

October 27, 2020

Structural Gravity

Structural Gravity of international trade can be stated as

$$X_{sd} = \frac{Y_s \, X_d}{X} \, \frac{1}{\tau_{sd}^{\theta}} \, \frac{1}{\Pi_s^{-\theta}} \, \frac{1}{P_d^{-\theta}}, \tag{1}$$

where X_{sd} is the bilateral trade flow from source country s to destination country d (James E. Anderson and Van Wincoop, Eric 2003). $Y_s = \sum_d X_{sd}$ is production (output) and country s's total income. $X_d = \sum_s X_{sd}$ is destination market size and country d's expenditure.[1] $X = \sum_s Y_s = \sum_d X_d$ is global income and expenditure. Similar to the gravity equation in physics, bilateral trade (the gravity force) X_{sd} is stronger the larger the source country's production Y_s and the larger the destination country's market X_d, adjusted for a gravity constant (reflecting the size of the global economy $1/X$).

Similar to gravity in physics, bilateral trade flows are inversely related to the countries' distance. The parameter $\tau_{sd} \geq 1$ measures relevant economic components of distance, including the physical distance between countries, their human distance in terms of language and legal barriers, their economic distance in terms of transport costs, and their economic policy distance in terms of political choices such as tariffs and non-tariff barriers. A concise way to think of τ_{sd} is to consider it the trade cost (such as a freight factor for transportation) for shipments from source s to destination d with $\tau_{ss} = 1$ for internal (domestic) trade and $\tau_{sd} > 1$ for international (foreign) trade. The coefficient $-\theta$ is the elasticity of bilateral trade with respect to trade cost—the *trade elasticity* for short.

The combined term τ_{sd}^{θ} is also called the bilateral resistance that trade flows have to overcome. The remaining two terms $\Pi_s^{-\theta}$ and $P_d^{-\theta}$ are then called the multilateral resistances (Anderson and Van Wincoop, Eric 2003) and serve as correction terms for bilateral resistance in the structural gravity equation. The multilateral resistances reflect the fact that, in general economic equilibrium, optimal trade flows depend on both the source country's remoteness and the destination country's remoteness, mitigating or aggravating the importance of the bilateral resistance for a country pair.

Concretely,

$$P_d^{-\theta} = CPI_d^{-\theta} = \frac{1}{X} \sum_s \frac{Y_s}{\Pi_s^{-\theta}} \tau_{sd}^{-\theta}$$

is the *inward multilateral resistance* of destination d and proportional to the country's properly defined Consumer Price Index (CPI). For a constant elasticity of substitution σ in consumer demand, and $\theta = \sigma - 1$, the

*University of California–San Diego, La Jolla, CA; *muendler@ucsd.edu*. Further affiliations: CAGE, ifo and IGC.

[1]Expenditure at destination d does not need to equal income of destination d. There can be trade imbalances: if $X_d < Y_d$ country d runs a trade surplus by spending less than its income, if $X_d > Y_d$ country d runs a trade deficit by spending more than its income. There is no extraterrestrial trade, so in the world aggregate $\sum_s Y_s = \sum_d X_d = X$.

properly defined CPI (the minimum expenditure on one unit of the consumption basket) is

$$CPI_d = \left[\sum_s \alpha_s (p_{sd})^{-\theta} \right]^{-1/\theta}, \tag{2}$$

where the price of a good from s at destination d is $p_{sd} = \tau_{sd} p_{ss}$. The inward multilateral resistance measures the destination's average remoteness from any source country in the world, weighted by the economic importance of the source countries $Y_s/\Pi_s^{-\theta}$. Conversely,

$$\Pi_s^{-\theta} \equiv \frac{1}{X} \sum_d \frac{X_d}{P_d^{-\theta}} \tau_{sd}^{-\theta}$$

is the *outward multilateral resistance* of source country s and measures the source's average remoteness from any destination in the world, weighted by the economic importance of the destinations $X_d/P_d^{-\theta}$. Take Antarctica as an example of a source of trade. It is a remote source, so its outward resistance $\Pi_s^{-\theta}$ is a small number (its $\tau_{s\cdot}$ are large). Any bilateral trade flow, such as the flow to the destination Denmark d say, will be larger than the bilateral distance alone τ_{sd} would predict on average because Antarctica is a remote source for any destination; hence $\Pi_s^{-\theta}$ corrects the bilateral trade prediction under τ_{sd} upwards. Similarly, consider Antarctica as a destination. Its inward resistance $P_d^{-\theta}$ is low (its consumer price index is high as purchasing goods from afar is expensive). So, any bilateral trade flow, such as the flow from the source country Sweden s say, will be larger than the bilateral distance alone τ_{sd} would predict on average; hence $P_d^{-\theta}$ corrects the bilateral trade prediction under τ_{sd} upwards.

Foundations of Structural Gravity

Many trade models can provide the theoretical foundations for structural gravity. One example of an important demand-side framework is the Paul S. Armington (1969) model, where source countries s are endowed with unique goods. Consumers anywhere in the world have an insatiable love for variety and will reduce their expenditure on all other countries' unique good to make room in their expenditure for the consumption of any given country's good. The Paul R. Krugman (1980) model considers a production side in addition to the demand feature that consumers have an inexhaustible love variety and introduces firms, all symmetric look-alikes of each other when it comes to production. However, each firm with its product variety occupies a unique niche in the global market and consumers will reduce their expenditure on all other firms' unique goods to make room in their expenditure for the consumption of any given firm's good. Given the unique sourcing of any variety from only one firm in the world, it may not be surprising that the Krugman (1980) model delivers the same gravity relationship as the Armington (1969) model, just that now a country's good is a bundle of all its firms' varieties. The resulting trade elasticity from the Armington (1969) and Krugman (1980) models is a function of the elasticity of substitution between goods: $-(\sigma - 1)$, where σ is the constant elasticity of substitution in consumer demand.

Instead of making firms symmetric look-alikes of each other in production, the Marc J. Melitz (2003) model allows firms to differ by their innate productivity. The demand side with consumers' inexhaustible love for variety remains the same and, similar to the Krugman (1980) model, the Melitz (2003) model implies unique sourcing: any variety comes from one unique firm in the world. The structural gravity equation results again. When specifying the distribution of the heterogeneous firms' productivities with a Pareto distribution that has a shape parameter θ (Elhanan Helpman, Marc J. Melitz and Stephen R. Yeaple 2004; Thomas Chaney 2008), the trade elasticity is no longer governed by the demand side but instead equals the negative of the Pareto shape parameter $-\theta$. Jonathan Eaton and Samuel Kortum (2002) establish a model with many countries and many industries, generalizing the early insight by David Ricardo (1817) for two industries and two countries that a global division of labor according to the principle of comparative advantage generates gains from trade for all participating nations. Similar to the original Ricardian example, complete specialization results so that a destination country's consumption of a given good is purchased from a unique country, implying unique sourcing like in the preceding three models—now at the level of

Table 1: **Theoretical Foundations of Structural Gravity**

Model	Exporter capability A_s	Importer effect X_d/Φ_d	Bilateral effect	Trade elasticity
Armington (1969)	$\alpha_s w_s^{-\theta}$	$X_d/P_d^{-\theta} = X_d/\left[\sum_k \alpha_s(\tau_{sd}p_{kk})^{-(\sigma-1)}\right]$	$\tau_{sd}^{-(\sigma-1)}$	$-(\sigma-1)$
Krugman (1980)	$N_s w_s^{-\theta}$	$X_d/P_d^{-\theta} = X_d/\left[\sum_k N_s(\tau_{sd}p_{kk})^{-(\sigma-1)}\right]$	$\tau_{sd}^{-(\sigma-1)}$	$-(\sigma-1)$
Melitz (2003)	$N_s w_s^{-\theta}$	$X_d/\Phi_d = X_d/\left[\sum_k N_s(\tau_{sd}p_{kk})^{-\theta}\right]$	$\tau_{sd}^{-\theta}$	$-\theta$
Eaton and Kortum (2002)	$T_s w_s^{-\theta}$	$X_d/\Phi_d = X_d/\left[\sum_k T_s(\tau_{sd}p_{kk})^{-\theta}\right]$	$\tau_{sd}^{-\theta}$	$-\theta$

Note: Source country's wage (per-capita income) is w_s. In Armington (1969), α_s is the source's global consumer appeal, and $w_s = p_{ss}$ follows for unitary unit labor requirements in production so output equals employment ($q_s = L_s$). In Krugman (1980) and Melitz (2003), N_s is the source's equilibrium number (measure) of firms (varieties). In Melitz (2003), θ is the Pareto shape parameter of the firm productivity distribution. In Eaton and Kortum (2002), T_s is the Fréchet location parameter and θ the Fréchet shape parameter of the country-industry productivity distribution.

industries. Eaton and Kortum (2002) specify the global distribution of country-industry productivities with a Fréchet distribution, and the trade elasticity now equals the negative of the Fréchet shape parameter $-\theta$.

All these models, and several more (Keith Head and Thierry Mayer 2014), imply a Structural Gravity equation of the form

$$X_{sd} = A_s \frac{X_d}{\Phi_d} \frac{1}{\tau_{sd}^{\theta}}, \tag{3}$$

where $A_s = Y_s/\Pi_s^{-\theta}$ and $\Phi_d \equiv \sum_k A_k \tau_{kd}^{-\theta} = P_d^{-\theta}$ so (3) is equivalent to (1). Stated differently, all these models imply that a source country s's share in a destination d's expenditure X_d is a fraction

$$\frac{X_{sd}}{X_d} = \frac{A_s \tau_{sd}^{-\theta}}{\sum_k A_k \tau_{kd}^{-\theta}}, \tag{4}$$

and Structural Gravity results as a consequence. Sufficient assumptions for a relationship like (4) to hold include (i) unique sourcing and (ii) a love-for-variety demand system.[2]

The term A_s can be thought of as the source country's export capability behind its bilateral trade flows. It is a function of the country's production cost (reflected in its wage), which in turn is a function of all the prices that its products command around the globe. As a consequence A_s reflects both the source country's export capability and its multilateral outward resistance. Table 1 provides a synopsis of alternative foundations for Structural Gravity.

Gains from Trade

To measure the welfare gains from trade for the country as a whole, consider real per-capita income (the average household's real income) at a destination country d:

$$W_d = \frac{w_d}{P_d},$$

where W_d is welfare, w_d is the nominal wage, and P_d is the well-defined (ideal) CPI consistent with the demand system. To simplify derivations, we can take labor at destination d as the *numéraire* in the sense that all other labor incomes around the world w_s ($s \neq d$) and all product prices p_{sd}, including p_{ss} at home, are relative to w_d. Then we can set the destination's nominal per-capita income to $w_d = 1$ for convenience and with no loss of generality.

For the derivation, we return to the Armington (1969) model. Similar derivations apply to the other

[2]For a discussion of necessary assumptions, see Head and Mayer (2014) and Arnaud Costinot and Andrés Rodríguez-Clare (2014).

models in Table 1.[3]

We are interested in the total percentage change (proportional change) in destination d's welfare when trade costs τ_{sd} around the world change:

$$\mathrm{d}\ln W_d = \mathrm{d}\ln w_d - \mathrm{d}\ln P_d = -\mathrm{d}\ln P_d, \tag{5}$$

where the second step follows because $\mathrm{d}\ln w_d = 0$ for the *numéraire*. The operator d stands for the total differential (the total incremental change).

Turn to the price index and its changes. Consider a change in all product prices around the globe p_{sd} and take the total differential of the log of the CPI in (2) to study the impact of changing prices around the globe on consumers at destination d:[4]

$$\mathrm{d}\ln P_d = \sum_s \frac{X_{sd}}{X_d}\,\mathrm{d}\ln p_{sd}. \tag{6}$$

Mirroring the derivation in Costas Arkolakis, Arnaud Costinot and Andrés Rodríguez-Clare (2012, Appendix A, step 3), it is useful for the derivation, and intuitive, to consider the change in foreign trade flows relative to the change in self trade (within-country sourcing). After all, it is the global trade flows that change when global trade costs change, and the comparison to self trade provides a normalizaiton. Take the total derivative of the log of (4) for $s \neq d$ and substract the total derivative of the log of (4) for $s = d$:

$$\mathrm{d}\ln\frac{X_{sd}}{X_d} - \mathrm{d}\ln\frac{X_{dd}}{X_d} = \mathrm{d}\ln\frac{\alpha_s\,(\tau_{sd}p_{ss})^{-\theta}}{\sum_k A_k\tau_{kd}^{-\theta}} - \mathrm{d}\ln\frac{\alpha_d\,(\tau_{dd}w_d)^{-\theta}}{\sum_k A_k\tau_{kd}^{-\theta}} = -\theta\,\mathrm{d}\ln p_{sd}$$

because $p_{sd} = \tau_{sd}\,p_{ss}$, w_d is the *numéraire* and constant, and $\tau_{dd} = 1$ is constant. Equivalently,

$$\mathrm{d}\ln p_{sd} = -\frac{1}{\theta}\left(\mathrm{d}\ln\frac{X_{sd}}{X_d} - \mathrm{d}\ln\frac{X_{dd}}{X_d}\right). \tag{7}$$

Using (7) in (6), the total change in the consumer price index can be restated as

$$\mathrm{d}\ln P_d = -\frac{1}{\theta}\sum_s\frac{X_{sd}}{X_d}\left(\mathrm{d}\ln\frac{X_{sd}}{X_d} - \mathrm{d}\ln\frac{X_{dd}}{X_d}\right) = \frac{1}{\theta}\sum_s\frac{X_{sd}}{X_d}\,\mathrm{d}\ln\frac{X_{dd}}{X_d} = \frac{1}{\theta}\,\mathrm{d}\ln\frac{X_{dd}}{X_d}\sum_s\frac{X_{sd}}{X_d} = \frac{1}{\theta}\,\mathrm{d}\ln\frac{X_{dd}}{X_d}. \tag{8}$$

The simplifications follow because the trade shares must sum to one, so $\sum_s X_{sd}/X_d = 1$, and therefore the relative changes in trade shares must sum to zero, $\sum_s (X_{sd}/X_d)\,\mathrm{d}\ln(X_{sd}/X_d) = 0$. Returning to welfare change in (5) and using (8), we find

$$\mathrm{d}\ln W_d = -\mathrm{d}\ln P_d = -\frac{1}{\theta}\,\mathrm{d}\ln\frac{X_{dd}}{X_d}. \tag{9}$$

This result, first established by Arkolakis, Costinot and Rodríguez-Clare (2012) for a family of models and demand systems, is often stated for the *self-trade share* (the share of expenditure on domestic goods),

$$\Lambda_d \equiv \frac{X_{dd}}{X_d},$$

and in terms of percentage changes. Define the percentage change (proportional change) of a variable with

[3] In the Krugman (1980) and Melitz (2003) models, the equilibrium number (measure) of firms N_s is endogenous and changes in response to changes in fundamental parameters. The derivations nevertheless extend to the case of endogenous N_s.

[4] The total differential of (2) is

$$-(\sigma-1)\,\mathrm{d}\ln P_d = \mathrm{d}\ln\sum_s\alpha_s(p_{sd})^{-(\sigma-1)} = \sum_s\frac{\alpha_s}{\sum_k\alpha_k(p_{kd})^{-(\sigma-1)}}\,\mathrm{d}(p_{sd})^{-(\sigma-1)} = -(\sigma-1)\sum_s\frac{\alpha_s\,\mathrm{d}(p_{sd})^{-(\sigma-1)}}{\sum_k\alpha_k(p_{kd})^{-(\sigma-1)}}\,\frac{\mathrm{d}p_{sd}}{p_{sd}}.$$

a hat on top of the variable: $\hat{x} \equiv \Delta x/x$. Then, integrating (9) over all incremental changes, the welfare consequences of large changes in global trade costs are given by the formula

$$\hat{W}_d = \left(\hat{\Lambda}_d\right)^{-\frac{1}{\theta}}. \tag{10}$$

The welfare formula establishes that, for any change in global trade costs, the change in country-wide welfare can be inferred from only two statistics: the trade elasticity $-\theta$ and the change in the share of expenditure on domestic goods $\hat{\Lambda}$.[5] This is a backward looking (ex post) result. If we know the realized change in self trade $\hat{\Lambda}_d$ we can completely infer the welfare consequence. The converse is not true for forward looking (ex ante) analysis. If we do not know the change in $\hat{\Lambda}_d$ but instead have to predict it, trade cost reductions can have different implications for $\hat{\Lambda}_d$ in different trade models, and the predicted welfare change will therefore depend on the theory.

In autarky, $\Lambda_d = 1$. The current level of Λ_d can therefore be used to compute the gains from trade relative to the counterfactual of a completely closed economy.

References

Anderson, James E., and Van Wincoop, Eric. 2003. "Gravity with Gravitas: A Solution to the Border Puzzle." *American Economic Review*, 93(1): 170–92.

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

Armington, Paul S. 1969. "A Theory of Demand for Products Distinguished by Place of Production." *International Monetary Fund Staff Papers*, 16(1): 159–178.

Chaney, Thomas. 2008. "Distorted Gravity: The Intensive and Extensive Margins of International Trade." *American Economic Review*, 98(4): 1707–21.

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

Eaton, Jonathan, and Samuel Kortum. 2002. "Technology, Geography, and Trade." *Econometrica*, 70(5): 1741–79.

Head, Keith, and Thierry Mayer. 2014. "Gravity Equations: Workhorse, Toolkit, and Cookbook." In *Handbook of International Economics*. Vol. 4, , ed. Elhanan Helpman, Kenneth Rogoff and Gita Gopinath, Chapter 3, 131–195. Amsterdam: Elsevier.

Helpman, Elhanan, Marc J. Melitz, and Stephen R. Yeaple. 2004. "Export versus FDI with Heterogeneous Firms." *American Economic Review*, 94(1): 300–316.

Krugman, Paul R. 1980. "Scale Economies, Product Differentiation, and the Pattern of Trade." *American Economic Review*, 70(5): 950–59.

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

Ricardo, David. 1817. *On the Principles of Political Economy and Taxation.* London: John Murray.

[5] The result is based on changes in the domestic CPI, an instance of Shepard's Lemma, and the fact that domestic prices are unaffected by the shock. Formula (10) therefore also applies to the case of a shock to foreign preferences or endowments beyond the case of changes in global trade costs.