Global Journal of Arts, Humanities and Social Sciences, 14(1),1-23, 2026
Print ISSN: 2052-6350(Print)
Online ISSN: 2052-6369(Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

Theoretical Foundations and Evolution of the Structural Gravity Model in International Trade: Policy Insights and Implications

Clinton Uzobor
Research Candidate at University of Gdansk, Poland
clintuzo@yahoo.com

doi: https://doi.org/10.37745/gjahss.2013/vol14n1123 Published January 11, 2026

Citation: Uzobor C. (2026) Theoretical Foundations and Evolution of the Structural Gravity Model in International Trade: Policy Insights and Implications, *Global Journal of Arts, Humanities and Social Sciences*, 14(1),1-23

Abstract: *This paper revisits the theoretical evolution and modern refinements of the gravity model of international trade, emphasizing its transformation from an empirical regularity into a structurally grounded analytical framework. It situates the gravity model within the broader landscape of trade theory, highlighting the limitations of traditional Ricardian and Heckscher–Ohlin models in explaining observed trade patterns. The paper traces the model's development from early Newtonian analogies to its microeconomic foundations, as formalized by Anderson, Bergstrand, and Deardorff, and to subsequent extensions by Eaton and Kortum and Melitz that incorporate technology differences, market structure, and firm heterogeneity. Central attention is given to the contribution of Anderson and van Wincoop, who resolved the border puzzle by introducing multilateral resistance terms that capture general equilibrium trade cost effects. The refined structural gravity model is shown to provide a coherent explanation of bilateral trade flows by accounting for both direct and indirect trade barriers. Overall, the paper concludes that structural gravity represents a unifying framework for empirical trade analysis and policy evaluation.*

Keywords: Gravity model, international trade theory, trade costs, multilateral resistance, structural gravity, globalization.

INTRODUCTION

Late 1970s marks the period when the differences between international trade theory and international trade facts were becoming rather more visible. Although the theoretical phase relying on comparative advantage was highly praised and intellectually appealing but was very much in disagreement with the observed patterns of trade. According to Deardoff (1984), there are many who thinks that the Ricardian and Heckscher-Ohlin theories do not provide a complete explanation of world trade. Several other authors have also alleged that empirical consistency in the international trade data cannot be explained in terms of these dominant traditional theories.

There are three important empirical regularities that seem to stand strongly in support of the later development in trade theories, especially the monopolistic competition model and other models such as the type proposed by Anderson (1979), Deardorff (1998), Eaton and Kortum (2002) and those explaining the importance of heterogenous firms. The first evidence in support of the new theories points to its flexibility in interpreting the large volume of intra-industry trade with respect to inter-industry trade (see Leamer & Levinsohn, 1995). The second evidence relates to the excellent empirical performance of the gravity equation for both intra-industry and inter-industry trade patterns. Finally, the third evidence adduce to the rising volume of trade among countries with similar technology and factor endowment relative to the amount of trade between countries with different levels of technology and factor inputs, which runs contrary to the predictions of the traditional theories.

This paper re-visits the contemporary refinement in theoretical gravity, as an empirically consistent framework in the analysis of international trade. The paper is structured as follows. The next section discusses the initial development on gravity model. The third section discusses prelude to the development of theoretical gravity. The fourth section presents the contemporary refinement to the theoretical framework. The fifth, sixth, and seventh sections presents the theoretical synthesis/implications, the policy insight and future directions and policy extensions of the model. The final section concludes.

Initial development on gravity model

The gravity model has received an increasing popularity over the years and has since become the workhorse of the applied international trade literature. It owes its origin to Newtonian physics, which explains the law of universal gravitation, and dates back to 1687. According to the law, the gravitational forces between two entities, are directly proportional to the product of their masses, and indirectly related to distance between them.

The equation can be represented as:

$$F_{ij} = \mathrm{G}\frac{M_i M_j}{d_{ij}^2} \qquad (1)$$

where F_{ij} is the force of gravity, defined as a direct function of the product of masses (M) of i *and* j entities, and indirect function of their squared differences (d) apart. The gravitational constant (G) is the autonomous function, and the value of which is statistically or econometrically determined.

Carey (1865) applied the gravity model in the study of migration, making it the very first time it was used in the field of social science. Isard (1960) also contributed to the regional science study through the application of the concept of gravity. The applicability of the gravity model in the analysis of international economics is the theoretical success that puts the law of universal gravity into an economic context. In the analysis of trade flows, the relationship gave rise to the simplified version of the gravity model defined as:

$$X_{ij} = K\frac{Y_i Y_j}{D_{ij}} \qquad (2)$$

The model explains trade flows (X_{ij}) between country i and j as an increasing function of their respective economic size (Y_i and Y_j), represented by each country's GDP and a decreasing function of distance (D_{ij}), represented by the geographical distance between the country's capital cities, as well as the gravitational constant captured by the K (Head and Mayer 2014). In the recent years, the gravity model has been used to analyse the impact of different flows such as equity, migration, and foreign direct investment (Grogger and Hanson, 2011; Head and Ries, 2008; Portes and Rey, 2005).

In terms of the standard gravity model of trade, the first application came to the fore in 1962 through the work of Tinbergen, marking it the first time the law of universal gravity was treated in a proper economic context. In the author's view, the trade flow equation can be defined by the following relationship:

$$X_{ij} = \mathrm{A}\frac{Y_i^{\alpha} Y_j^{\beta}}{D_{ij}^{\rho}} \qquad (3)$$

Where α is the exporter's GDP elasticity, β is the importer's GDP elasticity, and ρ represent the elasticity of distance. The elasticities α, β and ρ can take any different values than 1 and reveals the nature and how much influence each of the variable can exert on trade flows. This model is the same with the Newtonian equation of the universal gravitation if the elasticities of the GDP of partners are constant (i.e., $\alpha, \beta = 1$) and the distance between partners is considered to and from (i.e., $\rho = 2$).

In Tinbergen (1962), the trade flow between two entities is estimated by the gravity equation given as:

$$In(X_{ijt}) = A + \alpha In(Y_{it}) + \beta In(Y_{jt}) - \rho In(D_{ij}) + e_{ijt} \qquad (4)$$

The apriori expectation is that the exporter's GDP (Y_i) and the importer's GDP (Y_j) will have positive effects on trade flows by α and β per cent everything else held constant; distance between the exporter and importer is expected to negatively influence trade flows by ρ per cent ceteris paribus; e_{ijt} is the error term - usually assumed to be independent and log-normally distributed. Intuitively, if Y_i and Y_j increases by 1 unit, the impact on trade flow is by α and β amounts. In the same vein, if D_{ij} increases by 1 unit, it reduces trade flows by ρ, indicating an inverse relationship between trade and geographical distance.

In general, the model was developed on the ground that the amount of exports country i can supply depends on its size of the economy, and the amount of goods country j can demand from country i depends on its expenditure capability given the level of income. The economic size is often proxied by the GDP, gross national product, income per capita or the country's population size, which underscores both the production capacity and market potentials of each country. Precisely, the economic size represents the supply and demand condition of each country that determines trade flows. Distance is a trade barrier with similar effects as the function of a tax, which are linked to transportation cost. The influence of distance on trade can take different forms including transportation costs, cost of demurrage, synchronization costs, communication costs, transaction costs or costs resulting from cultural differences (Head, 2003). Because of the complexity in measuring these factors, geographical distance between capital cities (measured as the shortest distance over the earth's surface) is often used to approximate these costs. However, trade can also be encouraged by some elements of cultural heritage and similarity in political system. In the recent times, many studies have modelled different factors affecting trade. For

Global Journal of Arts, Humanities and Social Sciences, 14(1),1-23, 2026
Print ISSN: 2052-6350(Print)
Online ISSN: 2052-6369(Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

example, current discussion in trade literature have tended to ascertain whether countries in the same currency union, facing similar borders or enjoying some form of regional trade agreements have advantages in trading (see, Rose and van Wincoop, 2001; McCallum, 1995; Rose, 2004).

Evolution of the Theoretical Gravity Framework

Figure 1 illustrates the gravity model as a unifying framework that integrates classical and modern trade theories within a single general equilibrium structure. Rather than competing with existing theories, the gravity model operationalizes their core mechanisms through relative trade costs and multilateral resistance, yielding a common structural foundation for empirical trade analysis.

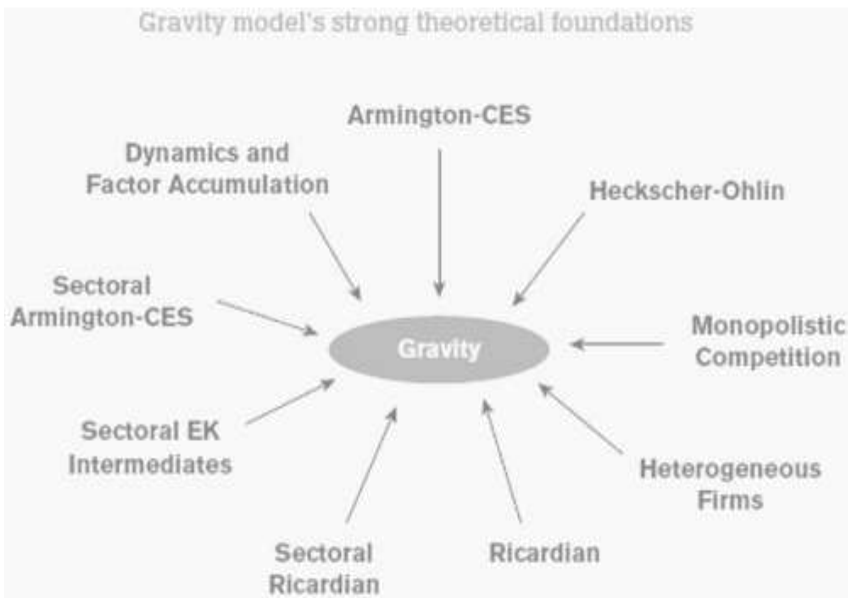

Figure 1: The Gravity Model as a Synthesis of Trade Theories

The first and most profound development in the theoretical gravity model of international trade is attributed to Anderson (1979). One of the key anchors of the model is the Armington's assumption[i], which argued that goods are nationally differentiated - the goods can either be traded intra-nationally or/and inter-nationally. The base assumption was that countries have identical Cobb-Douglas preferences[ii], which implies that income spent on tradable goods is the same for both exporting and importing countries, with prices considered to be fixed at equilibrium value. In a world where distance is not regarded and no shipment cost or trade barriers, export supplies from the origin countries are conditioned by the multiple value of income attributed to the exporting country i and importing country j. The standard or traditional gravity equation can be represented as:

$$X_{ij} = \frac{Y_i Y_j}{\sum_j Y_j} \qquad (5)$$

However, the assumption of a world without border and by implication trade relations without barriers is not well-grounded in economic theory and cannot be said to be very helpful when examining the holistic determinants of trade. Thus, with gradual modification to incorporate friction factors, Anderson extended his model to include transportation costs, which defines the value of export shipments of goods produced in country *i* to country j. This gives the final aggregate gravity of trade equation proposed by Anderson as:

$$T_{ij} = \underbrace{\frac{\varnothing_i Y_i \, \varnothing_j Y_j}{\sum_j \varnothing_j Y_j} \frac{1}{f(D_{ij})}}_{\text{distance from i to j in relation to world trade expenditure}} \underbrace{\left[\sum_j \frac{\varnothing_j Y_j}{\sum_j \varnothing_j Y_j} \frac{1}{f(D_{ij})}\right]^{-1}}_{\text{Distance from i to all other potential trading partners in relation to world trade}} \qquad (6)$$

The first part of the equation outside the bracket represents the economic distance from country *i* to country *j* in relation to aggregate world expenditure on trade. The second part in the bracket denotes the economic distance from country i to all other potential trading partners in relation to total world trade. Thus, Anderson (1979) argued that bilateral trade flows between countries depends on their income relative to a trade-weighted average of the economic distance.

Later development by Helpman and Krugman (1985) was based on Ricardian comparative advantage and factor endowments of Heckscher-Ohlin. At the firm level, the approach assumes increasing returns to scale and a condition based on monopolistic competition between firms[iii]. The trade explanation of Krugman (1979, 1980) was based on the concept of monopolistic competition as proposed by Chamberlin[iv]. The theory assumed similarity in endowment and technology and further argued that the most important factor of production is labour. He argued that because of similarity in production method (technology), wage rate in the two trading countries is equal and as a result, prices for all products will be the same. Production cost also falls subsequently due to increasing production (attributed to expanding trade-induced labour force) and the concept of economies of scale. Krugman (1979, 1980) argued that preferences among countries were identical – a proposition similar to Lindner (1961), who observed that nations with similar characteristics engage in trade of products belonging to the same class of industries. The vintage of the argument is the inherent concept of 'love of variety'[v] through which welfare of participating countries rises out of expanded product choice.

The model by Bergstrand (1985, 1989) is the representative gravity model that captures the monopolistic competition-based analysis, with differentiated products and economies of scale. Bergstrand's gravity equation combines both factor endowments, relating to Heckscher-Ohlin models and the CES preferences, reflecting Lindner's propositions of identical countries having similar preferences.

The micro-based model is represented as:

$$PX_{ij} = \varkappa_0 Y_i^{\varkappa_1} \left(\frac{K_i}{L_i}\right)^{\varkappa_2} Y_j^{\varkappa_3} \left(\frac{K_j}{L_j}\right)^{\varkappa_4} D_{ij}^{\varkappa_5} Z_{ij}^{\varkappa_6} \cup_{ij} \qquad (7a)$$

Bergstrand extends the micro-based model to incorporate the factor proportion theory of trade - by including factor endowment variable. This reflects the Heckscher-Ohlin inter-industry trade models, and CES preferences in the spirit of Lindner (1961)'s assumption of national similarity in preferences. By including further assumption of a competitively determined market price faced by the profit maximizing firm of good A, the final and extended gravity equation from Bergstrand (1989) yields:

$$PX_{ij} = \varkappa_0 Y_i^{\varkappa_1} \left(\frac{K_i}{L_i}\right)^{\varkappa_2} Y_j^{\varkappa_3} \left(\frac{K_j}{L_j}\right)^{\varkappa_4} C_{ij}^{\varkappa_5} \mathcal{T}_{ij}^{\varkappa_6} E_{ij}^{\varkappa_7} P_i^{\varkappa_8} P_j^{\varkappa_9} \cup_{ij} \qquad (7b)$$

In the model, trade flow PX_{ij} from country i to destination j is dependent on the respective income (Y_I and Y_j) of the trading partners, the exporter's capital-labor ratio $\left(\frac{K_i}{L_i}\right)$, the importers' per capita income $\left(\frac{K_j}{L_j}\right)$, C_{ij} represents the c.i.f./f.o.b. geographical distance factor, $\mathcal{T}_{ij}$ is the imposed tariff rate, E_{ij} stands for the exchange rate - denominated in the exporter's currency, and P_i *and* P_j captures the exporter's and importer's price indexes, and by convention a log-normally distributed stochastic term is included.

Bergstrand assumes that measure of the GDP in units of capital can serve as a proxy for the income Y_i of the exporter and thus gives information about the relative factor endowment (capital-labour ratio) of each country, which resonates with the inter-industry trade theory of Heckscher-Ohlin. Changes in the level of expenditure Y_j of the importer (often proxied by the GDP) is seen to be a function of changes in the preference defined by taste, and points to the intra-industry trade model of Helpman and Krugman.

Contrary to the flaws presented by Helpman and Krugman that the application of the Heckscher-Ohlin approach to the gravity model is theoretically unfounded, Deardorff (1998) provided a refute by arguing that two scenarios in trade pursuit exists: one that is frictionless and another with impediment. In the case of trade without friction, there are no shipment cost and other barriers to trade. Thus, due to product homogeneity in a perfect competition with zero shipment cost for trading partners, the origin of the product and efforts to make available the products in any required location does not face hindrances. In the discern of Deardorff, a world pool of goods is the sum of individual producers' goods. Individual consumers explore their preferences by choosing out of the pool of available goods. The equilibrating price for each good clears the world market pool of goods, because of perfect competition[vi]. The very tenet of Deardorff's frictionless model is that income must equal expenditure (defined as the multiplication of price and quantity of good). Hence the frictionless gravity model.

$$Y_i = \hat{p}x_i = \hat{p}c_i$$

The vector x_i and c_i are the production and consumption capabilities of country i $\hat{p}$ is the world price. Given the assumption that preferences are identical and homothetic for all countries, it is inferred that they spend the same amount of income, β_k, on good k. The demand of Country j of good k to meet its consumption needs becomes:

$$C_{jk} = \beta_k \frac{Y_j}{p_k}$$

And out of the world basket of good k from all manufacturing countries A, the contribution x of country i to the pool is represented by:

$$\varphi_{ik} = \frac{x_{ik}}{\sum_A x_{Ak}}$$

Thus, the actual consumption-meeting demand of country j from country i represents

$$C_{ijk} = \varphi_{ik} C_{jk} = \frac{x_{ik}}{\sum_A x_{Ak}} \left[\beta_k \frac{Y_j}{p_k} \right] \qquad (8)$$

For simplicity, let assume that the aggregated (world) output of good k, is given by X_i and with recourse to the identical proportion of income expended by respective countries on good k, the fraction of world expenditure on good k must equal the aggregate income of the world Y_w. Then, the value of imports of country j from country i can be given as

$$X_{ij} = \frac{Y_i Y_j}{Y_w} \qquad (9)$$

There are no geographical barriers. Distance barrier is not considered in the model, which signifies zero transportation cost. This equation is the simple frictionless gravity model, which is the same as equation (5)

In the case of impeded trade, Deardorff assumes a complete specialization with trade barriers (such as transportation cost) existing for every good and are considered strictly positive on all country pairs (i, j) bilateral trade. According to Deardorff, it is consistent with Armington preference and any monopolistic competition (the bases for which Bergstrand (1989) used "iceberg" form of transport costs). If there is a single price for all market, the pattern of bilateral trade will be such that countries specialize in the production of good for which they have relative factor cost advantage and then engage in trade. Transport cost relates to distance and tends to diminish trade. Though with the Cobb-Douglas preference, the bilateral trade flourish with distance. However, Deardorff considered preference underscored by CES utility function to arrive at the Heckscher-Ohlin world:

$$T_{ij}^{f.o.b} = \frac{Y_i Y_j}{t_{ij} Y_w} \left[\frac{\left(\frac{t_{ij}}{p_j^I}\right)^{1-\sigma}}{\sum_A \theta_A \left(\frac{t_{ij}}{p_j^I}\right)^{1-\sigma}} \right] = \frac{Y_i Y_j}{t_{ij} Y_w} \left[\frac{p_{ij}^{1-\sigma}}{\sum_A \theta_A p_{iA}^{1-\sigma}} \right] \qquad (10)$$ [vii]

Where θ_A is the fraction of total world income attributed to country A, and σ denotes the elasticity of substitution. The term $\frac{p_{ij}^{1-\sigma}}{\sum_A \theta_A p_{iA}^{1-\sigma}}$ is the distance factor (the multilateral resistance term), that defines the relative distance from the destination country j (the importer) to the origin country i (the exporter), and therewith a representation of the average of all destinations (importing countries') relative distance to the origin (exporter i). 1+ transport cost defining the transport factor. According to Deardorff (1998), If the relative distance between origin country i and destination country j is smaller (greater) than the average, trade between the two countries will be more (less) than trade in the standard gravity equation.

Eaton-Kortum (2002) develop Ricardian model of bilateral trade (with continuum of goods) to motivate an approach that captures the tension between comparative advantage (promoting trade) and geographic barriers (impeding it), through the instrumentality of differences in production technology. They combined the nature of technology in each country (reflecting the absolute advantage), the differences in technology (reflecting the comparative advantage) and the geographic barriers (reflecting the iceberg transportation cost due to distance)[viii] to estimate the elasticity of trade. According to Eaton-Kortum (2002), the position of the distribution of productivities is determined by technological capability of countries, such that the amount of goods country i sales to country j, measured by country j's expenditure on goods from country i, underscores the bilateral representation of the trade flows expressed as:

$$X_{ij} = \frac{\left(\overline{A}_j d_{ij}\right)^{-\theta} X_j}{\sum_M \left(\overline{A}_m d_{im}\right)^{-\theta} X_m} \mathbb{Q}_i = \frac{X_j \left(\frac{d_{ij}}{p_j}\right)^{-\theta}}{\sum_M \left(\frac{d_{im}}{p_m}\right)^{-\theta} X_m} \mathbb{Q}_i \qquad (11)$$

Where, X_j total spending of country j on purchase of goods from country i, X_{ij} is the fraction of goods that country j buys from country i, $\mathbb{Q}_i$ is the exporter's total quantity of goods sold (which reflects country i's income Y_i), d_{ij} is the geographic barrier (the bilateral resistance term) between the exporter and the importer and is deflated by the importer's price level p_j (the multilateral resistance term) – both of which are raised to the power of the variability factor $-\theta$, and the denominator $\left(\frac{d_{im}}{p_m}\right)^{-\theta} X_m$ represents the size of individual destination market m as perceived by the origin country i and $\sum_M X_m$ can be perceived as the world income Y_w. Analogously, they considered the share of country i in the total market of country j in relation to its home market share to be expressed by the equation:

$$\frac{X_{ij}/X_j}{X_{ii}/X_i} = \frac{\Phi_i}{\Phi_j} d_{ij}^{-\theta} = \left(\frac{p_i d_{ij}}{p_j}\right)^{-\theta} \qquad (12)$$

According to Eaton-Kortum (2002), the higher the variability factor (that is lower $-\theta$), the greater the likelihood for comparative advantage to exerts a stronger force on some goods to escape the geographical barriers. Alternatively, as the force of comparative advantages weakens (that is higher θ), the shares of normalized import become more elastic to the average relative price and to geographic barriers.

Melitz (2003) insisted on the importance of export behaviour of heterogeneous firm. He combined the differentiated firm-level productivity and fixed cost of shipment to develop a model that captures the major characteristics that distinguishes firms. Chaney (2008) and Redding (2011) modelled the productivity distributions of the firm, which according to the authors are pareto in nature. Thus, the Melitz model is useful in a situation where the productivity distribution can be characterised by the pareto distribution. The model assumes consumer's preferences to be characterised by the Dixit-Stiglitz CES-utility function that underscores the product varieties. Following the assumption of Free entry by firms, the model also yields a structural gravity model similar to those derived in the previous sections.

$$X_{ij} = \left(\frac{t_{ij}}{\Pi_i P_j}\right)^{-k} \left(f_{ij}{}^X\right)^{-\left(\frac{k}{\sigma-1}-1\right)} Y_i Y_j \qquad (13)$$

The first term in the bracket is function of fixed trade cost ($f_{ij}{}^X$) - conditional on the cut-off productivity, which captures the size of the firm, and can be thought of as the extensive margins following Redding (2011).Π_i *and* P_j are the price terms referred to as the multilateral resistance terms of Anderson and van Wincoop. The multilateral resistance helps to capture the relative trade resistance. The relative trade resistance defines the specific bilateral trade between country i and j in relation to possibilities of trade engagement with other potential trade partners.

Contemporary refinement to the theoretical gravity

The most profound refinement to the theoretical foundation of the gravity came as a response to McCallum's (1995) border effect publication on trade between Canadian provinces and the United States of America (U.S.) states. In McCallum (1995), the trade flow between two regions (either inter-provincial trade or province-state trade) was estimated through a gravity equation that failed to consider multiple equilibria in international trade, leading to omitted variable bias. The findings of McCallum resulted in a phenomenon often described as the border puzzle, and to which Obstfeld and Rogoff (2001) refer to as home bias in trade.

Motivated by the resulting border puzzle, Anderson and van Wincoop (2003, 2004) expanded on the Armington-CES model of Anderson (1979) to develop a method which allows estimating gravity equations more consistently by incorporating the general equilibrium effects of trade costs. The authors argued that the gravity estimation based on prior presentation of McCallum was bias, due to omitted variables and lack of comparative statics for the trade frictions. Their main aim was to identify the prevailing trade costs, because according to the authors, differences in prices observed by countries are due to trade costs, which are directly unobservable.

To capture the multiple and gravitating equilibria in international trade and solve the McCallum border puzzle, Anderson and van Wincoop included the multilateral resistance term in the gravity equation. Characteristically, these are the structural terms that emerged to form part of the trade costs. The costs correspond to the value of traded goods - usually defined by the amount of trade. The barriers can take the form of an ad-valorem tax equivalent, or an iceberg trade costs (Samuelson, 1954; Krugman, 1980; Anderson & van Wincoop, 2003). Anderson and van Wincoop assumed that trade costs are c.i.f., which implies that the cost burden is shouldered by the exporter, who finally shifts the burden to the importer.

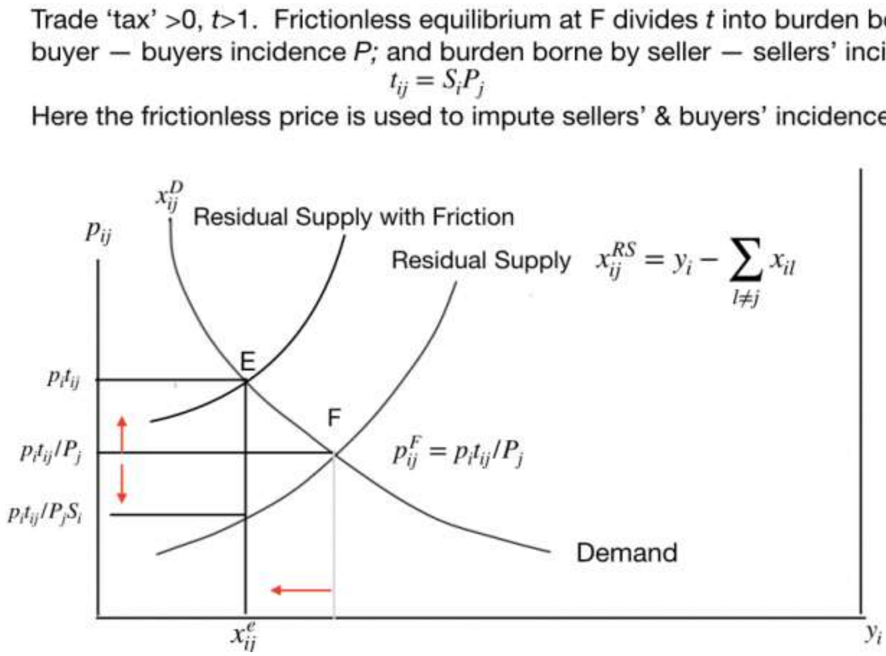

Figure 2: Trade Cost Incidence and Price Formation under Trade Frictions

Trade costs are modeled as iceberg or ad-valorem equivalents that create a wedge between producer and consumer prices. These wedges determine the incidence of trade barriers and enter the gravity system through **relative** price indices and multilateral resistance terms.

The model of Anderson and van Wincoop is built on three main assumptions: first, is the assumption that goods are differentiated by their place of origin, in accordance with Armington preferences, individual countries specialize in the production of one good. Second, is the assumption of trade separability, indicating that agents maximize their utility by consuming tradable goods, which are allocated and separately analysed across countries. Third is that countries possess identical and homothetic preferences, implying that countries maintain similar demand structure and given their income, spend the same amount on the good from the exporting country.

Thus, they suggested that the demand-side preferences lead to utility function that exhibits a constant elasticity of substitution (CES), with the trade elasticity being $\sigma > 1$, over all the national products, such that the importer country j region (the consumer) maximize utility given as:

$$U_j = \left(\sum_i (A_i Q_{ij})^{\frac{(\sigma-1)}{\sigma}} \right)^{\frac{\sigma}{(\sigma-1)}} \qquad (14)$$

where A_i represent a utility distribution parameter that can be thought of as an index of the quality of country i's product (technology) (Head and Mayer, 2015). For the sake of simplicity, the parameter A_i can be neglected since its suppression does not in any way change the interpretation of the resulting gravity equation. Hence, the new CES utility function now takes the form:

$$\cup_j = \left(\sum_i (Q_{ij})^{\frac{(\sigma-1)}{\sigma}}\right)^{\frac{\sigma}{(\sigma-1)}} \quad (15)$$

Subject to the budget constrain:

$$\sum_i P_i Q_{ij} t_{ij} = y_j \quad (16)$$

The budget constraint represents total expenditure of country j for the purchase of goods from country i. In the equation, P_i is the selling price (not including the transportation or other costs) charged by the suppliers in country i. While t_{ij} (said to be $t_{ij} > 1$) is the trade cost factor between origin i and destination j. Thus, the total expenditure of country j on all the goods imported from other countries i is represented by the multiple of $P_i Q_{ij} t_{ij}$ - as the total cost, which is equal to the total income y_j earned in country j.

Total expenditure (total costs) is equal to value of the goods sold. The costs-value correspondence defines the barriers, which can take the form of an ad-valorem tax equivalent or iceberg trade costs. Note that in equation (16), what would have been P_{ij} was substituted by $P_i t_{ij}$, since $P_{ij} = P_i t_{ij}$. Therefore, P_i is the price that country j the importer pays to country i the producer or exporter. While t_{ij} is the trade cost factor, such that $t_{ij} - 1$ defines the tax equivalent. Anderson and van Wincoop assumed that trade costs are *c.i.f.*, which implies that the cost burden is shouldered by the exporter, who finally shifts the burden to the importer. Thus, the nominal import of country j from country i can be summarised as:

$$X_{ij} = P_{ij} Q_{ij} \quad (17)$$

The market-clearing condition for the exporter, becomes the total income in country i, which is made up of all exports from country i to all other importing countries j, represented as:

$$Y_i = \sum_j X_{ij} \quad (18)$$

Thus, it has been established that the structural gravity relies on two important conditions. The first relates to spatial allocation of expenditure for the importer. The second relates to the imposition of market-clearing for the exporter. Therefore, the next step is substituting the aggregate demand in the market clearing condition in order to aggregate trade flows and estimate equilibrium trade.

Now, maximizing the CES utility function subject to the budget constraint results in the maximization problem as:

$$\text{Langrange (L)} = \left(\sum_i (Q_{ij})^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{(\sigma-1)}} + \lambda \left(y_j - \sum_i P_i Q_{ij} t_{ij}\right) \quad (19)$$

The aim is to maximize utility (i.e., consumption Q_{ij}), and following the first order condition with respect to Q_{ij} requires that $\frac{\partial L}{\partial Q_{ij}} = 0$, which gives:

$$\frac{\partial L}{\partial Q_{ij}} = \frac{\sigma}{(\sigma-1)} \left(\sum_i (Q_{ij})^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{(\sigma-1)}-1} \frac{\sigma-1}{\sigma} (Q_{ij})^{\frac{\sigma-1}{\sigma}-1} - \lambda P_i t_{ij} = 0$$

Further expression and transpositions give:

$$\lambda P_i t_{ij} = \left(\sum_i (Q_{ij})^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{(\sigma-1)} - \frac{\sigma-1}{\sigma-1}} (Q_{ij})^{\frac{\sigma-1}{\sigma}-1}$$

$$\lambda P_i t_{ij} = \left(\sum_i (Q_{ij})^{\frac{\sigma-1}{\sigma}}\right)^{\frac{1}{(\sigma-1)}} Q_{ij}^{\frac{-1}{\sigma}} \tag{20}$$

Expressing further by multiplying through with Q_{ij} , and then summing over all *i* will yield:

$$\lambda P_i C_{ij} t_{ij} = \left(\sum_i (Q_{ij})^{\frac{\sigma-1}{\sigma}}\right)^{\frac{1}{(\sigma-1)}} Q_{ij}^{\frac{-1}{\sigma}+\frac{\sigma}{\sigma}}$$

$$\lambda \sum_i P_i Q_{ij} t_{ij} = \left(\sum_i (Q_{ij})^{\frac{\sigma-1}{\sigma}}\right)^{\frac{1}{(\sigma-1)}} \sum_i Q_{ij}^{\frac{\sigma-1}{\sigma}} \tag{21}$$

Replacing λ with $\left(\sum_i (Q_{ij})^{\frac{\sigma-1}{\sigma}}\right)^{\frac{1}{(\sigma-1)}} Q_{ij}^{\frac{-1}{\sigma}} (P_i t_{ij})^{-1}$ from equation (20) and substituting $\sum_i P_i Q_{ij} t_{ij}$ in equation (21) with Y_j. It is considered here that the market clearing condition (18) can be transferred on the total income of country j, to obtain:

$$\left(\sum_i (Q_{ij})^{\frac{\sigma-1}{\sigma}}\right)^{\frac{1}{(\sigma-1)}} Q_{ij}^{\frac{-1}{\sigma}} (P_i t_{ij})^{-1} Y_j = \left(\sum_i (Q_{ij})^{\frac{\sigma-1}{\sigma}}\right)^{\frac{1}{(\sigma-1)}} \sum_i Q_{ij}^{\frac{\sigma-1}{\sigma}}$$

If the first term in both sides of the equation cancels out, it gives:

$$Q_{ij}^{\frac{-1}{\sigma}} (P_i t_{ij})^{-1} Y_j = \sum_i Q_{ij}^{\frac{\sigma-1}{\sigma}}$$

Solving for $P_{ij} = P_i t_{ij}$, one obtains

$$P_i t_{ij} = \frac{Q_{ij}^{\frac{-1}{\sigma}} Y_j}{\sum_i Q_{ij}^{\frac{\sigma-1}{\sigma}}}$$

Expressing further, by properly decomposing the exponent $(\ldots)^{-\sigma}$ at the numerator will result in:

$$(P_i t_{ij})^{-\sigma} = \frac{Q_{ij} Y_j^{-\sigma}}{\sum_i Q_{ij}^{\left(\frac{\sigma-1}{\sigma}\right)^{-\sigma}}}$$

Multiplying both sides of the equation by $P_i t_{ij}$, gives:

$$(P_i t_{ij})^{1-\sigma} = \frac{P_i t_{ij} Q_{ij} Y_j^{-\sigma}}{\sum_i Q_{ij}^{\left(\frac{\sigma-1}{\sigma}\right)^{-\sigma}}} \tag{22}$$

As a demand-side approximation, the CES utility function is price characteristic. The CES price index is needed to determine equilibrium prices, which defines the price level. The models presented by Anderson

(1979), Bergstrand (1989) and Eaton & Kortum (2002) also utilised price structure. The purpose here is to establish a price structure that can be interpreted as a price index. To obtain this price would require summing up all the i from the left-hand side of the equation and also substitute $\sum_i P_i Q_{ij} t_{ij}$ with Y_j, this gives:

$$\sum_i \left(P_i t_{ij}\right)^{1-\sigma} = \frac{Y_j^{1-\sigma}}{\sum_i Q_{ij}^{\left(\frac{\sigma-1}{\sigma}\right)^{-\sigma}}} \qquad (23)$$

Next is to realize the consumer price index of Anderson and van Wincoop (2003) for country j:

$$P_j = \left(\sum_i \left(P_i t_{ij}\right)^{1-\sigma}\right)^{\frac{1}{(\sigma-1)}} \qquad (24)$$

The equation can further be simplified with an equivalence given as:

$$P_j^{1-\sigma} = \frac{Y_j^{1-\sigma}}{\sum_i Q_{ij}^{\left(\frac{\sigma-1}{\sigma}\right)^{-\sigma}}} \equiv \sum_i Q_{ij}^{\left(\frac{\sigma-1}{\sigma}\right)^{-\sigma}} = \frac{Y_j^{1-\sigma}}{P_j^{1-\sigma}}$$

Substituting equation $\sum_i Q_{ij}^{\left(\frac{\sigma-1}{\sigma}\right)^{-\sigma}}$ into equation (22) leads to:

$$\left(P_i t_{ij}\right)^{1-\sigma} = \frac{P_i t_{ij} Q_{ij} Y_j^{-\sigma}}{Y_j^{-\sigma}} P_j^{1-\sigma}$$

By further replacing $P_i t_{ij} Q_{ij}$ with X_{ij} (see equation (17) describing the expenditure and income equilibrium of the importer and the exporter) one obtains:

$$\left(P_i t_{ij}\right)^{1-\sigma} = \frac{X_{ij}}{Y_j^{-\sigma}} P_j^{1-\sigma}$$

Solving for X_{ij} leads to determining the demand function of country j for imports from country i, which is

$$X_{ij} = \frac{\left(P_i t_{ij}\right)^{1-\sigma}}{P_j^{1-\sigma}} Y_j \qquad (25)$$

By restoring the shift parameter A_i, equation (25) is equivalent to the equation of Anderson and van Wincoop (2004) given as: $X_{ij} = \left(\frac{A_i P_i t_{ij}}{P_j}\right)^{1-\sigma} y_j$. The main purpose here is to determine the aggregate demand and the general equilibrium trade flows. Through the market clearing condition, all the countries can be aggregated. Thus, by substituting equation (25) into equation of the market clearing condition (18), one obtains:

$$Y_i = \sum_j \left(\frac{P_i t_{ij}}{P_j}\right)^{1-\sigma} Y_j$$

The market equilibrium price can now be determined as:

$$(P_i)^{1-\sigma} = \frac{Y_i}{\sum_j \left(\frac{t_{ij}}{P_j}\right)^{1-\sigma} Y_j} \qquad (26)$$

Substituting the price $(P_i)^{1-\sigma}$ back into the import demand equation (25), X_{ij} , demand becomes:

$$X_{ij} = \left(\frac{P_i t_{ij}}{P_j}\right)^{1-\sigma} Y_j = \left(\frac{t_{ij}}{P_j}\right)^{1-\sigma} Y_j \, Y_i \left(\sum_j \left(\frac{t_{ij}}{P_j}\right)^{1-\sigma} Y_j\right)^{-1}$$

But the expenditures of individual countries (Y_j) and the income of each country (Y_i) are both shares in the world aggregate income (Y_w), hence the equation changes to:

$$X_{ij} = \left(\frac{t_{ij}}{P_j}\right)^{1-\sigma} \frac{Y_j \, Y_i}{Y_w} \left(\sum_j \left(\frac{t_{ij}}{P_j}\right)^{1-\sigma} \frac{Y_j}{Y_w}\right)^{-1} \qquad (27)$$

Expressing the equation in terms of Y_w: i.e., $\frac{1}{Y_w}\left(\frac{1}{Y_w}\right)^{-1} = 1$, and to keep the unfolding equations simple and easy, the term $\left(\frac{t_{ij}}{P_j}\right)^{1-\sigma}$ on the right hand-side of equation (27) is defined as $(\Pi_i)^{1-\sigma}$ (see Anderson and van Wincoop 2004). Hence:

$$(\Pi_i)^{1-\sigma} = \sum_j \left(\frac{t_{ij}}{P_j}\right)^{1-\sigma} \frac{Y_j}{Y_w}$$

Putting back the term $(\Pi_i)^{1-\sigma}$ into equation (27) gives the final structural gravity model as:

$$X_{ij} = \frac{Y_j \, Y_i}{Y_w} \left(\frac{t_{ij}}{P_j \Pi_i}\right)^{1-\sigma} \qquad (I)$$

$$(\Pi_i)^{1-\sigma} = \sum_j \left(\frac{t_{ij}}{P_j}\right)^{1-\sigma} \frac{Y_j}{Y_w} \qquad (II)$$

$$\left(P_j\right)^{1-\sigma} = \sum_i \left(\frac{t_{ij}}{\Pi_i}\right)^{1-\sigma} \frac{Y_i}{Y_w} \qquad (III)$$

If the process leading to $\left(P_j\right)^{1-\sigma}$ is known, then the actual values of P_j and Π_i can be determined. This is done by first substituting the isolated price term of equation (26) into the CES price index in equation (24) to give:

$$P_j = \left(\sum_i \left(P_i t_{ij}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}} \left(\sum_i \left(t_{ij}\right)^{1-\sigma} Y_i \left(\sum_j \left(\frac{t_{ij}}{P_j}\right)^{1-\sigma} Y_j\right)^{-1}\right)^{\frac{1}{1-\sigma}}$$

Expressing the equation further by dividing and multiplying with world income, leads to:

$$P_j = \left(\sum_i \left(t_{ij}\right)^{1-\sigma} \frac{Y_i}{Y_w} \left(\sum_j \left(\frac{t_{ij}}{P_j}\right)^{1-\sigma} \frac{Y_j}{Y_w}\right)^{-1}\right)^{\frac{1}{1-\sigma}}$$

Substituting Π_i in place of the term on the hand side and multiplying the superscript with $(1 - \sigma)$ gives:

$$(P_j)^{1-\sigma} = \sum_i \left(\frac{t_{ij}}{\Pi_i}\right)^{1-\sigma} \frac{Y_i}{Y_w}$$

Hence, a simple solution to the gravity model exists where Anderson and van Wincoop (2003) find a symmetric transportation cost $t_{ij} = t_{ji}$. Given that P_j is the CES price index, then Π_i can be defined as the factory gate price and a solution to $(\Pi_i)^{1-\sigma}$ and $(P_j)^{1-\sigma}$. By transposing the superscript $(1-\sigma)$ at the left hand-side, leads to the multilateral resistance of Anderson and van Wincoop, as: [ix]

$$P_j = \left(\sum_i \left(\frac{t_{ij}}{\Pi_i}\right)^{1-\sigma} \frac{Y_i}{Y_w}\right)^{\frac{1}{\sigma-1}} \tag{28}$$

The same expression can be done for $(\Pi_i)^{1-\sigma}$ and yields:

$$\Pi_i = \left(\sum_j \left(\frac{t_{ij}}{P_j}\right)^{1-\sigma} \frac{Y_j}{Y_w}\right)^{\frac{1}{\sigma-1}} \tag{29}$$

Equation (I) represents the gravity equation and equation (II) and (III) are the inward multilateral resistance and the outward multilateral resistance terms, respectively. This implies that trade flow between country i and j depends on three resistances:

- The bilateral trade resistance reflected by the trade cost t_{ij}
- Outward multilateral resistance faced by the exporting country i at the factory gate, to all other trading partners. It determines how easily the exporter country i can ship goods to markets in country j.
- Inward multilateral resistance faced by the importing country j from other trading partners. It determines how easily the importer country j can import goods form country i.

Thus, this implies that the vector of bilateral trade costs relative to the inward and outward multilateral resistance terms is what actually matters for the volume of trade. Finally, from an empirical standpoint, the trade elasticity in Armington model is considered to be constant and is determined by the elasticity of substitution across goods (σ). The multilateral resistance terms transmit the micro level effects of trade policy at the cross-country level to single-country impacts on both consumer and producer prices (Anderson and van Wincoop, 2003). Trade flows are also affected by the direct impact of trade costs. A good example can be illustrated by how NAFTA and European Union (EU) traded among one another, which run contrary to how the trade unions traded with other non-member countries.

Global Journal of Arts, Humanities and Social Sciences, 14(1),1-23, 2026
Print ISSN: 2052-6350(Print)
Online ISSN: 2052-6369(Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

The Synthesis

The synthesis of contemporary gravity literature reveals that the evolution of the model has been characterized by a progressive deepening of its theoretical and empirical coherence. Each refinement from the early intuitive formulations to the structural and micro-founded versions has incrementally enhanced the model's capacity to explain real-world trade patterns and to inform policy analysis.

The findings establish that while the original gravity model, as proposed by Tinbergen (1962), possessed strong empirical validity, it lacked a firm theoretical foundation. The later integration of microeconomic principles, particularly through Anderson (1979), provided the first formal justification of the model by introducing Armington's assumption of national product differentiation and incorporating expenditure shares as determinants of trade flows. This laid the groundwork for the transition from empirical correlation to theoretically grounded structural modeling.

The successive developments by Deardorff (1998) and Bergstrand (1985, 1989) reinforced the centrality of distance and trade impediments as determinants of bilateral trade. Their inclusion of transportation costs, tariff structures, and factor endowment variables demonstrated that the gravity model could encompass both inter- and intra-industry trade under conditions of imperfect competition. Deardorff's distinction between frictionless and impeded trade provided a framework for analyzing how shipment costs and factor mobility shape trade volumes and patterns.

The introduction of monopolistic competition and firm-level heterogeneity through Helpman and Krugman (1985), Melitz (2003), and Chaney (2008) marked a fundamental theoretical breakthrough. These models established that trade flows are driven not only by aggregate country characteristics but also by firm-level productivity differences and fixed costs of exporting. The implication is that the composition and direction of trade depend on both the intensive margin (volume of trade among existing exporters) and the extensive margin (number of exporting firms). This micro-foundation improved the explanatory and predictive strength of the gravity model.

Eaton and Kortum (2002) advanced the Ricardian interpretation of the gravity model by demonstrating that bilateral trade flows arise from the interplay between technological heterogeneity (reflecting absolute advantage) and geographic barriers (reflecting transport costs). Their model linked comparative advantage directly to technology distributions across countries, providing a quantifiable explanation for how technological capabilities and productivity differentials influence trade elasticity. This linkage bridged the conceptual divide between Ricardian and Heckscher-Ohlin frameworks.

A critical empirical advancement emerged from Anderson and van Wincoop (2003, 2004), who addressed McCallum's (1995) "border effect" puzzle in Canada–U.S. trade. Their inclusion of multilateral resistance terms corrected for omitted variable bias and established that bilateral trade depends on relative, not absolute, trade costs. This innovation transformed the gravity model into a fully general equilibrium system, capable of accounting for all direct and indirect trade linkages. The multilateral resistance framework also clarified why geographically proximate or institutionally similar economies often experience disproportionate intra-bloc trade relative to external trade partners.

Relevance and Key Findings

The structural gravity model captures the interaction between trade barriers, price levels, and income distributions across countries, enabling comprehensive welfare analysis. It provides a theoretical mechanism for translating micro-level trade cost changes into macro-level trade and welfare effects. Thus, the model serves not only as a descriptive tool but also as a quantitative framework for policy simulation allowing for the assessment of how changes in tariffs, non-tariff measures, or transportation infrastructure affect national and global welfare.

The core findings from the theoretical trajectory indicate that:

- Bilateral trade is a function of economic mass, relative trade resistance, and multilateral linkages rather than simple physical distance.
- Trade costs both observable and implicit act as policy-contingent variables that can be shaped by institutional reforms and technological advancements.
- Firm heterogeneity and technological variation introduce structural asymmetries that explain cross-country and intra-industry trade diversity.
- The structural gravity model effectively reconciles diverse trade theories (Ricardian, Heckscher-Ohlin, and new trade theory) under a single analytical framework.

Policy Insight and Implications

The evolution of the gravity model from a descriptive empirical framework to a structural, theory-consistent model has transformed it into a vital instrument for contemporary trade policy reform. The refined gravity approach anchored in the general equilibrium framework of Anderson and van Wincoop (2003), complemented by firm heterogeneity and technological differentiation offers a rigorous analytical basis for evaluating how policy-driven and structural frictions shape international trade. Its insights provide a powerful platform for designing evidence-based reforms aimed at fostering efficiency, inclusivity, and resilience in the global trading system.

A core implication of the structural gravity model is that trade policy effectiveness depends on the relative, not absolute, reduction of trade costs. The introduction of the multilateral resistance terms redefines how policymakers should interpret the outcomes of bilateral or regional liberalization. Trade gains from reduced tariffs or improved border procedures depend on how these reforms alter a country's trade costs relative to those of all other potential trading partners. This finding calls for coordinated, system-wide reforms that transcend bilateral agreements and target the overall efficiency of international trade networks. For developing and middle-income economies, this perspective reinforces the need for harmonization of trade facilitation measures, convergence in standards, and policy coherence between regional and multilateral frameworks.

Second, the model establishes that trade costs are policy variables amenable to reform. These costs encompass tariffs, non-tariff barriers, transport inefficiencies, regulatory heterogeneity, and institutional frictions. Structural gravity analysis enables policymakers to decompose these components and simulate the welfare and distributional effects of specific reforms such as reducing non-tariff barriers, modernizing customs procedures, or improving infrastructure. Such counterfactual simulations have become essential tools for evidence-based trade policymaking, enabling governments to predict how incremental reforms translate into measurable welfare gains and competitiveness improvements. This methodological evolution strengthens the empirical foundation for policy evaluation and prioritization.

Third, the incorporation of heterogeneous firms and fixed export costs (Melitz, 2003; Chaney, 2008) into gravity theory introduces an important dimension for trade reform design. Policy instruments that reduce aggregate trade barriers are necessary but not sufficient; they must also lower firm-level entry and compliance costs to ensure that smaller and medium-sized enterprises can participate in global markets. Institutional reforms that streamline export procedures, expand access to trade finance, and promote technological upgrading can amplify the extensive margin of trade. Thus, microeconomic reforms supporting firm productivity and innovation are complementary to macro-level trade liberalization efforts.

Fourth, the technological interpretation of comparative advantage advanced by Eaton and Kortum (2002) reinforces the strategic role of innovation policy in trade reform. Differences in productivity distributions across countries explain much of the variation in global trade flows. Accordingly, countries aiming to expand their export base and move up the value chain must invest in human capital, digital infrastructure, and R&D ecosystems. Trade liberalization should be accompanied by policies that strengthen absorptive capacity and technological learning. Such complementary measures enhance a nation's capacity to internalize the benefits of globalization and mitigate exposure to external shocks.

Fifth, the general equilibrium structure of the gravity model highlights the interdependence between domestic institutional quality and trade outcomes. Weak institutions, inefficient logistics, and macroeconomic instability amplify multilateral resistance and undermine the effectiveness of trade reforms. Consequently, policy reforms must target not only tariff and non-tariff measures but also the institutional foundations of trade such as transparency, regulatory predictability, and governance of trade-related infrastructure. Empirical applications of the structural gravity model provide diagnostic insights into how institutional reforms can reduce hidden trade costs and foster integration into global markets.

Finally, the model underscores the redistributive and welfare dimensions of trade reform. Because the benefits of liberalization are transmitted through the network of trade resistances, gains are uneven across sectors and regions. Policymakers must therefore complement liberalization measures with adjustment policies including worker retraining, regional diversification, and fiscal mechanisms that ensure inclusive participation in the gains from trade. The gravity framework's ability to simulate welfare effects across countries and income groups offers a practical instrument for designing equitable reform packages that align efficiency with social sustainability. In summary, the contemporary gravity model provides not only a theoretical foundation for understanding trade patterns but also a policy-engineering framework for reforming trade systems in a globally consistent manner. Its strength lies in integrating microeconomic behaviour, macroeconomic equilibria, and institutional constraints into a unified model of trade adjustment. For policymakers, it delivers actionable insights into where and how reforms yield the greatest welfare impact, how institutions mediate trade efficiency, and how liberalization can be sequenced to balance openness with resilience. In an era of shifting global trade architectures and emerging economic nationalism, the structural gravity model remains an indispensable guide for designing evidence-based, reform-oriented, and institutionally grounded trade policies that promote sustainable globalization.

Future Directions and Policy Extensions

The contemporary refinement of the gravity model has opened new frontiers for both theoretical inquiry and applied trade policy analysis. While this paper has focused primarily on the model's conceptual evolution, several promising directions remain for advancing research and deepening its policy relevance particularly in the Canadian and subnational contexts where globalization and regional integration intersect. A priority area for future research lies in the empirical operationalization of structural gravity equations. Building upon Anderson and van Wincoop's (2003) general equilibrium framework, future studies could employ Poisson Pseudo-Maximum Likelihood (PPML) estimations or related structural methods to assess the magnitude of multilateral resistance and trade elasticity effects using Canadian provincial data. Such empirical validation would enable quantification of trade frictions within and across provinces, thereby translating theoretical constructs into measurable policy variables. Simulation-based approaches could further test counterfactual scenarios—such as reductions in interprovincial trade costs, tariff changes under regional trade agreements, or infrastructure improvements providing actionable insights for policymakers.

Another valuable extension involves comparing the structural gravity framework with alternative modeling traditions, including New Quantitative Trade Models (NQTMs), Computable General Equilibrium (CGE) approaches, and Global Value Chain (GVC) analysis. This comparative perspective would enrich understanding of how the gravity model complements or diverges from other frameworks in estimating welfare effects and policy multipliers. By integrating firm heterogeneity, technological diffusion, and services trade into gravity formulations, future research can align the model more closely with contemporary trends in digitalization and global production fragmentation. At the subnational level, the model's policy potential is especially significant for evaluating interprovincial trade barriers and internal market integration. The persistent evidence of border effects between Canadian provinces suggests that internal trade costs remain substantial despite existing federal coordination. Structural gravity estimations can help identify the sources of these inefficiencies whether regulatory divergence, transportation bottlenecks, or institutional frictions and guide targeted policy reforms. Strengthening trade infrastructure, harmonizing standards, and simplifying logistics procedures could thus enhance both domestic efficiency and international competitiveness. Furthermore, provincial-level applications can reveal how globalization's benefits are distributed regionally, informing more equitable and evidence-based trade policies. From a policy standpoint, future research should also explore how trade and innovation policies interact within the gravity framework. Incorporating technological asymmetries and productivity dynamics (Eaton & Kortum, 2002) would allow for the assessment of how R&D investment, digital trade facilitation, and human capital development influence comparative advantage. This would position trade liberalization not merely as a mechanism for market expansion but as part of a broader strategy for productivity-led growth and regional diversification.

Finally, a comprehensive research agenda should acknowledge both the limitations and future adaptability of the gravity model. Assumptions of constant elasticity, homothetic preferences, and friction symmetry may constrain its realism in some contexts. Expanding the model to incorporate non-linear cost structures, policy uncertainty, and climate-related trade disruptions would enhance its analytical power. Moreover, greater integration of subnational data, firm-level microdata, and policy simulations will strengthen its role as a practical instrument for trade reform design. In sum, the structural gravity model's evolution from a descriptive framework to a theoretically consistent and policy-relevant tool provides fertile ground for continued empirical testing and methodological innovation. Future studies that bridge its theoretical precision with real-world applications particularly in the contemporary Canadian context will not only refine our understanding of trade mechanisms but also enhance the capacity of policymakers to design reforms that are evidence-based, inclusive, and globally coherent.

CONCLUSION

The evolution of the theoretical gravity model represents one of the most significant advancements in modern international trade theory. From its empirical origins in the 1960s to its contemporary structural formulations, the gravity model has transcended its early descriptive nature to become an empirically validated, theoretically consistent, and policy-relevant analytical tool. The review demonstrates that no single theory can singularly account for the observed complexity of international trade flows. Instead, the gravity framework integrates insights from multiple theoretical traditions comparative advantage, factor proportions, monopolistic competition, and heterogeneous firms into a unified system capable of explaining both inter- and intra-industry trade. This theoretical pluralism underlines the model's robustness and adaptability. The findings reaffirm that trade flows are shaped by both economic fundamentals (such as income, production capacity, and technology) and policy-determined frictions (such as tariffs, distance-related costs, and regulatory barriers). The inclusion of multilateral resistance has refined the model's predictive power by embedding bilateral trade within a global equilibrium context. This allows researchers and policymakers to estimate not only the direct effects of trade policies but also their spillovers across trading partners. From a policy standpoint, the refined gravity model offers an empirical framework for evidence-based trade policymaking. It enables the quantitative assessment of how trade costs and policy reforms affect welfare, production, and market integration. The model's equilibrium-consistent nature ensures that policy analysis accounts for feedback effects and cross-country dependencies making it indispensable for evaluating trade agreements, tariff reforms, and globalization's subnational impacts.

In essence, the contemporary gravity model transforms theoretical trade analysis into an operational policy tool. Its ability to connect micro-level decisions with macroeconomic outcomes makes it particularly valuable for designing reform-oriented, inclusive, and sustainable trade strategies. As global trade patterns continue to evolve amid technological disruption and geopolitical shifts, the structural gravity model remains the cornerstone of modern trade analysis linking theoretical refinement with practical policy relevance.

REFERENCES

Anderson, J. E. (1979). A theoretical foundation for the gravity equation. *American Economic Review, 69*(1), 106–116.

Anderson, J. E., & van Wincoop, E. (2003). Gravity with gravitas: A solution to the border puzzle. *American Economic Review, 93*(1), 170–192.

Armington, P. S. (1969). A theory of demand for products distinguished by place of production. *IMF Staff Papers, 16*(1), 159–178.

Bergstrand, J. H. (1985). The gravity equation in international trade: Some microeconomic foundations and empirical evidence. *The Review of Economics and Statistics, 67*(3), 474–481.

Bergstrand, J. H. (1989). The generalized gravity equation, monopolistic competition, and the factor-proportions theory in international trade. *The Review of Economics and Statistics, 71*(1), 143–153.

Carey, H. (1865). *Principles of social science.* Philadelphia, PA: J. B. Lippincott and Company.

Chamberlin, E. (1933). *The theory of monopolistic competition.* Cambridge, MA: Harvard University Press.

Chamberlin, E. (1962). *The theory of monopolistic competition: A re-orientation of the theory of value.* Cambridge, MA: Harvard University Press.

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

Deardorff, A. V. (1984). Testing trade theories and predicting trade flows. In R. W. Jones & P. B. Kenen (Eds.), *Handbook of international economics* (Vol. 1, pp. 467–517). Amsterdam: North-Holland.

Deardorff, A. V. (1998). Determinants of bilateral trade: Does gravity work in a neoclassical world? In J. A. Frankel (Ed.), *The regionalization of the world economy* (pp. 7–22). Chicago, IL: University of Chicago Press.

Dixit, A. K., & Stiglitz, J. E. (1977). Monopolistic competition and optimum product diversity. *American Economic Review, 67*(3), 297–308.

Eaton, J., & Kortum, S. (2002). Technology, geography, and trade. *Econometrica, 70*(5), 1741–1779.

Feenstra, R. C. (2004). *Advanced international trade: Theory and evidence.* Princeton, NJ: Princeton University Press.

Grogger, J., & Hanson, G. H. (2011). Income maximization and the selection and sorting of international migrants. *Journal of Development Economics, 95*(1), 42–57.

Head, K., & Ries, J. (2008). FDI as an outcome of the market for corporate control: Theory and evidence. *Journal of International Economics, 74*(1), 2–20.

Head, K., & Mayer, T. (2014). Gravity equations: Workhorse, toolkit, and cookbook. In G. Gopinath, E. Helpman, & K. Rogoff (Eds.), *Handbook of international economics* (Vol. 4, pp. 131–195). Amsterdam: North-Holland.

Head, K., & Mayer, T. (2015). Gravity equations: Workhorse, toolkit, and cookbook. *Handbook of International Economics, 4*, 131–195.

Helpman, E., & Krugman, P. R. (1985). *Market structure and foreign trade: Increasing returns, imperfect competition, and the international economy.* Cambridge, MA: MIT Press.

Isard, W. (1960). *Methods of regional analysis: An introduction to regional science.* Cambridge, MA: MIT Press.

Krugman, P. R. (1979). Increasing returns, monopolistic competition, and international trade. *Journal of International Economics, 9*(4), 469–479.

Krugman, P. R. (1980). Scale economies, product differentiation, and the pattern of trade. *American Economic Review, 70*(5), 950–959.

Krugman, P. R. (1998). Space: The final frontier. *Journal of Economic Perspectives, 12*(2), 161–174.

Krugman, P. R., & Obstfeld, M. (2006). *International economics: Theory and policy.* Boston, MA: Pearson Addison Wesley.

Leamer, E. E., & Levinsohn, J. (1995). International trade theory: The evidence. In G. Grossman & K. Rogoff (Eds.), *Handbook of international economics* (Vol. 3, pp. 1339–1394). Amsterdam: Elsevier Science BV.

Lindner, S. B. (1961). *An essay on trade and transformation.* New York, NY: John Wiley and Sons.

McCallum, J. (1995). National borders matter: Canada–US regional patterns. *American Economic Review, 85*(3), 616–623.

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

Pepall, L., Richards, D., & Norman, G. (2008). *Industrial organization: Contemporary theory and empirical application.* Malden, MA: Blackwell.

Portes, R., & Rey, H. (2005). The determinants of cross-border equity flows. *Journal of International Economics, 65*(2), 269–296.

Redding, S. J. (2011). Theories of heterogeneous firms and trade. *Annual Review of Economics, 3*(1), 77–105.

Rose, A. K. (2004). Do we really know that the WTO increases trade? *American Economic Review*, *94*(1), 98–114.

Rose, A. K., & van Wincoop, E. (2001). National money as a barrier to international trade: The real case for currency union. *American Economic Review*, *91*(2), 386–390.

Samuelson, P. A. (1952). Spatial price equilibrium and linear programming. *American Economic Review, 42*(3), 560–580.

Samuelson, P. A. (1954). The transfer problem and transport costs: The terms of trade when impediments are absent. *Economic Journal*, *64*(254), 264–289.

Varian, H. R. (2004). *Grundzüge der Mikroökonomik.* München, Germany: Oldenbourg.

Endnotes

[i] Armington (1969) distinguishes products not only by their peculiarity, but also by their place of formation, implying that the supplier's origin is an important factor in determining the characteristics of the product. Armington argued that the demand of two products of similar characteristics but originating from different places are imperfect substitutes. This is an ad hoc assumption, and it completely ignores the "classical" trade forces such as increased specialization due to comparative advantage.

[ii] Cobb-Douglas utility functions take the form ∪ (T, NT) = $A\ x_1^{\alpha} x_2^{\beta}$ where A is a constant term and x_1 and x_2 represents the tradable and non-tradable goods. As a homothetic preference, the elasticities are such that $\alpha + \beta = 1$, indicating that trading regions with a rising income spends the same proportion in tradable goods in relation to their total income due to a constant marginal rate of substitution (Varian, 2004; Pepall, Richards and Norman, 2008).

[iii] Monopolistic competition is an idea that goes back to Chamberlain (1933). Dixit-Stiglitz (1977) developed the model of monopolistic competition, which has become remarkable workhouse in many areas of economics. It is a simple equilibrium model that deals with the study of optimum product diversity (where goods are close substitutes within the market but are not necessarily substitutes for the rest of the goods in the economy).

[iv] The essence of monopolistic competition is that there are static equilibrium profits from monopoly power, but there are no ex-ante profits, such that the static monopoly profits are just enough to meet entry costs. So, profits will equal to the fixed entry or innovation costs the firm has incurred previously – that is at a breakeven point, where profits are driven to zero (Chamberlin, 1962).

[v] Love of variety is defined by Dixit-Stiglitz preferences, which means more variety of goods preferred by the consumer, or the tendency for increased productivity in final good sector, due to more variety of intermediate goods (Dixit and Stiglitz, 1977).

[vi] The main contribution under this circumstance is that the gravity model can essentially evolve from an Heckscher-Ohlin world (without any recourse to monopolistically competitive settings - as in Bergstrand (1989).

[vii] The term p_j^I is the CES index, representing the transport cost factors of county j denoted as $p_j^I = \left(\sum_{\mathrm{i}} \beta_{\mathrm{i}} t_{\mathrm{ij}}{}^{1-\sigma} p_{\mathrm{ij}}{}^{1-\sigma}\right)^{1/1-\sigma}$, and defined as the average of the supplier's distance $\delta_j^S = \left(\sum_{\mathrm{i}} \beta_{\mathrm{i}} t_{\mathrm{ij}}{}^{1-\sigma}\right)^{1/1-\sigma}$. For more clarification on the derivation, see Deardorff, 1998, 18-20.

[viii] Samuelson (1952 &54) assumed that in order to export goods produced in a particular origin location to another destination location a constant fraction of the good melts away in transit depending on the size of the iceberg, such that total transportation costs equal the cost of producing the melted goods (see Eaton and Kortum, 2002). The iceberg transportation cost has become one of the key ingredients of contemporary trade and economic geography model. Krugman (1998, 164) referred to the concept as an important "trick of the genre" in his treatment of economic geography model.

[ix] Anderson and van Wincoop (2003) coined out the multilateral resistance term and defined it as the theoretically appropriate average trade barrier. Whereas the inward multilateral resistance term represents the ease at which the importer country j can have access to the market, the outward multilateral resistance term measures the ease at which exporter country i can have access to the market (Anderson and van Wincoop, 2003). Given that trade is assumed to be separable, the supply-side market structure is not presented. The analysis captures the market structure of the demand-side, duly represented by the CES utility function (see equation 14). Therewith, Π_{i} – the supply price is perfectly elastic and becomes equal to the marginal cost in a perfectly competitive market setting. In the case of monopoly, the supplier price will be equal to (Π_{i} + mark-up).