RSCAS 2014/74
Robert Schuman Centre for Advanced Studies
Global Governance Programme-118

Specification and Estimation of Gravity Models: A Review of the Issues in the Literature

Fatima Olanike Kareem and Olayinka Idowu Kareem

European University Institute
Robert Schuman Centre for Advanced Studies
Global Governance Programme

Specification and Estimation of Gravity Models: A Review of the Issues in the Literature

Fatima Olanike Kareem and Olayinka Idowu Kareem

EUI Working Paper **RSCAS** 2014/74

ISSN 1028-3625

Printed in Italy, June 2014
European University Institute
Badia Fiesolana
I – 50014 San Domenico di Fiesole (FI)
Italy
www.eui.eu/RSCAS/Publications/
www.eui.eu
cadmus.eui.eu

Robert Schuman Centre for Advanced Studies

The Robert Schuman Centre for Advanced Studies (RSCAS), created in 1992 and directed by Brigid Laffan since September 2013, aims to develop inter-disciplinary and comparative research and to promote work on the major issues facing the process of integration and European society.

The Centre is home to a large post-doctoral programme and hosts major research programmes and projects, and a range of working groups and *ad hoc* initiatives. The research agenda is organised around a set of core themes and is continuously evolving, reflecting the changing agenda of European integration and the expanding membership of the European Union.

Details of the research of the Centre can be found on:
http://www.eui.eu/RSCAS/Research/

Research publications take the form of Working Papers, Policy Papers, Distinguished Lectures and books. Most of these are also available on the RSCAS website:
http://www.eui.eu/RSCAS/Publications/

The Global Governance Programme at the EUI

The Global Governance Programme (GGP) is research turned into action. It provides a European setting to conduct research at the highest level and promote synergies between the worlds of research and policy-making, to generate ideas and identify creative and innovative solutions to global challenges.

The GGP comprises three core dimensions: research, policy and training. Diverse global governance issues are investigated in *research* strands and projects coordinated by senior scholars, both from the EUI and from other internationally recognized top institutions. The *policy* dimension is developed throughout the programme, but is highlighted in the GGP High-Level Policy Seminars, which bring together policy-makers and academics at the highest level to discuss issues of current global importance.The Academy of Global Governance (AGG) is a unique executive *training* programme where theory and “real world” experience meet. Young executives, policy makers, diplomats, officials, private sector professionals and junior academics, have the opportunity to meet, share views and debate with leading academics, top-level officials, heads of international organisations and senior executives, on topical issues relating to governance.

For more information:
http://globalgovernanceprogramme.eui.eu

Abstract

The gravity model has become an efficient tool in the analysis of international economic relations due to its theoretical derivation and ability to explain these relationships. The contending issue now is the appropriate specification and estimation techniques. This paper presents a review of current controversy surrounding the specification and estimation of gravity model with zero trade data, which we called ‘gravity modeling estimation debate’. Different positions in the literature were enunciated with the view of bringing the readers to the frontier of knowledge in this area of empirical strategies revolving on the gravity modeling in the presence of zero trade. By and large, the identification of the most appropriate estimation technique in the presence of zero trade is still an empirical issue. This paper deduced that the choice of the estimation technique should largely be based on the research questions, the model specification and the choice of data to be used for the analysis.

Keywords

Gravity Model, Specification, Estimation, Debate

JEL Classification: C13, C51, F10

1.0 Introduction

Analysis of International economics, especially on the effects of bilateral, regional and multilateral relationships has given credence to the potency of gravity models in explaining impact of these relationships. Gravity models have emerged as important and popular model in explaining and predicting bilateral trade flows. In fact, it is among the most robust empirical regularities in economics (Chenery, 2014). The model has been used to analyze the economic impacts of trade, investment, migration; currency union, regional trade agreements, etc. It has become the workhorse or toolkit in international trade (see Head and Mayer, 2013), in which the proven popularity are primarily due to its exceptional success in predicting bilateral trade flows and the theoretical foundations given to it by both the old, new and “new” new trade theories. However, prior to its general acceptance, there have been several criticisms about its lack of strong theoretical application, which was later justified by the notable work of Anderson (1979), Bergstrand (1989), Deardorff (1998), Helman and Krugman (1985), etc, all of whom gave theoretical justifications to the model.

While the theoretical justification for the model is no longer in doubt, nonetheless, its empirical application has however generated several unresolved controversies. These controversies revolve around the appropriate estimation technique and specification of the gravity equation, in which the former has generated several debates in the literature. The first concern is the estimation challenges which revolve around the validity of the log linear transformation of the gravity equation in the presence of heteroscedasticity and zero trade observation. The challenges posed by the validity of the log linear gravity equation arise from the conventional practice in the literature which is to log linearizing the multiplicative gravity equation. This is estimated using ordinary least square (OLS) or by employing panel data techniques with the usual assumption of homoscedasticity across country pairs or countries (Gomez-Herrera, 2013). However, Santos Silva and Tenreyro (2006, 2011) pointed out that due to the logarithmic transformation of the equation, OLS estimator may be inconsistent in the presence of heteroscedasticity and non-linear estimators should be used.

Also, there are challenges presented by the appropriate choice of the estimation techniques in the presence of zero trade values that is very common in trade data, and particularly pervasive in disaggregated data. Usually, the common practice in the literature in dealing with these zero trade observations are by employing the truncation method where the zero trade observations are deleted completely from the trade matrix, or censoring method where the zeros are substituted by a small positive constant an arbitrary small value. However, Flowerdew and Aitkin (1982), Eichengreen and Irwin (1998), Linders and de Groot (2006) and Burger et al. (2009) posit that these methods are arbitrary, are without any strong theoretical or empirical justification and can distort the results significantly, leading to inconsistent estimates. In addition, Heckman (1979) posit that if the zeros are not random, deleting can lead to loss of information; while including arbitrary constants to the zero observations are tantamount to deliberately introducing measurement error which can lead to selection bias.

To this end, more appropriate estimation techniques are increasingly employed to deal with the estimation challenges posed by the logarithm transformation and zero trade flow issues in the context of gravity trade literature. The models proposed by Tobit (1959), Heckman (1979) and Helpman, Melitz and Rubinstein (2008) have all been used to deal with the problem associated with zero value trade flows. For instance, the Tobit model was employed by Rose (2004) and Baldwin and DiNino (2002) to deal with the problem of zero valued trade flows which resulted either because the actual trade flows are not observable or due to measurement errors from rounding. However, several studies, notable among them is Linder and de Groot (2006) have argued that the appropriateness of using the Tobit model to estimate zero valued trade flows in a gravity model depends on whether rounding up of trade flows is important or whether the desired trade could be negative. They posit that the desired trade cannot be negative since the zeros do not reflect unobservable trade flows; therefore, one cannot censor trade flows from below it. Likewise, sample selection models were developed by Heckmam (1979) and Helpman et al., (2008) to deal with selection bias resulting from the non-random elimination of zeros from the trade matrix. The sample selection models have also been criticized on the ground that it is difficult to satisfy the exclusion restriction. Further, Santos Sliver and Tenreyro (2009) and Flam and Nordström (2011) show that Helpman et al., (2008) model does not control for heteroscedasticity which is usually pervasive in most trade data, consequently casting doubts on the validity of inferences drawn from the model.

More so, the influential paper by Santos Siliva and Tenreyro (2006) suggest that non-linear estimators, precisely the poisson pseudo maximum likelihood (PPML) should be used to deal with the zero trade observations as it provides unbiased and consistent estimates that are robust to the presence of heteroscedasticity in the data and naturally take care of the zeros observations of the dependent variable. This influential work of Santos Siliva and Tenreyro (2006) has generated a lot of debates in the literature, which we called ‘gravity model estimation debate (GMED)’. The debate centered on the appropriateness of the PPML as the best estimator of gravity model in the presence of zero trade, as advocated by Santos Siliva and Tenreyro (2006). This assertion was contested and faulted, in which alternative estimation techniques have been proposed to accommodate zero trade values in the data (c.f. Burger et al., 2009; Martinez-Zarzaso, 2013; Helpman et al., 2008; Martin and Pham, 2008). In the effort of these studies to identity the best performing estimator, alternative estimation techniques were compared, however, they obtained divergent outcomes. This has further led to rise in the debate in the literature about which of the different alternative estimators performs best. For instance, Santos Siliva and Tenreyro (2006) propose the usage of the PPML as against the usual OLS technique, with the justification that it is consistent in the presence of heteroscedasticity and deals naturally with the zero trade flows. However, Martinez-Zarzaso (2013) found that, although the PPML is less affected by heteroscedastic compared to other estimators, nevertheless, the PPML estimator proposed by Santos Silva and Tenreyro (2006) is not always the best estimator as its estimates are outperformed by both the OLS and FGLS estimates in out of sample forecast.

In response to this, Santos Siliva and Tenreyro (2008) posit that although the other estimators might outperform the PPML in some cases, however, the PPML should be a benchmark against which other alternative estimators be compared due to its identified advantages. Study by Burger et al., (2009) has also challenged that of Santos Siliva and Tenreyro (2006) with the fact that PPML is vulnerable to the problem of overdispersion in the dependent variable and excessive zeros and propose the use of the Negative Binomial Pseudo Maximum Likelihood (NBPML) to correct for the overdispersion in the dependent variable. In addition, they also found PPML and NBPML to be inconsistent in the presence of excessive zero trade observations and propose the usage of the Zero-inflated models which are Zero-inflated Pseudo Maximum Likelihood technique (ZIPML) and Zero-inflated Binomial Pseudo Maximum Likelihood technique (NIBPML) as they are noted to be consistent in the presence of excessive zeros. Similar result has been found by Martinez-Zarzaso (2013) and Martin and Pham, (2008), with the latter claiming that the Heckman model is appropriate for dealing with this issue. Therefore, these raging arguments and counter-arguments in the literature are the focus of this paper in order to bring to fore the recent development in the estimation strategies of zero trade.

To this end, this paper reviews the recent work on the application of gravity models to zero trade. This review does not claim to have exhaustively reviewed the zero trade gravity estimation strategies, but rather to take the readers as close as possible to the current frontier of knowledge in this segment of gravity modeling.

1.1 The Motivation

In line with the aforementioned studies’ positions on the raging issues of the best estimation technique in the presence of zero trade and the appropriate gravity model specifications, this paper review the

GMED with respect to the contributions of these scholars to the frontier of knowledge in the area. The trend in the derivation of gravity models from different trade theories were also shown in this paper.

This paper departs from the work of Head and Mayer (2013) that examines different ways by which the gravity models could be specified and provided a workhorse or toolkit for gravity modeling in trade in goods and beyond. This study specifically focuses on the review of the controversy surrounding the gravity modeling and estimation of zero trade. Similar to the Head and Mayer (2013) was the evaluation of the appropriate gravity model specification by Baldwin and Taglioni (2007) that identified three common mistakes in gravity modeling in the literature, in which they gave each mistake a ‘medal’[1]. Fugazza (2013) reviews the modeling of non-tariff barriers with gravity models and the computed general equilibrium model (CGE) as well as the different conclusions in the literature. The focus of his paper was not the review of issues arising from gravity model specification and estimation.

Evenett and Keller (1998) examine the theoretical derivation of gravity equation from Heckscher – Ohlin (H-O) and Increasing Returns to Scale (IRS) trade theories. They concluded that only few production is perfectly specialized as a result of the differences in factor endowments and that the increasing returns to scale causes perfect product specialization and the gravity equation, while the extent of imperfection in production across countries gives support for the H – O and IRS models. Basically, the paper evaluated and derived the gravity model from these theories, while also determining the reason behind the variation in international production patterns and trade volume. However, this is not the focus of our study, which reviews the specification and estimation issues in zero trade modeling in the literature. A theoretical contribution was made recently by Cheney (2014) when he offers an explanaion of the roles of economic size and distance in a gravity model. He confirms the fact that the size distribution of the firms is empirically well approximated by Zipf’s law and finds a new evidence that larger firms export over longer distances than smaller ones. His explanation for the role of economic size is not new, but confirms existing facts, however, innovation was brought in through the role of distance in a gravity model. He asserted that if the distribution of firm size is pareto, and if the avearge distance squared of a firm’s exports is an increasing power function of its size, then the distance elasticity of trade is constant and equals -1 in the special case of Zipf’s law. This article gave a theoretical validation to the coefficient and sign of distance in gravity model but did not consider zero trade and other specification issues as we have done in this survey. De Benedictis and Taglioni (2011) show the extent to which some of the issues raised 50-year ago by Tinbergen have been the step stones of research agenda over the years. The paper also discusses how many of the empirical and theoretical contributions that followed Tinbergen has dealt with the old problems, among which are the issue of zero trade specification and estimation that the study reviews in one of the sub-sections. However, among the studies reviewed, recent studies such as Martinez-Zarzoso (2013), Helpman et al. (2008) etc., were not considered and their contributions to the discussion on the specification of gravity models in the presence of zero trade were not included. This might be due to the coverage period of the paper, but our paper has put these studies into consideration for review. More so, our paper actually focus on the specification and estimation issues in gravity modeling, particularly the raging debate right from the thought – provoking work of Santos-Silva and Tenreyro (2006) on the best estimator of the gravity model in the presence of zero trade. The conclusion of De Benedictis and Taglioni (2011) was that Heckman two-step procedure and count data modeling were the two main strategies to dealing with the zero trade, however, some criticisms have been leveled against the estimators (see Santos-Silva and Tenreyro, 2006; Helpman et al. 2008; Martinez – Zarzoso, 2013), which our paper considered and reviewed.

Demaria, Rau and Schlueter (2011) examine the state of the art in gravity modeling, especially that relates to the non-tariff measures. The paper reviewed gravity model estimation techniques such as the

---

[1] The medals are gold, silver and bronze depending on the type error or mistake committed (see Baldwin and Taglioni, 2007).

Heckman, Poisson, Negative Binomial and the Zero Inflated models as possible solution to the estimation problems in the log-normal gravity equation. They concluded that the Zero Inflated Negative Binomial Poisson Maximum Likelihood (ZINBPML) regression supercedes other estimators, especially Heckman procedure. However, the study did not give consideration to the feasible generalized least square (FGLS) as proposed by Martinez – Zarzoso (2013), the Gamma Pseudo Maximum Likelihood (GPML) of Manny and Mullay (2001) and Frankel and Wei (1993) non – linear least square. Besides, there are other gravity model estimators like the Tobit model as used by Anderson and Marcoller (2002), Marttin and Pham (2008), Rose (2004) etc. that need to be adequately considered before making the conclusion. More so, the choice of ZINBPML needs to be evaluated in the presence of model misspecification as argued by Staub and Winkeelman (2012), which makes it inconsistent. All these arguments and counter – arguments in the literature are reviewed, which is our focus in this study in order to bring to fore the ongoing debate and current research on the estimation of gravity models with zeros.

Similarly, Gomez – Herrera (2012) surveyed gravity literature with respect to the specifications and estimation techniques. He proceeded to test for the most appropriate estimator using trade data for 80 countries that accounted for 80% of world trade. The conclusion of the study gave credence to the efficacy of Heckman sample selection model among other estimators. The difference in his study and this present comprehensive review of the literature on zero trade estimation is that, his study excluded the Negative Binomial Pseudo Maximum Likelihood and the Zero Inflated Models among the most recently used gravity model with zeros estimators as was considered in our paper. However, our study did not perform any empirical estimation to compare and select the best estimator, since the identification of the most appropriate estimator is not the focus of the paper, but to review the recent development in the zero trade gravity model literature, in terms of the specification and estimation of the models. This will enable users and prospective users of these estimation techniques to know the pros and cons of the estimators and provide them with estimation options that they can choose from in line with their research questions and the available trade data.

2.0 An Overview of Gravity Model

The gravity equation were first used in the nineteenth century by Ravenstein (1885) and then by Zipf (1946), which is contrary to what majority of trade economists believe and rarely mentioned in the literature. However, the formal usage of the model dated back to Tinbergen (1962) and Pöyhönen (1963), both of whom suggest that the functional form of Newtonian gravity could also be used to explain bilateral trade flows between distant countries. This notion of gravity equation is based on Isaac Newton’s proposition of the law of universal gravitation, which states that the gravitation force between two objects ‘i’ and ‘j’ is directly proportional to the multiplication of the masses of the objects and inversely related to the distance between these two objects. The Newtonian gravity equation is given as:

$$GF_{ij} = C\frac{M_i M_j}{D_{ij}} \qquad \text{……………(1)}$$

Where GF is the gravitational force between two masses; C is the gravitational constant; M_i and M_j are the masses and D is distance between the two masses.

The early version of the model may also be expressed roughly in the same notation as:

$$X_{ij} = \beta_0 (Y_i)^{\beta_1} (Y_j)^{\beta_2} (D_{ij})^{\beta_3} \mu_{ij} \qquad \text{……………(2)}$$

Where X_{ij} is the value of bilateral import\exports in current dollars; Y_i, Y_j are respectively the exporters and importers economic masses proxy by their income; D_{ij} is the distance between country-pairs, μ_{ij} is the disturbance term; and β_s are the unknown parameters of the equation.

This specification was first used by Tinbergen (1962) and Pöyhönen(1963) and later used by other scholars. However, Linnemann (1966) used the same specification but augmented it with importer and exporter population. His theoretical basis for the gravity equation was based on the Walrasian general equilibrium framework and the equation is derived as a reduced form equation from a four-equation partial equilibrium model of import demand and export supply function. Here, prices are excluded as they only adjust to equalise demand and supply (Linnemann 1966; Leamer and Stern, 1970). Leamer and Stern (1970) however argued that this theoretical approach is loose, it lacks a compelling economic justification and fails to explain the multiplicative functional form of the gravity equation. Subsequently, Leamer developed a hybrid version of the gravity equation, which has also been faulted as being atheoretical.

Strong criticisms were made against gravity equation due to its lack of strong theoretical foundations; and this made the model to be neglected between late 1960s and late 1970s. Nevertheless, in recent years, the gravity model has again become very popular in explaining trade relations due to two factors. One of these is due to the rigorous theoretical foundation given to it with the advent of trade theories especially the new trade theory. The second and most important is that, it is now very popular due to its notable empirical success in predicting bilateral trade flows of different commodities under different situations (Deardorff, 1984; Leamer and Levinsohn, 1995). This is the reason that most recent studies in trade often adopt the model in explaining bilateral, multilateral and regional trade agreements. In fact, the use of the model has been applied beyond trade; evidence has shown that it has been applied to currency union (Rose 2000, Baldwin, 2006), health (Manning and Mullahy, 2001; Staub and Winkelmann, 2013)**,** foreign direct investment (FDI) (Linnemann, 1966; Egger, 2004, 2007; Egger and Pfaffermayr, 2004) and so on. Thus, the equation has now become a toolkit in international economics.

2.1 Theoretical Foundations for the Gravity Equation

The theoretical basis for the gravity model was formally introduced by Anderson (1979) and later extended by Bergstrand (1985, 1989, 1990), Deardorff (1998), Eaton and Kortum (2002), and Anderson and Van Wincoop, (2003) etcetera. Specifically, the gravity equation has been derived under the classical or standard trade theory, the new and new new trade theories. Under the standard trade models, the explanations and pattern of international trade rely heavily on comparative advantage and differences in production technology (Ricardian model of trade) and differences in relative factor endowments (Heckcher-Ohlin model). These models assume perfect competition and therefore constant returns to scale in production and no attention is paid to increasing returns to scale, imperfect competition and transport costs. However, with the advent of new trade theories, the equation has also been derived under imperfect competition markets and increasing returns to scale (Helpman and Krugman approach).

2.1.1 Major Development in Gravity Equation

A perusal of the literature shows that several developments have occurred to gravity modeling, in terms of their derivation from trade theories. In this section, we examine the different theoretical frameworks that have been used to situate gravity models in the literature.

A. Gravity Equation under Perfect Competition

The derivation of the gravity model from the standard trade theory was pioneered by Anderson (1979), who derived the gravity equation from the trade share expenditure system model that assume that products are differentiated by regions of origin "Armington assumption[2]" (Armington, 1969) and identical homothetic preferences exist across regions, and utility functions are weakly separated between traded and non-traded goods. To justify the theoretical basis of the gravity model, he applied product differentiation framework with identical Cobb Douglas or constant elasticity of substitution (between domestic and imported goods) preference function for all countries, implying a gravity equation of income elasticities of unity. Using a general equilibrium frame work from which reduced forms equations were derived, his final derivation gives the gravity equation for aggregate imports which is a log-linear function in exporter and importers income and population size with a scale term added.

B. Monopolistic Competition and Economies of Scale Gravity Equation Derivation

Using the monopolistic and economies of scale framework of the new trade theory, Helpman and Krugman (1985) gave theoretical basis for the gravity equation by relaxing the strong assumption of perfect competition. They founded the model using a monopolistic competitive frame work in which firms produces slightly differentiated goods, and operate under increasing returns to scale in production. With monopolistic competitive model, product differentiation occurs in line with economies of scale; each firm produce a uniquely differentiated product under increasing returns to scale and distributes its output to all markets including the domestic market under diminishing returns to scale. Assuming consumers have Dixit-Stiglitz [3] preferences, they derive a gravity equation of intra-industry trade which is identical to Anderson (1979). A major limitation of Anderson (1979) and Helpman and Krugman (1985) is the absence of trade barriers (both policy induced such as tariff and natural geographical barriers such as transportation costs) in their gravity equations. This is because they assume that goods are perfectly or costlessly substituted between importing and exporting countries, which gives rise to a frictionless gravity equation of bilateral trade (Bergstrand, 1985). Thus, they cannot be termed a full theoretical foundation of the gravity equation *(ibid).*

Further theoretical justification of the gravity equation of bilateral trade flows were made by Bergstand (1985, 1989, 1990) in a series of papers in which the general equilibrium was derived using monopolistic competitive model with differentiated products and economies of scale, in which he allow a role for transport cost in his gravity equation. In 1985, he applied the microeconomics foundations to the gravity equation using the framework of a general equilibrium model of world trade from which he derived a gravity model that assumes a single factor of production in each country and product differentiation according to the constant elasticity of substitution (CES) utility function. He developed a general equilibrium framework of world trade from the utility and profit maximizing economic agent behaviour. He modeled demand by assuming that utility maximizing consumers in each country are assumed to share a CES preference function and on the supply side, profit maximizing firms in each country have a constant elasticity of transformation (CET) production technology function. Solving both functions produce the bilateral aggregate import demand equations and bilateral aggregate export supply equations respectively.

The equilibrium condition of these functions gives the general equilibrium model of world trade, which is in form of some reduced form equations with only endogenous variables; as the reduced form[4]equations eliminates endogenous variables out of the explanatory part of each equation. Conditioning on further assumptions that utility and production functions are identical and constant

---

[2] Products are differentiated by origin where products produced in different countries (in the same countries) are seen as imperfect substitutes by consumers.

[3] Dixit-Stiglitz preference refers to love of variety where consumers value varieties and their utility increases for all differentiated varieties of the goods that exist (Dixit and Stiglitz, 1977).

[4] Since the reduce form eliminates endogenous exporter and importer income out of the explanatory part of each equation, the solution cannot be a gravity equation, as "a bilateral trade flow equation must include (both) exporter and importer incomes as exogenous variables to be a gravity model by definition" (Bergstrand, 1985: p 475).

across all countries and assumption that country *i* is a small open economy relative to the other markets, allows foreign income and foreign price level to be treated as exogenous. Imposing these assumptions, the general equilibrium system model is solved using a partial equilibrium approach, which produces a 'generalized[5] gravity equation' that includes exporters and importers incomes and prices as exogenous explanatory variables. Transport cost is introduced into the equation and it is given as a price term.

An important distinction of Bergstrand (1985) gravity equation is that he derived a gravity equation that includes an exogenous price variable in the specification. According to him, if aggregate flows are differentiated by country of origin as suggested by the data (perfect substitutability of goods across countries is unlikely to be costless) then, the previous gravity equations are mis-specified as they omit price variables. Thus, he differs from previous studies by allowing for the costs of distribution, marketing and tailoring each country's output to importing markets into his gravity model. He captures these potential distribution by a CET function in which the elasticity of transformation of production is greater than zero. This implies that because firms face distribution, marketing and costs of tailoring a product to its destination markets, each country's exports which is a differentiated product is unlikely to be costlessly substituted between foreign markets; they would rather be imperfectly substituted.

Following Bergstrand (1985), Bier and Bergstrand (2001) also allow for the cost of distribution, marketing and tailoring each countries' goods to importing countries markets into their model. Since differentiated products are not costlessly substituted between markets, this allows imperfect substitution across home and foreign markets. They also posit that models that allows for monopolistic competition, economies of scale, positive transportation costs and asymmetric country sizes would not yield unity relative prices. Thus, they concluded that the common assumptions of setting all prices to unity is not realistic as larger countries tend to have higher relative prices and wage rate levels.

They therefore, develop a model which assumes that, the optimising consumer maximises a CES utility function subject to a budget constraint in which the imported goods' prices reflect an iceberg transportation cost and advalorem tariff and goods are differentiated in line with the Dixit-Stiglitz preference. Maximising the constrained utility gives the import demand function for the destination country. On the firm's side, the representative firm in the exporting country is assumed to maximise profit subject to two technology constraints. The first is that, it faces both fixed costs and constant marginal costs. The second is that the existence of cost in distributing the products to each market makes the products to be imperfect substitute across domestic and foreign markets. This is captured by a CET function whose elasticity of transformation of production is greater than zero. The general equilibrium condition yields a gravity model which allows tariff and transport costs to be non-zero, prices to be non-unitary and the elasticity of transformation to be non-infinity.

C. Gravity Equation under Heckscher-Ohlin and Linder Theories

Further theoretical justifications revealed that the gravity equation can also be derived from other trade theories. Bergstrand (1989) derived the gravity equation using both the Heckscher-Ohlin (HO) and Linder trade models. He extended the microeconomics foundation of the generalized gravity equation in Bergstrand (1985) to include differences in relative factor endowment based on non-homothetic tastes in line with the Linder theory, and the factor proportion theory of international trade within the HO model of inter-industry trade and the Helpman-Krugman-Makursen models of intra-industry trade. From these, he developed a general equilibrium model of trade which now has two different products or industries that are produced using two factors of production - labour and capital which are assumed to be fixed in each country, such that each firm produces a uniquely differentiated product in a market

---

[5] The gravity equation termed the "generalized gravity equation" because it includes price terms.

characterised as a Chamberlinian[6] monopolistic competitive market. Assuming monopolistic competition in one industry and perfect competition in the second industry, he model demand by assuming that the utility maximising consumer maximises a nested Cobb Douglas CES-Stone-Geary utility function subject to an income constraint which results in a bilateral set of 'Armington-like' bilateral import demand function. The aggregate demand function in each country for the two differentiated products A and B relates bilateral trade flows to national incomes, prices and per capita income and implies that the income elasticity of demand for product A(B) will be greater(or less) than unity, if per capita income rises. In other words, the monopolistic competitive assumption does not impose the unity constraints.

On the supply side, each country has two differentiated industries, with profit maximising firms in each industry. Also, countries have identical CET production technology function, while distributing its output among both domestic and foreign markets according to this function. Thus, producing a uniquely differentiated product in a market characterised as a Chamberlinain monopolistic competitive market. The firm incurs fixed costs and constant marginal costs, and therefore realises internal increasing returns to scale in production. The equilibrium condition gives a set of reduced forms equation whose solution gives a generalized gravity equation, which includes exporters and importers incomes, exporter and importer per capita incomes and prices.

One important distinction between his work and past theoretical derivations is that the latter specified the gravity equation as a function of exporters and importers incomes multiplicatively, while ignoring exporter and importers per capita income or exporters and importers population. Bergstrand (1989) work was unique due to two reasons. One, he became the first person to fully attempt to integrate the gravity equation into the HO model (factor proportion theory of international trade). Two, he provides a theoretical foundation for the inclusion of exporter and importer per capita incomes, and exporters and importers income which is consistent with both traditional trade theories and new trade theories.

However, Bergstrand (1990) gave a formal theoretical justification to the empirical correlations found between the share of intra-industry trade among country pairs and their average levels of their gross domestic product, per capita income and tariffs. Extending the theoretical gravity model developed in Bergstrand (1989), he provide theoretical framework for these six determinants of the pattern and volume of bilateral intra-industry trade. He used the analytical framework in Bergstrand (1989), with the usual utility and production assumptions, but with the exception that the high income elasticity good (which was capital intensive in Bergstrand, 1989) does not need to be capital intensive in production.

In addition, he relaxes the assumption of two differentiated products/industries and two factors by assuming that one industry produces homogenous non-manufactured products under constant returns to scale and the other industry produces differentiated but symmetric manufactured commodities, but are imperfect substitutes in demand and are also differentiated by firms and country of origin. Maximizing the utility and profit functions, their analytical solution gives the gravity equation similar to Bergstrand (1989) but not identical to it as the number of firms is endogenous in the equation.

D. Gravity Equation under HO Model with both Friction and Frictionless Trade

In contrast to the Helpman and Krugman's thesis that the HO model is inconsistent with the multiplicative form of the gravity model, Deardoff (1998) also made a theoretical derivation of the gravity equation from the HO model of international trade within the Neoclassical framework.

---

[6] In a Chamberlinian monopolistic competition model, firms have monopoly power as products sold have no identical substitute. However, new firms entry drives profits to zero (Chamberlin, 1962).

Contrary to previous studies (e.g. Bergstrand[7] 1985; 1989; and 1990) that derived the gravity equation from a model that incorporate the monopolistic competitive market where products are differentiated by country of origin (according to the Armington assumption), Deardorff argued that the equation can also be derived using the HO model and with perfect competitive assumptions, where products differentiation and specialization occur due to non-factor price equalization among countries, rather than the Armington assumption. In addition, he shows that the gravity model is also consistent with several variants of the HO model and standard trade theories, and can therefore be derived as well as justified from them as the gravity equation seems to characterize a large set of models.

He derived the gravity equation assuming both frictionless trade and trade barriers. First, using a HO model which can incorporate any number of factors and goods that are homogenous, he assumes frictionless trade equilibrium. With frictionless trade, foreign trade is as cheap as domestic transaction as consumers face the same prices for the goods and consumers are indifferent between purchasing domestic and foreign goods, which are of equally priced sources of supply. Producers are also indifferent about the destination of their products sales. Conversely, consumption, production and net trade follows the maximization condition of perfect competitive markets, as they face the same prices as a result of the frictionless trade. Resolving this actual level of transaction yields a simple frictionless gravity equation which gives bilateral level of trade flows in which preferences are identical and homothetic.

Finally, he derived the gravity equation under the HO framework, allowing for trade impediments. Each country produces differentiated products and trade barriers exist for every good in form of transport costs which are strictly positive on all international transactions. In the presence of transport cost, factor prices are not equalized for each country and this allows non-factor price equalization between countries (factor price equalization version of the HO). Under further assumptions that there are many goods than there are factors of production, a gravity equation of bilateral trade flows with the Cobb Douglas and the CES preferences is derived.

E. Gravity Equation under Heckscher-Ohlin and Increasing Returns to Scale (IRS)

Drawing from the HO and IRS theories of international trade, Evenett and Keller (1998) gave further theoretical foundation to the gravity equation by determining whether they can actually account for the empirical success of the gravity equation. Two different versions of these two trade models can theoretically predict the general equation; these are the perfect specialization under HO and IRS models, and the imperfect specialization under HO and IRS models.

They imposed the assumption that both countries have identical production technologies and their consumers share identical homothetic preferences, both countries produce differentiated goods which are identically produce using increasing returns to scale. With no transport costs, they predicted the gravity equation with perfect specialization of production with both the HO and the IRS models. Furthermore, they also predicted a gravity equation with imperfect specialization of production for both the HO and IRS models under the assumption that the two goods are produced by two sectors; the first good is produced as a homogenous good under constant returns to scale, and it is more labour intensive in production. However, the second sector produces differentiated good using increasing returns to scale, and it is more capital intensive.

However, their findings revealed that: first, increasing returns to scale is an important cause of perfect product specialization and the gravity equation, and it is important in explaining the volume of North-North bilateral trade flow. Second, they however found no empirical support for the perfect product specialization HO model as little production is perfectly specialized as a result of differences in factor proportion, thus, making the perfect specialization version of the HO model unable to

---

[7] Bergstrand (1989; 1990) have a hybrid of one sector perfect competitive HO model and the second sector being a monopolistic competitive model.

explaining the empirical success of the gravity equation. Third, because production is not imperfectly specialized across countries due to differences in factor proportion, they find supports for both imperfect specialization versions of the HO and IRS models of trade in homogenous goods, and is said to be quite able to explain North-South trade.

F. Gravity Equation under Reciprocal Dumping Model

Feenstra, Markusen and Rose (1998) predict some other theoretical foundations for the gravity equation based on the argument that the empirical performance of the equation is specific to the type of good considered. They noted that the existing theories for the gravity equation derived depend on the assumption of differentiated goods which allows for product specialization in different goods. They further argued that while specialization might characterize manufacturing products (which explains why the gravity equation predicts the trade among industrialized (OECD) countries and empirically well), however, specialization is not a feature of homogenous primary products in which most developing countries trade in. Puzzled by the fact that most developing countries trade more in homogenous primary product, and that the gravity equation also work empirically well for these set of countries, they therefore show that the gravity equation can also arise from a wide range of models other than those specified before although they might generate subtle differences in the estimated coefficients. Consequently, they therefore derived the equation for both differentiated and homogenous products.

First, using models of product differentiation, they derive a theoretical gravity equation from both the monopolistic competition-like product differentiation and a country of origin (Armington) product differentiation, both of which yield subtle differences in the gravity equation. Theoretically, the monopolistic competition product differentiation gives a gravity equation that has larger domestic income elasticity of exports than importers income elasticity of export – known as the home market effect. The converse is the case in the model with an Armington product differentiation.

Second, using models that allow for homogenous products, they derived gravity equations from the reciprocal dumping model of international trade. However, it was shown that alternative conditions of firm entry for the reciprocal dumping model generate subtle differences in the gravity equation. Theoretically, the model with free entry predict a gravity equation in which domestic income elasticity of exports for homogenous products is larger than importers' income elasticity of export and the reverse is the case for a reciprocal dumping model with restricted firm entry.

G. Gravity Equation under Ricardian Model

Eaton and Kortum (2002) give theoretical foundation to the gravity equation using a Ricardian model of international trade that incorporates technology and geographic barriers[8] into a general equilibrium system of demand and supply. In contrast to previous studies, their model allows both geographical barriers and technology to determine specialization. The analytical solution of the general equilibrium model then gives a simple structural gravity equation which relates bilateral trade volumes, first, to deviations from the purchasing power parity and second, to technology and geographical variables. Technology creates comparative advantage and promotes trade, while the gains are attenuated due to geographical barriers.

H. Gravity Equation under an Incomplete Specialization Model

In contrast to the conventional way of deriving the gravity model, Haveman and Hummels (2004) show that the gravity equation can also be derived from a model with incomplete specialization and trade costs which is in sharp contrast to the early theoretical gravity equation that were derived from

---

[8] Their notion of geographical barriers includes tariffs, transport costs, quotas, delays, problems associated with deal negotiations from distance afar

general equilibrium models that assume identical preferences and that product are either differentiated by origin according to the Armington, which lead to complete specialization. Using a neoclassical trade model (HO) and allowing for incomplete specialization such that each homogenous product is produced by more than one country, they derive a gravity equation of bilateral trade flows that predict a much lower trade level and whose income elasticity of trade is similar to that of complete specialization model.

They posit that the predominance of zero bilateral trade flows is inconsistent with the assumption of complete specialization as was typically used in deriving gravity equation, but it is however consistent with the incomplete specialization assumption. The intuition is that models with complete specialization where each good is produced by one county implies that consumers highly value these goods, and therefore purchase every one of them. Therefore, there is no possibility for zero trade between the countries. In contrast, in incomplete specialization models, multiple countries produce each identical good, which gives no room for complete specialization. In addition, due to trade frictions, importers will tend to buy from only a small number of exporters, thus, there is the possibility for some bilateral trade flows to some countries to be zero, which strongly consistent with trade data (Haveman and Hummels, 2004).

2.1.2 Recent Development in the Theoretical Foundation

After more than two decades of an influx of models providing theoretical justification for the empirical success of the gravity equation, emphasis thereafter turned to ensuring that the empirical results of the gravity equation is well defined on theoretical grounds. One important contribution in this regard relates to the structural form of the equation and the implication of misspecification or omitted variable bias. These relate to way trade costs and firm heterogeneous behavior is incorporated into the gravity equation. The work of Anderson and vanWincoop (2001 or 2003) and Helpman, Metliz and Rubeinstein (2008), etc are deemed to be influential here.

Modeling Trade Costs - Multilateral Trade Resistance

The concept of multilateral trade resistance cost was discovered by Anderon and van Wincoop (2001) in his seminar paper following the controversial study by McCallum (1995) who find that in 1988, US-Canadian border led to a trade between Canadian provinces which is 22 (2200%) times more than trade between the US states and the Canadian provinces. This is termed the '*border puzzle*'[9] or a home bias in trade, which makes it one of the six puzzles of open macroeconomics (Obstfeld and Kenneth Rogoff, 2001).

Motivated by the resulting border puzzle of McCallum (1995), Anderson and van Wincoop (2001, 2003) gave the gravity model a new theoretical underpinning to explain and solve this border puzzle by incorporating the multilateral resistance term. They posit that McCallum's ratio of inter-provisional trade to province-state trade is very large because of omitted variables bias, (multilateral resistance terms term) and the small size of the Canadian economy. They however got a smaller border effects than in McCallum (1995) after controlling for multilateral trade resistance in their regression model.

Extending Anderson 1979 theoretical derivation, they derive that economic distance between countries *i* and *j* is not only determined by a bilateral resistance term between these two countries as

---

[9] Border puzzle is the tendency for a country to trade with and buy domestic products originating from domestic home country - a strong preference or bias for domestic goods. This phenomenon is termed border puzzle by McCallum (1995) and arises because countries borders are supposed to have a significant effect on the trade patterns between the countries especially if the countries are similar in terms of same language, culture and economic institutions as in the case of the US and Canada. However, the estimated patterns of trade indicates strong inter-provincial trade and less province-state (international trade) between Canada and the US, implying that national borders constraint trade among countries even though the countries are similar to one another (McCallum, 1995).

shown by previous derivations, but also in relation to a weighted average of economic distance to all other trading partners of the given country. The latter is what they termed the multilateral resistance term, and theoretically appropriate average trade barrier.

They employ a monopolistic competition framework which is built on the Armington assumption that each country produce differentiated goods and trade is therefore driven by consumers' love for varieties such that all domestic and foreign goods are imported by the variety loving consumers. Optimizing consumers; preferences across countries and this is captured by CES preference. Goods are also assumed to be differentiated by region of origin such that each country specialises in the production of only good which is fixed in supply; and all goods produced by both domestic and foreign firms are consumed by the variety loving consumers. A key feature of the model is the introduction of exogenous bilateral trade costs into the gravity model. This incorporation of trade costs, which are directly observable, ensures that prices of the goods can differ across countries, and non-price equalisation implies that elasticity of substitution across products is non-unitary which is in contrast to Anderson (1979) that assumes a unitary elasticity of substitution.

The equilibrium condition results in a general equilibrium model and assuming trade barriers are symmetric and imposing a market clearing condition, yields a micro-founded gravity equation which relates bilateral trade flows to size and trade costs where the trade costs are decomposed into 3 components: the bilateral trade barriers between exporting country *i* and importing countries *j*; exporting country'*s* resistance to trade with all countries (outward multilateral resistance); and importing country*'s* resistance to trade with all countries (inward multilateral resistance).The resulting micro-founded gravity equation then relates bilateral trade flows to country's size, bilateral trade barriers and multilateral trade resistance variables. Specifically, it predicts that bilateral trade flow is explained by income of exporters and importers, an elasticity of substitution across goods which is greater than unity, bilateral trade costs, and exporters and importers prices indices which they termed multilateral trade resistance term (ratio of outward to inward multilateral trade resistance) also known as relative trade term or average trade costs.

Sequence to Anderson and van Wincoop (2001) influential seminar paper, Feenstra (2002) also noted the exaggerated and biased estimate of the Canada-US border effects in McCallum (1995). To avoid this bias, he re-derived the gravity equation allowing for trade barriers (such as tariff and transport costs) across countries such that they have different prices. He therefore deviated from the conventional gravity equation (like that of McCallum, 1995), which did not incorporate price indexes, which have the effect of overstating the border effect for Canada and understating it for the US. According to Feenstra (2002), with the introduction of border effects (tariffs and transport costs) price equalization across countries no longer holds.

Following Anderson (1979), Feenstra also derived a gravity equation from a monopolistic competitive model in which consumers face CES utility function. To allow for non-factor price equalization across countries, he made a further assumption. Each country is assumed to produce unique product varieties with the products exported by the exporting country selling for the same price in the foreign importing country, where these prices are sold in importing market inclusive of transport costs while prices in exporting countries are exclusive of any transport costs (fob). Optimizing the utility of the representative consumer in destination countries and solving the equation further gives a gravity equation which relates total bilateral trade values to aggregate income in destination country, number of products, relative price index of each country and elasticity of substitution factor.

More recently, Novy (2011) also derived a gravity equation which incorporates multilateral trade resistance. Building on Anderson and van Wincoop (2003) gravity framework, he derived an analytical solution for the multilateral trade resistance (both time varying and observable multilateral resistance variables) from which bilateral trade costs can be directly predicted. He noted that there are some drawbacks in Anderson and van Wincoop (2003) assumptions used in solving the multilateral resistance terms, as they abstract strongly from reality. For instance, they assume bilateral trade costs to be function of two trade costs proxies – bilateral geographical distance and a border barrier, he further assume that these bilateral trade costs are symmetric for country-pairs. He noted that drawbacks arise, first, because there is the possibility of trade cost function being mis-specified as it omits an important trade cost – tariff; and secondly, trade costs might turn out to be asymmetric as counties impose higher tariffs than others. Novy, therefore, overcome these drawbacks by deriving an analytical solution for the multilateral resistance variables using a method that neither imposes symmetric trade costs nor any particular trade cost function. This gives a micro-founded gravity equation which allows for unobservable trade costs.

New 'New' Trade Theory

Another major area of new contribution relates to the methodological issue associated with the presence and behavior of heterogeneous firms operating in international markets which were spearheaded by Melitz (2003) and Bernard et al., (2003). Firm heterogeneity arises since not all existing firms in a country exports; only a minority of these firms participate in international market (Bernard et al, 2003; Mayer and Ottaviano, 2008). Furthermore, not all exporting firms export to all the countries in the rest of the world; they are only active in just a subset of countries and may choose not to sell specific products to specific markets (or their inability to do so). The reason for the heterogeneity in firm behavior is because fixed costs are market specific and higher for international trade than for domestic markets. Thus, only the most productive firms are able to cover these costs, and firms' inability to exports may be due to the high cost involved. Consequently, the bilateral trade flows matrix will not be full as many cells will have zero entries. This case is seen at the aggregated level of bilateral trade flows but more often in greater levels of product data disaggregation such as HS6 and HS8.

The prevalence of zero bilateral trade flows has important implication for modeling the gravity equation as zero trade between several country-pairs might signal a selection bias problem. In addition, the observed zeros might contain important information about the countries (such as why they are not trading) which should be exploited for efficient estimation. Thus, more recent waves of theoretical contribution relate to deriving the gravity equation that allows for firm heterogeneity into the equation and the development of an influx of estimation techniques that would take care of the zero trade records.

Standard gravity equation usually neglect the issue of the prevalence of zero bilateral trade flows and predict theory consistent with only positive bilateral trade flows. However, Helpman, Melitz and Rubinstein (2008); Novy (2011, 2012), etc derived theoretical gravity equation which highlight the presences of zero trade records and gives theoretical interpretations for them. The new new trade model of international trade with firm heterogeneity which is spear-headed by Metlitz (2003) is usually adopted in giving the gravity equation theoretical basis which is elaborated below.

Helpman et al. (2008) argue that "by disregarding countries that do not trade with each other, these studies give up important information contained in the data" (Helpman et al. 2008 p442), and that symmetric relationship imposed by the standard gravity model biases the estimates as it is inconsistent with the data. To correct for this bias, Helpman et al. (2008) provides a theoretical gravity equation that incorporates firm heterogeneity and positive asymmetric and was thus, able to predict both positive and zero trade flows between country-pairs. Given firm level heterogeneity, they assume products are differentiated and firms are faced with both fixed and variable costs of exporting. Firms vary by productivity, such that only the more productive firms find it profitable to export; with the profitability of exports varying by destination. Since not all firms found it profitable, this gives rise to positive and zero trade flows across country-pairs. Furthermore, this difference in productivity gives rise to asymmetric positive trade flows in both directions for some pairs of countries. These positive asymmetric trade and zero bilateral trade flows then determine the extensive margin of trade flows (number of prospective firms). Moreover, given that firms in country '*j*' are not productive enough to enable them profitably export to country i, this implies that, there will be zero trade flows from country *j* to *i* for some pairs of countries. This generates a model of firm heterogeneity that predicts zero trade flow from countries j to i but positive exports from country *i* to *j* for some pairs of countries, and zero bilateral trade flows between countries in both direction.

Sequent to Helpman et al. (2008), others have also derived the gravity equation allowing for firm heterogeneity (c.f. Chaney, 2008; Melitz and Ottaviano, 2008; Chen and Novy 2011).For instance, Chaney (2008) derives an industry level gravity equation using a model that assume firm level heterogeneous productivity across firms and fixed costs of exporting. Chen and Novy (2011) however argued that apart from variations in trade costs across industries, industry specific elasticities of substitution are also important in capturing the cross industry variations. So they derive a model that allows for both industry specific bilateral trade costs and industry specific elasticities of substitution. Employing the monopolistic competition framework used in Anderson and van Wincoop (2004) that allows for only heterogeneous cross country trade costs, they also included heterogeneous elasticities of substitution across industries in the model, and generate a micro-founded gravity equation of bilateral trade flows that controls for cross industry heterogeneity but nets out multilateral resistance terms.

Chen (2012) deviate from the standard gravity equation that assumes CES, in which trade costs have similar effects across country-pair, which gives rise to gravity equations with constant elasticity of trade with respect to trade costs. This implies that ceteris paribus, a change in trade cost has similar proportionate effect on bilateral trade flows irrespective of whether the tariffs faced by the countries were initially low or high, or whether a given country pair traded a lot or little. He justified that in reality, trade costs have heterogeneous trade impeding impact across countries as the effect on trade flows depend on how intensive pairs trade with each other. The trade flows of exporting countries that provide only a small portion of the destination country’s total import is more sensitive to bilateral trade costs. Likewise, trade is more sensitive to bilateral trade flows for countries that import very little from a given exporter. Consequently, trade costs might have a heterogeneous impact across country-pairs, with some trade flows being zero. Based on this justification, he then use the translog gravity equation in which trade costs have a heterogeneous trade impeding effect across country-pairs, which is also consistent with zero trade demand.

Recently, Chaney (2014) validated the role of economic size in gravity model and confirms the size elasticity of trade to be approximately 1, as often seen in most gravity literature. Beyond this confirmation, he gave new evidence that larger firms export over longer distances than small ones. This theoretical paper first explains the reason behind the fact that the size elasticity of trade is approximately 1, while the distance elasticity is -1. Although, he gave no new evidence or reason for the size elasticity besides the conventional trade model, but his explanation of the distance elasticity is new. This was done by showing that if the distribution firm sizes is Pareto, and if the average squared distance of a firm’s exports is an increasing power function of its size, then the distance elasticity of trade is constant, and equal to-1 in the special case of Zipf’s law. Second, the paper built a model that is micro-founded where the distribution of firm size is Pareto, while the average squared distance elasticity of the firms’ exports is a power function of its size, such that the gravity equation emerges endogenously. These firms were geographical distributed in the model such that the theory assumes that the firms combine, produce and trade in intermediate inputs. Given the fact that inputs are imperfect substitutes that are combine in a CES production function, firms have the incentives to acquire more upstream suppliers. Also, since consumers’ value differentiated goods produced by firms, this gave firms the incentives to acquire more downstream consumers. Thus, assuming that the information about the potential suppliers and consumers is costly, which is acquired overtime; the firms gradually built network of suppliers and consumers spanning increasingly long distances. This generates an invariant Pareto distribution of firm sizes with larger firms shipping their export over longer distances. The two predictions of a Pareto distribution of a firm sizes and a distance of exports that increases with firm size, generates a constant distance elasticity of firms’ trade that is equal to -1 when the distribution firm sizes conform with Zipf’s law.

In sum, gravity equation can arise from a wide range of trade models both standard, new and new new trade theories. They are usually offered as theoretical substitutes and the choice of the equation depend on the preferred set of assumptions and models (Bier and Bergstrand, 2001). Nonetheless, there are some differences in the underlying assumptions and models and such differences could probably explain the various specifications in the literature and the diversity in the empirical results (Martinez-Zarzoso and Nowak-Lehmann, 2002). While the theoretical basis is no longer in doubt, emphasis is now on ensuring that its empirical applications is well rooted on its theoretical ground and that it can be linked to anyone of the available and appropriate theoretical frameworks. However, irrespective of the theoretical framework adopted, most of the subsequent justifications of the gravity equations are variants of the one first derived in Anderson (1979).

3.0 The Gravity Model Estimation Debate

A review of the literature indicates that gravity model is often use in explaining bilateral, regional and multilateral relations. First, this is due to the rigorous theoretical foundation given to it with the advent of trade theories, especially the new trade theory. Second and more important, this is due to its empirical success in the analysis of bilateral relations. However, in spite of the popularity it enjoys, there are still questions about the appropriate specification of the model and estimation technique(s) to use. Here, we shed light on the specification and estimation techniques issues involved in gravity modeling. Particular attention is focused on the GMED as it concern the merit and demerit of each techniques in the presence of zero trade flows that occurred prominently due to the disaggregated dataset in which over 50% of trade values are found to be zero.

3.1 The Debate

Early empirical studies rely on cross sectional data to estimate the gravity model, in which the economic framework for the model was cross-sectional analysis, (c.f. Anderson, 1979; Bergstrand, 1985, 1989; McCallum, 1995; and Deardorff, 1998; etcetera). For such cross-sectional analysis, the ordinary least square (OLS) estimation technique or pooled OLS technique is normally employed. However, the traditional cross-sectional approach is affected by severe misspecification problems and thus, previous estimates are likely to be unreliable (Carrerè, 2006). This is because, the traditional cross sectional gravity model usually include time invariant variables (e.g. distance, common language, historical and cultural dummies, border effects), but the model suffers from misspecification problems as it fail to account for country specific time invariant unobservable effects. This unobservable country specific time invariant determinants of trade are therefore captured by the error term. These unobserved variables are likely to be correlated with observed regressors and since OLS technique is usually used, this renders the least square estimator to be inconsistent, which makes one of its classical assumptions invalid. In addition, OLS does not control for heterogeneity among the individual countries, which has the potential of resulting into estimation bias as the estimated parameters may vary depending on the countries considered. Therefore, estimating cross sectional formulation without the inclusion of these country specific unobservable effects gives a bias estimate of the intended effects on trade. This renders the conclusions on cross sectional based trade estimates problematic *(ibid)*.

Thus, over the last decade, there is the increasing use of panel data in gravity modeling and the use of panel econometric methods (c.f. Egger, 2000; Rose and van Wincoop, 2001; Baltagi, 2003; Egger and Pfaffermayr, 2003, 2004; Melitz, 2007; and many others). The panel specification is much more adequate as the extra time series data points gives more degree of freedom, results in more accurate estimates. A unique advantage of panel data is that the panel framework allows the modeling of the evolvement of variables through time and space which helps in controlling for omitted variables in form of unobserved heterogeneity, which if not accounted for can cause omitted variable bias (Baltagi, 2008). In addition, with panel data, the time invariant unobserved trade effects can easily be modeled by including country specific effects such as time dummies, and thus avoiding the consistency issue mentioned above.

With the availability of panel data, the two common techniques used in fitting the data are the fixed effects and random effect estimation techniques, where the choice between the two hinges on their apriori assumptions. The fixed effect assumes that the unobserved heterogeneity is correlated with the error term. In contrast, the random effect assumes that the unobserved heterogeneity is strictly exogenous i.e. it does not impose any correlation between the unobserved heterogeneity (individual effects) and the regressors. Under the null hypothesis of zero correlation, the random effect model is efficient; both models are consistent, but the random model is more consistent. If however, the null hypothesis is rejected, the fixed effect is consistent and the random effect is neither consistent nor efficient. There are however, some drawbacks in the fixed effect model in the sense that all time invariant explanatory variables (are deem to be perfectly collinear with the fixed effects) would be dropped from the model. Consequently, fixed effect model eliminates some important theoretically relevant variables from the gravity equation which are distance, common language, common borders, and the effects of these variables cannot be established. In addition, studies have also applied the OLS technique to panel data. However, pooled OLS can only give precise estimators and test statistics with more power if the relationship between the dependent variable and the regressors remain constant over time.

Early gravity model estimation technique was to estimate the equation by least squares, where the model is usually log linearized as a common practice. Their position is that for the validity of a log-linear gravity model hinges on the homoscedastic assumption, as the error term must be statistically independent of the regressors. However, in recent times, Santos Silva and Tenreyro, (2006) have identified flaws with this practice. Their position is that due to the nature of trade data that are intrinsic to heteroscedasticity and pervasive zero trade observation, log linearizing the gravity equation and then applying OLS is problematic.

First, problems arise in logarithmic transformation due to heteroscedasticity, which is usually present in trade data. As noted by Santos Silva and Tenreyro (2006) in their influential paper, the common practice of log linearizing the gravity equation and then estimating using OLS is inappropriate because, the expected values of the log linearized error term will depend on the covariates of the regression, and hence, OLS will be inconsistent even if all observations of the dependent variables are strictly positive. This is because logarithmic transformation of the gravity model changes the property of the error term. In other words, OLS will produce consistent estimates as long as the error term (ε_{ij}) of the log linear specification $(\ln \varepsilon_{ij})$ is a linear function of the regressors, i.e., if $E[\ln(\varepsilon_{ijt} \mid x_{ijt})] = 0$, which is the homoscedasticity assumption. However, logarithmic transformation generates estimates of $E(\ln \varepsilon_{ij})$ and not $\ln E(\varepsilon_{ij})$, but, where $\ln E(\varepsilon_{ijt} \mid x_{ijt}) = 0; E(\ln \varepsilon_{ijt} \mid x_{ijt}) \neq 0$, which is the well-known Jensen’s inequality[10]. ’

Consequently, due to Jansen’s inequality, the error term (ε_{ijt}) is not equal to the log of the error term $(\ln \varepsilon_{ij})$ as the error terms in the log linear specification of the gravity equation are not statistically independent of the regressors but are rather heteroskedastic, leading to inconsistent estimates of the elasticity coefficients. Given this Jansen’s inequality, Santos Silva and Tenreyro (2006) argue that the log linear transformation of the gravity model is intrinsic to heteroscedasticity.

---

[10] Jensen’s inequality is named after Johan Jensen, the Danish mathematician who in 1906 discovered that: the secant line of all convex function (i.e., the means of the convex function) lies above graph of the function (i.e., the convex function of the weighted means) at every point. The reverse is true for a concave function. His inequality has appeared in many contexts and an example in this case is the arithmetic mean inequality. Thus, in simplified terms, his inequality states that the convex (or concave) transformation of a mean is less or equal to (greater or equal to) to the mean after a convex (concave) transformation. Thereafter, Economists have adopted his intuition to show that the logarithm transformation of an equation generates the expected value (mean) of the logarithmic transformation of the dependent variable $E(\ln Y_{ij})$ and not the logarithm of the mean of the dependent variable $\ln E(Y_{ij})$; and $E(\ln Y_{ij}) \neq \ln E(Y_{ij})$.

Thus, applying OLS results into biased and inefficient estimates. They argue that even though, economists have long known about Jensen's inequality and that the concavity of the logarithm function could create a download bias when employing OLS, this important drawback has, however, been overlooked in bilateral trade studies. They confirm their argument as they found evidence of the presence of heteroskedastisity and inconsistency in the normal log-linear representation of the gravity model; which renders the estimates of elasticity obtained from least squares estimation technique to be both inefficient and inconsistent.

Second and more importantly is the presence of zero trade flows in the trade matrix and the appropriate estimation technique. While the Newtonian gravity theory from which the gravity model of trade was derived allows for very small gravitational force, but not zero force, however, in trade, there are frequent occurrences of zero[11] valued bilateral trade flows. The practice of estimating the log linear gravity model in the presence of such zero trade flows implies both theoretical and methodological problems; especially in cases where the presence of such zero values are excessive. In estimating the gravity model, the gravity model is log linearized and estimated using these linear regression techniques. However, given the predominance of zero trade records in the trade matrix, particularly at the more disaggregated level, where zero records can account for about 50% of trade flows, the logarithm transformation of the dependent variable is therefore problematic. This is so because the logarithm of zero is indeterminate or not feasible.

The common practice in the literature employed to deal with the problem of zero records in the data are the truncation and censoring methods and thereafter applying linear estimation techniques. In the case of truncation method, the zero valued trade flows are dropped completely from the trade matrix, whereas, the censoring method involves substituting the zeros by a small positive arbitrary value. These methods are however, arbitrary and are without any strong theoretical or empirical justification and can distort the results significantly, leading to inconsistent estimates (c.f. Flowerdew and Aitkin, 1982; Eichengreen and Irwin, 1998; Linders and de Groot, 2006; Burger et al., 2009; Gomez-Hefrera, 2011). In addition, Flowerdew and Aitkin[12] (1982) show that the results are sensitive to (small) differences in the constant substituted, which can cause serious distortion in the results. Eichengreen and Irwin (1998) noted that deleting these zero values led to loss of information as important information on the zero trade levels is left out of the model and this can generate biased results if the zero trade flows are not randomly distributed; while Heckman (1979), Helpman et. al, (2008) posit that omitting these zero trade records can result into sample selection bias. The loss of information is said to reduce efficiency and omission of data produces biased estimates (Xiong and Beghin, 2011; Gomez-Herrera, 2011). In addition, Xiong and Beghin (2011) noted that deleting the zero trade observations prevents the possibility of exploring the extensive margin of trade – the creation of new bilateral trade relations. This implies that the estimates are conditioned on trade that already took place – the intensive margin of trade. They concur that ignoring zeros limits the economic interpretation of the model as nothing can be said on the implication for new trade.

Likewise, Linder and de Groot (2006) kicked against truncating and censoring trade data by arguing that, zero trade observation may provide important information for understanding the bilateral trade patterns and therefore should not be eliminated apriori. Disregarding the zeros trade flows can bias the results if they do not randomly occur. This is because, zero trade flows provided information about the probability to engage in bilateral trade. thus, if distance, low levels of GDP, the lack of historical or cultural links, etcetera, make trade to be non-profitable, thereby reducing trade or bringing about no trade, then eliminating zero flows from the analysis is tantamount to sample

---

[11] Frankel (1997) argued that these zero values arises as a result of lack of trade between countries, or from rounding errors when trade between countries does not reach a minimum value or can arise when they are rounded-down as zero, it can also results from measurement errors where observations are mistakenly recorded as zeros.

[12] They vary the substituted constant between 0.01 and 1 and found that the regression coefficient decreases with the size of the chosen constant.

selection bias and applying OLS will lead to underestimating of the gravity equation coefficients (downward bias).

Therefore, in recent years, attention has been on the appropriateness of the estimation technique especially those relating to the problems of zero trade costs and logarithmic transformation of the gravity equation, and the constant emphasis on the inappropriateness of linear estimators in taking care of these two problems. Consequently, more appropriate estimation techniques are being increasingly employed to deal with these two issues in the context of gravity trade literature. The Tobit and Probit models, truncated regression, Poisson and modified Poisson models, Nonlinear Least Square (NLS), Feasible Generalized Least Square (FGLS) and the Helpman, Melitz and Rubinstein (2008) approach have all been used to deal with the problem associated with log normal formulation and the excessive zero valued trade flows.

Early studies have relied on the Tobit model to deal with the zero trade problems. For instance, the Tobit model has been employed by Rose (2004) and Andersen and Marcoiller (2002) to deal with the problem of zero valued trade flows that resulted either because the actual trade flows are not observable or due to measurement errors from rounding. The Tobit estimator is applied to fit dataset when outcome/data are only observable over some range. It is applied in cases of measurement errors (e.g rounding up) or when actual outcomes cannot seem to reflect the desired outcomes. The Tobit censoring method involves rounding (censoring) part of the observation to zero or rounding up the zero trade flows below some positive value.

Nevertheless, Linder and de Groot, (2006) has debated on the appropriateness of using the Tobit model to fit zero valued trade flows in a gravity model. However, the fitness of Tobit model will depend on whether the desired trade could be negative or whether rounding up of trade flows is important. Their argument is that in the gravity model, the zero trade flows cannot be censored at zero as the desired trade cannot be negative in the gravity equation; this can only occur if the GDP of one or country pair is equal to zero which is unlikely in real life. They further argue that censoring at a positive value is not also appropriate. The intuition is that the UN COMTRADE data reports trade values, even for very small values (up to $1), indicating that rounding to zeros is not an important cause of zero observation as most zeros are caused by economic reasons such as lack of profitability .This implies that zero trade flows is likely to occur from binary decision making about the profitability of engaging in trade, and not from rounding up (censoring), thus the model might not be appropriate for taking care of zero trade flows. In addition, Frankel (1977) and Rose (2000) noted that the Tobit estimator involves an artificial censoring of positive albeit small trade values, however, the trade flow is subject to measurement errors, and they may have a high influence on the regression results.

Furthermore, Martin and Pham (2008) show that, although both truncated OLS and censored Tobit model lead to bias results but the censored method generally produced much worse results in comparison to the truncated method, and suggested that Eaton and Tamura (1994) threshold Tobit model gives the lowest bias and outperform all other estimators in a simulation exercise. However, in contrast, in a simulation exercise, Santos Silva and Tenreyro (2011) found the Tobit model of Eaton and Tamura (1994) has large bias, which increases with sample size, which also confirm its inconsistency as an estimator.

Attention has also been shifted to the use of the Poisson and the modified Poisson specifications of the gravity model. Santos Silva and Tenreyro (2006; 2011) used the Poission Psuedo Maximum Likelihood (PPML) method to deal with the zero valued trade flow and the logarithm transformation. According to them, in the presence of zero valued observations and also due to the logarithm transformation of the gravity equation, OLS (both truncated and censored OLS) are inconsistent and have very large bias which do not vanish as the sample size increase which confirm that they are inconsistent (Santos Silva and Tenreyro 2011). However, the PPML estimates the gravity equation in levels instead of taking its logarithms and this is said to avoid the problem posed by using OLS under logarithm transformation. According to them, this model is appropriate: first, the Poisson model takes account of observed heterogeneity. Second, the fixed effects PPML estimation technique gives a natural way to deal with zero valued trade flows because of its multiplicative form. Third, the method also avoids the under-prediction of large trade volumes and flows by generating estimates of trade flows and not the log of the trade flows. In their 2006 influential paper, they find the PPML estimator, which need not be log-linearized, to be the best performing estimator that naturally deal with zero trade flows, consistent and gives the lowest bias among the other estimators. They therefore suggest it as the new workhorse for the estimation of the typical constant elasticity models, such as the gravity model.

However, their influential paper has however generated some controversies in the literature (c.f. Martinez 2007; Martin and Pham 2008; Burger et al., 2009; etcetera). For instance, Burger et al. (2009) identified some important limitations of the PPML model. They noted that the model is vulnerable to the problem of overdispersion in the dependent variable and excess zero flows. They posit that the model only takes account of observed heterogeneity and not unobserved ones and this is an important limitation of the PPML model. While an important condition of the PPML is the assumption of equidispersion (the conditional variance is equal to the conditional mean) in the dependent variable, however, due to the presence of unobserved heterogeneity which are not accounted for in the model, there is an over-dispersion in the trade flows (dependent variable). The over-dispersion is said to generate consistent but inefficient estimates of trade flow (Burger, et al. 2009; Turkson, 2010).

Contrary to Burger et al. (2009) who noted that the model is vulnerable to the problem of overdispersion in the dependent variable and excess zero flows, which generate consistent but inefficient trade estimates, Santos Sliver and Tenreyro (2011), find that PPML is consistent and generally well-behaved even in the presence of overdispersion in the dependent variable (i.e. when the conditional variance is not equal to the conditional mean). Also, the predominance of large proportion of zeros does not affect its performance. In addition, Soren and Bruemmer (2012) find that the PPML performs quite well under over-dispersion, and show that the PPML is well-behaved under bimodal distributed trade data.

Nonetheless, attempts have also been made to correct for the over-dispersion in the dependent variable and the vulnerability of the PPML to excessive zero flows using other estimation techniques apart from the PPML. These are the Negative Binomial Pseudo Maximum Likelihood (NBPML) and the Zero-inflated models which are Zero-inflated Pseudo Maximum Likelihood technique (ZIPML) and Zero-inflated Binomial Pseudo Maximum Likelihood technique (NIBPML) (Burger et al. 2009). They posit that the NBPML corrects for the overdispersion the estimator incorporates unobserved heterogeneity into the conditional mean and thus, takes care of unobserved heterogeneity. However, an important drawback of the NBPML and PPML relates to the excessive number of zero in the observation which means that the number of zero flows is greater than what the models predicts; where excessive zeros is said to be derived from the 'non-Poissoness'[13] of the model (Johnson and Kotz, 1969). Thus, Burger et al. (2009) posit that even though the Poisson model and the NBPML model can technically handle with zero flows, both models are however not well suited to handle cases where the number of observed zero valued trade flows is greater than the number of zeros predicted by the model.

They posit that the zero inflated models (ZIPPML and ZINBPML) perform better and correct for excess zeros and overdispersion in the dependent variable. They also noted that zero-inflated models have an added advantage as they theoretically well suited in modeling the origin of zero counts because the models account for two different types of zero trade flows, which are countries that have

---

[13] Burger et al. (2009) identified that one important cause of non-Poissoness is when some zeros in the observation are produced by a different process compared to the remaining observations (including some other zeros)

never trade (the non-poisson group), implying a data that strictly have zero counts; and countries that presently do not trade but potentially could, i.e. those that have a non-zero probability of having non-zero counts (the poisson group[14]). Thus, these models make allowances for the possibility to separate the probability to trade from trade volume as it provides additional information on the causes of the probability of the different kinds of zero valued flows. Given these, Turkson (2011) argued that the choice of the model to use will depend on whether the sample has excessive zero trade flow or not. However, Burger et al. (2009) posit that the Poisson model and the NBPML model are not well suited to handle cases where the number of observed zero valued trade flows is greater than the number of zeros predicted by the model.

Contrary to Burger et al. (2009), Staub and Winkelmann (2012) however, find that the PPML is consistent even when zeros are excessive. They also show that both ZIPPML and ZINBPML are inconsistent if the underlying assumptions of the distribution of model are violated, i.e. if the models are misspecified. They instead recommend the use of zero inflated Poisson Quasi Likelihood (PQL) estimator which was shown to be consistent in the presence of excessive zeros and it is unaffected by unobserved heterogeneity and found to be robust to misspecification as it consistently estimate the regression coefficients irrespective of the true distribution of the counts while ZIPPML and ZINBPPML demonstrate considerable bias in medium sample. They also noted that the PQL can be less efficient compared to zero inflated estimators if the zero inflated model is correctly specified.

Similar to Burger et al., (2009), Martinez-Zarzoso (2013) also find out that the PPML estimator proposed by Santos Silva and Tenreyro (2006) is not always the best estimator as its estimates are outperformed by both the OLS and FGLS estimates in out of sample forecast. In addition, the PPML assumption regarding the pattern of heteroscedasticity is rejected by the data in most cases. However Santos Silva and Tenreyro (2008) responded by justifying the use of PPML as the best estimator in the context of gravity model, but also acknowledged that PPML estimator can be outperformed by other estimators in some cases.

Furthermore, Martinez-Zarzoso et al. (2007)[15] also finds the PPML to be outperformed by both the OLS and FGLS estimates in out of sample forecast and deduced that it is not always the best estimator. She finds that PPML assumption regarding the pattern of heteroscedasticity is rejected by the data in most cases. She opined that even in the presence of unknown form of heteroscedasticity, FGLS can still be applied, because as FGLS is an efficient estimator within the class of least squared estimators, but the variance of the disturbances should then be re-estimated to correct for heteroscedasticity errors. They pointed out that FGLS is well suited to estimating parameters in the presence of heteroscedasticity, so, the comparison[16] of the best performing estimator should be between FGLS and the class of generalized linear models[17] (GLM) such as the Non-linear least square (NLS), Gamma Poisson Maximum Likelihood (GPML), and PPML. However Santos Silva and Tenreyro (2008) in their response, provided justification for the PPML estimator in the context of log linear gravity model, and acknowledged the fact that in some specific situations, the PPML estimator can be outperformed by other estimators.

---

[14] The zero inflated models consider two different groups within the population: the poisson group and the non-poisson group. The non-poisson group are countries which have strict zero probability of trading but do not trade at all. The non-poisson zeros might be caused by lack of trade due to bans or other trade embargoes or simply the lack of resources. The poisson group consist of those countries with non-zero probability to trade and are actually trading,m and countries that have non-zero probability to trade but however do not trade. The poisson zeros might be caused by huge distance or large differences in country pairs preferences and specialization.

[15] This influential discussion paper by Martinez-Zarzoso, Nowak-Lehman and Vollmer (2007) was later published in Martinez-Zarzoso (2013)

[16] Santos Sliver and Tenreyro (2006) paper have majorly centred on comparing OLS to the class of GLS, particularly PPML

[17] Generalized linear models are class of multiplicative models.

Martinez-Zarzoso (2013) compares the performance of different estimators via a Monte Carlo simulation exercise and find that although PPML to be less affected by heteroscedasticity compared to FGLS, NLS and GPML, nonetheless, its performance is found to be similar both in terms of bias and standard errors to the performance of the FGLS estimator. Particularly for small sample size; with the lowest bias and standard errors found in the GPML in the simulations which has non-zero values in the dependent variable. Further empirical analysis using three different real datasets[18] reveal that the choice of the performance of the model is sensitive to the sample size; for small sample size, FGLS could be perfect way to deal with the heteroscedasticity problem, while the PPML will be appropriate when the sample size is large and there is measurement error in the dependent variable. However, for large sample size, PPML bias is found to decrease in large sample size while FGLS bias is found to remain almost constant. In addition, the PPML standard error falls considerably, but it remains twice the FGLS standard errors. Conclusively, Martinez-Zarzoso (2013) find that the choice of the best estimator is dependent on the specific dataset, and there is no generally best estimator for these three datasets; thus the appropriate estimator for any application is data specific, which could be determined using a number of model selection tests.

Martin and Pham (2008) has also challenge Santos Sliver and Tenreyro (2006) findings and posit that, although the PPML estimator is less subject to bias resulting from heteroscedasticity problem, however, it is not robust to the joint problems of zero trade flows and heteroscedasticity. Based on this, they conclude that the estimator could be appropriate for other multiplicative models[19] which have relatively few zero observations. They proposed that the Eaton and Tamura (1994) threshold Tobit model perform better than the PPML and other estimators considered as it recorded the smallest bias in a simulation exercise.

The Monte Carlo simulation done by Santos Sliver and Tenreyro (2006), has also generated some debates. Although the authors find that the PPML is able to deal with zero trade flows, interestingly, their simulation done in order to determine the best performing model were without any zeros, except where the dependent variable was contaminated with measurement errors. This has made some studies to question the performance of the PPML in cases where there are excessive zeros in the dependent variable (c.f. Martinez-Zarzoso, 2013; Martin and Pham, 2008). Martin and Pham (2008) therefore used a data generation process[20] different from that used by Santos Sliver and Tenreyro (2006), which include a high proportion of zero values and show PPML to be highly vulnerable to bias in the presence of high percentage of zero values in the dependent variable. Similar result has been found by Martinez-Zarzoso (2013). However these results have been challenged by Santos Silva and Tenreyro (2011).

In response to these studies, Santos Sliver and Tenreyro (2011), argued that both of the simulations done by Martinez-Zarzoso et al (2007) and Martin and Pham (2008) reveal no information on the performance of the PPML model of constant elasticity model as the data used in their simulation exercises are not generated by a constant elasticity model. Santos Sliver and Tenreyro (2011), however, further investigate the performance of the PPML estimator when the dependent variable has large percentage of zeros and when the data generating process is given by a constant elasticity model (both of which are typical in trade data used in gravity modeling). Similar to their 2006 findings, they also find the PPML estimator to be consistent and generally well-behaved in the presence of high

---

[18] The 3 dataset consist of about 13%, 15%, 25% of zero trade values.

[19] For instance the Cobb-Douglas production function, the consumer-demand systems and the Stochastic impact by regression on population, affluence and technology, which is a popular model used in environmental economics.

[20] Santos Silva and Tenreyro (2006) used a data generating process that generates no zero values but only positive values. Martin and Pham adopted similar design to Santos Silva and Tenreyro (2006) Monte Carlo simulation but however modified it by including a threshold trade level that must be exceeded before positive trade levels are observed. Where the chosen threshold generates zero trade frequencies, which is similar to those observed in studies using aggregate trade flows.

proportion of zeros, and to be more robust to departures from the heteroscedasticity assumption (overdispersion); as its performance is not affected even with the overdispersion in the dependent variable and the presence of excessive zero values.

Among the class of the generalized linear models, the Gamma Pseudo Maximum Likelihood (GPML) technique has also been used in taking care of the zero trade values and associated problem of the logarithm transformation (c.f. Manny and Mullay, 2001). Similar to the log linear model, the GMPL is said to be a more efficient estimator under the assumption that the conditional variance is a function of higher powers of the conditional mean, as it gives more weights to the conditional mean. Santos and Sliver and Tenreyro (2011) found that the GPML is consistent and well behaved under Monte Carlo simulation in the presence of excessive zero values whose data generation process follows the constant elasticity model. However, it is found to have a larger bias than the PPML, suggesting that the PPML is the best performing estimator (c.f. Santos Sliver and Tenreyro, 2011). In addition, Martinez- Zarzaso (2013) noted that the GPML may also suffer from substantial loss of precision, particularly, if the variance function is misspecified or if the log-scale residuals have high kurtosis.

Another class of the generalized linear model is the nonlinear least square (NLS) technique, which has also been used in the trade literature (c.f. Frankel and Wei, 1993) or used in comparison with other non-linear estimators (e.g. Santos Silva and Tenreyro 2006; Gomez-Herrera, 2011; Martinez-Zarzaso, 2013). Santos Silva and Tenreyro (2006) however show that although both GPML and NLS can take care of these two problems, the PPML is still the preferred estimator as the NLS technique assigns more weight to noisier observations, which reduces the efficiency of the estimator. This is because, while PPML gives the same weights to all observations, and assumes that the conditional variance is proportional to the conditional mean, however, GPLM and NLS give more weights to observations with large mean. This is because the curvatures of the conditional mean is more pronounced here, which are also generally observations with large variance, implying nosier observations. In addition, *ibid* noted that the estimator can also be very inefficient because it generally ignores the heteroscedasticity in the data**.**

Heckman (1979) sample selection model[21] has also been frequently used in the literature. Noting that the standard practice of excluding zero bilateral trade observations can potentially give rise to sample selection bias, especially if the eliminated zeros are not randomly done, and estimating non-randomly selected sample is a specification error and can potentially bias the results. Heckman, therefore, developed a model that corrects for this sample selection bias which is a two-step statistical approach in which the model is estimated under the normality assumption. The first step of the Heckman model involves estimating an equation (Probit regression) for the probability of exporting at the firm level based on the decisions of the firms and then using it in estimating the volume of trade. Heckman (1979) correction model allows one to correct for selection bias in non-randomly selected samples and has also been frequently used in the gravity model literature to correct for problems relating to zero valued trade flows (c.f. Linder and Groot, 2006; Munasib and Roy, 2011). Linder and Groot, (2006) noted that sample selection model uses the information provided by the zero valued trade observations; thus, providing information on the underlying decision process regarding the zero trade flows, while arbitrary truncating and censoring are ad-hoc crude methods and they do not give accurate results compared to the sample selection model. They argued that unlike truncated OLS, without sound theoretical background, the samples election model is theoretically sound and offers an econometrically elegant solution to estimate gravity equation that includes zero trade flows.

Further, in a methodological paper, Helpman, Melitz and Rubinstein (2008) (thereafter HMR), noted that the estimation of bilateral trade flows using the gravity equation is not only subjected to

---

[21] Heckman model is also referred to as sample selection or Tobit II model. The model makes a selection of trading and non-trading country pairs – sample selection.

sample selection bias (if the non-zero exports do not occur randomly), but that estimates may also be vulnerable to omitted variable bias if the number of exporting firms within an industry (extensive margin of trade) is not accounted for. The idea is that due to trade costs, firms differ in productivity (firm heterogeneity) and only firms with productivity level beyond a threshold end up exporting.

HMR therefore, extended Heckman (1979) procedure by controlling for both sample selection bias and firm heterogeneity bias and solve the zero problem by also developing a two-step estimation procedure which exploits the non-random presence of zero trade flows in the aggregate bilateral trade data. The aim of the HMR two-step procedure is to correct both the sample selection bias resulting from eliminating zero trade flows when estimating the logarithmic form of the gravity equation and the bias caused by unobserved firm heterogeneity that result from omitted variable, which also measures the effect of the number of exporting firms (extensive margin). The first step involves estimating an equation (Probit regression) for the probability of exporting at the firm level based on the decisions of the firms and then using it in estimating the effects on the extensive margin of trade (the decision to export from country i to j). The second step is a gravity equation estimated in its logarithm form and involves using the predicted probabilities obtained in the first step to estimate the effects on the intensive margin of trade (the number of exporting firms from country i to j).

Helpman et al., (2008) posit that the excluded variable must not be correlated with the error term of the second stage equation but must be correlated with trade volume (the dependent variable). In addition, the excluded variable must influenced trade through fixed trade cost and not through variable trade cost because the latter impact on the extent of trade volume, and as such, is not uncorrelated with the second stage equation. However, Burger et al., (2009) noted that one important drawback of the Heckman (1979) and Helpman et al. (2008) models is that, it is difficult to satisfy the exclusion restriction as the instrumental variable is most often difficult to find. Examples of exclusion variables used in the literature are common religion and common language variables (Helpman et al., 2008); governance indicators of regulatory quality (Shepotylo, 2009); historical frequency of positive trade between country pairs (Linder and de Groot, 2006; Haq et al., 2010 and Bouet et al., 2008). However, both Linder and de Groot (2006) and Haq et al., (2010) include the excluded variable in both equations and impose the normality of the error term in the two equations – an identification condition implying a zero covariance between both equations.

Notwithstanding the aforementioned advantages of the HMR, some other limitations have been identified regarding its application. Both the Heckman (1979) and the HMR trade flow equations are usually transformed to the logarithmic form before estimated and might cause biased coefficient (Haworth and Vincent, 1979; Santos Silva and Tenreyro. 2006). In addition, Santos Silva and Tenreyro (2009) and Flam and Nordström (2011) also show that HMR does not control for heteroscedasticity which is usually pervasive in most trade data. For instance, Santos Silva and Tenreyro (2009) show that the assumption of homoscedasticity[22] error term for all country pairs by the HMR results in serious misspecifications as HMR does not control for heteroscedasticity, consequently casting doubts on the validity of inferences drawn from the model. They also pointed out that in contrast to models which can be made robust to the presence of heteroscedasticity, the consistency of the HMR model is only possible under the 'unrealistic' homoscedasticity assumption, which they identified as the most important drawback of the model as it is too strong to make it applicable or practicable to trade data in which heteroscedasticity is pervasive. They therefore posit that the presence of heteroscedasticity in the data preclude the estimation of any model that purports to identify the effects of the covariates in the intensive and extensive margins, at least with the current econometric technology (Santos Silva and Tenreyro, 2009).

---

[22] The Helpman et al. (2008) model hinges heavily on both the homoscedasticity and normality assumptions to be consistent.

In sum, as noted in the review, each technique has its pros and cons and the ‘workhorse’ or best performing model for the estimation of the gravity equation still remains unclear as the consensus on a commonly accepted solution has not yet been reached. Therefore, given the pros and cons of each estimator, the determination of the best performing estimator remains an empirical issue.

4.0 Review of the Model Specifications

In general, in line with the various estimation techniques previously discussed, the volume of bilateral trade flow between countries i and j in year t can be represented in either the multiplicative or logarithmic forms. For the sake of comparison and completeness, we adopt the Anderson and van Wincoop (2003) equation as the starting point of the review. First, it is widely accepted in the literature; second, it ensures the modeling of multilateral trade resistance, which if omitted can bias the estimated gravity coefficients (c.f. Baldwin and Taglioni, 2006; Fenstra 2006, etc).

4.1 Log-Linear Models

We begin with the following multiplicative gravity equation:

$$y_{ijt} = \beta_0 GDP_{it}^{\beta_1} GDP_{jt}^{\beta_2} D_{ij}^{\beta_3} lang_{jt}^{\beta_5} Col_{ij}^{\beta_6} Llock_{ij}^{\beta_7} RTA_{ij}^{\beta_8} \varepsilon_{ijt} \tag{1}$$

Taking the natural logarithm of both sides of equation (1), yields a log linear gravity model given as:

$$\begin{aligned} \ln y_{ijt} &= \beta_0 + \beta_1 \ln GDP_{it} + \beta_2 \ln GDP_{jt} + \beta_3 \ln D_{ij} \beta_4 lang \\ &+ \beta_5 Col_{ij} + \beta_6 Llock_{ij} + \beta_7 RTA_{ij} + \varepsilon_{ijt} \end{aligned} \tag{2}$$

Where ln denotes the natural logarithms of the variables; *i* and *j* are exporter and importer subscripts respectively while t denotes time period; y_{ijt} is exports value from country *i* to country *j* in time *t* in current US \$; GDP_{it} and GDP_{jt} are the gross domestic products of countries *i* and *j* in time *t* in current PPP US \$, respectively; whose coefficients are expected to be positive. D_{ij} is the geographical distance between the major cities of countries *i* and *j*; $lang_{ij}$ is a dummy that take the value of 1 when countries *i* and *j* speak the same official language, zero otherwise; Col_{ij} is a dummy variable that takes the value of 1 when countries *i* had colonized country j in the past, zero otherwise; $Llock_{ij}$ takes the value of 1 when at least one of the country-pair is a landlocked countries, zero otherwise. RTA_{ij} takes the value of 1 when trading countries belong to similar trade agreement, zero otherwise. Finally, ε_{ijt} is the two-way error component term of the model, $\varepsilon_{ijt} = \nu_i + \gamma_t + \mu_{ijt}$; ν_i is the unobserved individual effects; γ_t unobserved time effect; and μ_{ijt} is the remaining part of the stochastic disturbance term. Specifically, ν_i captured the country specific unobservable effects – the exporter and importer fixed effects δ_i and δ_j respectively. Thus, δ_i and δ_j are the exporter and importers fixed effect – the multilateral resistant term, while γ_t is the time effect, all of which correct for the biases from estimating panel data (Baldwin and Tagloni, 2006).

Apriori, we expect β_1 to be positive as high level of income in the exporting country denotes a high level of production ceteris paribus, which increases the exports goods; the coefficients on β_2 is also expected to be positive as high income level in importing countries stimulates higher imports. The distance coefficient is however expected to be negative as it is a proxy of all trade cost. Chaney (2014) assert that, in the special case of Zipt’ law, the distance elasticity is constant and equals -1 if firm sizes is ‘Pareto’ and the average distance squared is an increasing power function of its size. The coefficients on lang, Col, Llock and RTA are all expected to be positive.

Equation (2) is generally estimated by pooled OLS and other estimators such as fixed effect and random effect estimators, Tobit, Heckman and Helpman models. Where the log linear equation is consistent when the conditional variance $V[y_i \mid x]$ is proportional to the square of the mean $E[y_i \mid x]$, (that is $V[y_{ijt} \mid x] \propto E[y_{ijt} \mid x]^2$)

Pooled regression model

The OLS estimation of equation (2) is specified as either censored OLS in which case, we add a constant 'c' to replace the entire zero trade observation or by using the truncated OLS where all zero records are deleted. Here, the model assumes the error term to be linearly and independently distributed with zero mean and constant variance $\varepsilon \sim N(0,\sigma^2)$.

Fixed effects model

An alternative way to estimate equation (2) is to control for unobserved heterogeneity using panel data estimators such as the fixed effects technique. Assuming the variables are correlated with the unobserved heterogeneity, the fixed effects estimator becomes:

$$\ln y_{ijt} = \beta_0 + \beta_1 \ln GDP_{it} + \beta_2 \ln GDP_{jt} + \beta_3 P_{it} + \beta_4 P_{jt} + \beta_5 \ln S_{it} + \beta_6 RTA + \varepsilon_{ijt} \qquad (3)$$

Where ε_{ijt} is the two-way error component model $\varepsilon_{ijt} = \nu_i + \gamma_t + \mu_{ijt}$; ν_i is the unobserved individual effects which is represented or captured by country specific unobservable effects – the exporter and importer fixed effects δ_i and δ_j respectively; γ_t unobserved time effect; and μ_{ijt} is the remaining part of the stochastic disturbance term. Both ν_i and γ_t are assumed to be fixed parameters to be estimated while μ_{ijt} is assumed to be $IID(0,\sigma_\mu^2)$. Also, P_{it} , P_{jt} and S_{it} represent population of exporting, importing and trade policy variables.

Equation (3) assumes that the explanatory variables and the unobserved heterogeneity are correlated: $E(\nu_i \mid x_{ijt}) \neq 0$ and $E(\gamma_t \mid x_{ijt}) \neq 0$ that the explanatory variables are independent of the residual error term μ_{ijt} for all *i*, *j*, and *t* - $E(x_{ijt} \mid \mu_{ijt}) = 0$; where x_{ijt} is defined as the explanatory variables of the gravity equation in (3) above. All other variables remained as earlier defined. All time invariant explanatory variables are perfectly collinear with the fixed effects and are dropped from the model.

Random effects model

Alternatively, equation (2) can be estimated using the FGLS estimator which on the contrary assumes orthogonally between the explanatory variable and the unobserved heterogeneity (Baltagi, 2008). The random effects model is specified as equation (2) with the difference that the explanatory variables (x_{ijt}) now contains both time invariant and time varying explanatory variables; and $\nu_i \sim IID(0,\sigma_\nu^2)$, $\gamma_t \sim IID(0,\sigma_\gamma^2)$, and $\mu_{ijt} \sim IID(0,\sigma_\mu^2)$; and x_{ijt} is independent of the unobserved heterogeneity $\nu_i and \gamma_t$ as well as the remainder of the error term μ_{ijt} for all *i* and *t* - that is, $E(x_{ijt} \mid \nu_i, \gamma_t, \mu_{ijt}) = 0$

Heckman model

The Heckman approach is a two-stage estimation procedure consisting of two separate equations: the selection equation and the trade flow equation. Following Heckman (1978), we specify a two stage equation as:

$$\begin{aligned}\rho_{ij} = P(y_{ijt} = 1/x_{ijt}) = \Phi(\beta_0 + \beta_1 \ln GDP_{it} + \beta_2 \ln GDP_{jt} + \beta_3 P_{it} + \beta_4 P_{jt} + \beta_5 \ln S_{it} \\ + \beta_6 \ln D_{ij} + \beta_7 lang + \beta_8 Col + \beta_9 RTA + \varepsilon_{ijt}) \ldots\ldots\ldots\ldots\ldots\ldots\ldots(4)\end{aligned}$$

$$E[y_{ijt} / y_{ijt} = 1] = \beta_0 + \beta_1 \ln GDP_{it} + \beta_2 \ln GDP_{jt} + \beta_3 P_{it} + \beta_4 P_{jt} + \beta_5 \ln S_{it} + \beta_6 \ln D_{ij} + \beta_7 Col + \beta_8 RTA + \sigma_{12}\lambda_{ij} + \varepsilon_{ijt} \text{...}(5)$$

Equation (4) is the selection equation given as a Probit maximum likelihood regression model, which determines the binary decision of whether to trade or not. ρ_{ij} is the probability that country *i* exports to country *j*, conditional on the observed variables; y_{ijt} which is our dependent variable is export from country *i* to *j* is now a binary variable which is equal to 1 if country *i* exports to country *j* $(y_{ijt} = 1)$ and zero when it does not $(y_{ijt} = 0)$; x_{ijt} is the vector of all explanatory variables earlier defined in equation (2) potentially including some fixed effects; and $\Phi(.)$ is the cumulative distributive function of the bivariate normal distribution.

The second equation, a log linear model (equation 5), is the trade flow equation, which gives the conditional trade flow given that the observation on trade flows is positive. Where y_{ij} is the exports from country i to country j in logarithmic form, given that observed trade flow y_{ij} is positive; s_i is the vector of the same explanatory variables used in equation (4) in logarithmic form minus an exclusion restriction variable which does not enter the second stage regression ; σ_{12} is the covariance of the unobserved errors or unobserved trade costs of the selection equation (4), where the unobserved errors are assumed to be bivariate normally distributed; and λ_{ij} is the Heckman's lamda, also called the inverse Mills ratio, which is obtained from the first stage regression and added to the second equation to controls for sample selection bias as a result of non-randomization of the sample of nonzero exports. It is given as

$$\lambda_{ij} = \frac{\phi(x_{ijt}\beta)}{\Phi(x_{ijt}\beta)} \qquad \text{.............................}(6)$$

In the trade flow equation, we included the same set of explanatory variables contained in the selection equation, except for the exclusion variable, which in our case is the common language as used in Helpman et al. (2008). This selection variable is assumed to be correlated with the fixed costs of trade but weakly or negligibly correlated with the variable trade costs.

Helpman, Melitz and Rubinstein (HRM) model

Similar to the Heckman procedure, the HRM model is also a 2-stage procedure where the first stage equation is same as that of Heckman's sample selection equation (4), while the trade flow equation[23] (estimated as a log linear model) is extended by including an additional variable to control for unobserved heterogeneity. The impact of trade barrier is thus decomposed into the intensive margin (trade volume per exporter) and the extensive margin (number of exporters). The trade flow equation becomes:

$$E[y_{ijt} / y_{ijt} = 1] = \beta_0 + \beta_1 \ln GDP_{it} + \beta_2 \ln GDP_{jt} + \beta_3 P_{it} + \beta_4 P_{jt} + \beta_5 \ln S_{it} + \beta_6 \ln D_{ij} + \beta_7 Col + \beta_8 RTA + \sigma_{12}\lambda_{ij} + \textstyle\int_{ij} + \varepsilon_{ijt} \text{...}(7)$$

Where $\int_{ij}$ controls for unobserved firm heterogeneity - the number of firms exporting from country *i* to *j*, which can possibly be zero.

---

[23] In this second equation, HMR controls for both sample selection bias through the mill ratio and also control for omitted variable bias in the estimates by also accounting for unobserved firm heterogeneity through the inclusion of additional variable which accounts for selection of films into the export markets.

4.2 Multiplicative Models' Estimators – The Generalized linear models (GLM)

The generalized linear models estimate the constant elasticity gravity model in its multiplicative form as:

$$y_{ijt} = \exp(x_{ijt}\beta)\varepsilon_{ijt} = \tag{8}$$

Where $E(\varepsilon_i \mid x) = 1$; x_{ijt} are the explanatory variables of the gravity equation earlier defined in equation (1) above; β is the parameters and ε_{ijt} is the composite error term which contains the importer and exporter fixed effects, time effects and the remainder of the error term.

The Poisson Pseudo Maximum Likelihood (PPML) Estimator

The PPML estimates β by solving the following first-order conditions:

$$\sum_{i=1}^{n} \left[y_{ijt} - \exp(x_{ijt}\beta)\right]x_{ijt} = 0 \tag{9}$$

Equation (10) is the PPML estimator, which is consistent[24] under the estimator's equidispersion[25] assumption that the conditional mean $E[y_{ijt} \mid x]$ given as $\exp(x_{ijt}\beta)$ is equal to the conditional variance $V[y_{ijt} \mid x]$ - this is implied by equation (11) which imposes restrictions on the conditional moments of the dependent variable.

$$E[y_{ijt} \mid x] = \exp(x_{ijt}\beta) \propto V[y_{ijt} \mid x]\ldots \qquad \ldots\ldots\ldots\ldots\ldots(10)$$

However, the equidispersion assumption is unlikely to hold (Santos Sliver and Tenreyro, 2006; Martinez-Zarzaso, 2013) as the estimator does not fully account for the presence of heteroscedasticity in the model. In other words, the estimator does not fully take account of the presence of unobserved heterogeneity caused by the unobserved trade costs, thus making the conditional variance to be greater than the conditional mean$_{26}$. Thus, inferences are based on the Eicker-White robust covariance matrix estimator (Eicker, 1963; White, 1980).

Gamma Pseudo Maximum Likelihood Estimator (GPML)

The GPML estimator also belongs to the class of the GLM and it is obtained by solving the following first-order conditions of the following likelihood function:

$$\sum_{i=1}^{n} \left[y_{ijt} - \exp(x_{ijt}\beta)\right]\exp(-x_{ijt}\beta)x_{ijt} = 0 \qquad \ldots\ldots\ldots\ldots\ldots(11)$$

Similar to the log linear model, this estimator assumes that the conditional variance $V[y_{ijt} \mid x]$ is proportional to the square of the conditional mean $E[y_{ijt} \mid x]^2$.

---

[24] To obtain consistent estimates, while the trade flow variable is assumed to follow a Poisson distribution, however, the data need not follow a Poisson distribution, and the independent variable needs not be an integer (Gourieroux, Monfort, and Trognon, 1984).

[25] PPML gives the same weights to all observations, such that all the observations have the same information on the parameters because the additional information about the curvature of the mean which comes from observations with large mean is offset by their large variance (Santos Sliva and Tenreyro, 2006).

[26] Although the PPML specification hinges on the assumption of equidispersion of the dependent variable, however, SST 2006 show that the PPML is still well-behaved and consistent even with departure from this assumption.

The GPML is a more efficient estimator[27] given that the conditional variance is a function of higher powers of the conditional mean, as it gives more weights to the conditional mean. Imposing this gives equation (13) which is the consistency assumption of the estimator.

$$V[y_{ijt} \mid x] \propto E[y_{ijt} \mid x]^2 \qquad \text{………………(12)}$$

Non linear Least Square Estimator (NLS)

The NLS estimator also specifies the gravity equation in a multiplicative form; where the first order condition for this estimator is given as:

$$\sum_{i=1}^{n} \left[y_{ijt} - \exp(x_{ijt}\beta) \right] \exp(x_{ijt}\beta) x_{ijt} = 0 \qquad \text{……………....(13)}$$

Here, the necessary condition for consistency of the NLS is for the variance of the error to be constant such that $V[y_{ijt} \mid x] = 1$ ……… ………(14)

Negative Binomial Poisson Maximum Likelihood (NBPML) Estimator

Since the equidispersion assumption does not always holds for the Poisson model, the Negative Binomial (NB) model, a modified Poisson model is alternatively employed to deal appropriately with the occurrence of overdispersion in the dependent variables (c.f. Burger et al., 2009). Following Winkelmann (2008), the negative binomial probability distribution function for y is given as:

$$\Pr\left[y_{ijt} \right] = \frac{\Gamma(\alpha + y_{ijt})}{\Gamma(\alpha)\Gamma(y_{ijt} + 1)} \left(\frac{\alpha}{\alpha + \exp(x_{ijt}\beta)} \right)^{\alpha} \left(\frac{\exp(x_{ijt}\beta)}{\alpha + \exp(x_{ijt}\beta)} \right)^{y_{ijt}} \qquad \text{…………………(15)}$$

Where Γ is the gamma function, α is the dispersion term which allows the conditional variance to exceed the conditional mean and also determines the degree of variance dispersion (Verbeek, 2004; Cameroon and Trivedi, 1986).The larger α is, the larger the degree of overdispersion in the dependent variable. A likelihood ratio test on α can be used to test if the NBPML is more appropriate model compare to the PPML (Cameroon and Trivedi, 1986; Winkelmann, 2008).

We consider two variants of the NBPML[28] model here: the Negbin I and Negbin II models which are obtained when the variance is a linear or quadratic function of the mean respectively.

The *Negbin I* estimator of $\hat{\beta}$ is given by solving the following first-order condition for $\hat{\beta}$:

$$\sum_{i=1}^{n} \left[\sum_{j=1}^{yitj} \left(\frac{\alpha \exp(x_{ijt}\beta)}{\alpha \exp(x_{ijt}\beta) + j - 1} \right) x_{ijt} + \alpha \exp(x_{jti}\beta) x_{ijt} \right] = 0 \qquad \text{…………………(16)}$$

The model is obtained by solving equation (17) under the estimator's overdispersion assumption that the conditional variance $V[y_{ijt} \mid x] = (1 + \alpha^{-1}) \exp(x_{ijt}\beta)$ is greater than the conditional mean $E\left[y_{ijt} \mid x \right]$ implied by equation 16.

---

[27] It down weights observations with larger conditional means (Santos Sliva and Tenreyro, 2006).

[28] There are also other variants such as the Negbin$_k$ and Negbin$_x$ models which are also asymptotically efficient if specified correctly (see Winkelmann, 2008)

$$V[y_{ijt} \mid x] = (1+\alpha^{-1})\exp(x_{ijt}\beta) > E(x_{ijt}\beta) \qquad \text{……………….(17)}$$

The *Negbin II* estimator of $\hat{\beta}$ is obtained by solving the following first-order conditions:

$$\sum_{i=1}^{n} \frac{y_{ijt} - \exp(x_{ijt}\beta)}{1+\alpha^{-1}\exp(x_{ijt}\beta)} x_{ijt} = 0 \qquad \text{…………………(18)}$$

The Negbin II model involves solving for equation (19) and assuming that the conditional variance $V[y_{ijt} \mid x] = \exp(x_{ijt}\beta) + \alpha^{-1}\left[\exp(x_{ijt}\beta)\right]^2$ is greater than the conditional mean $\exp(x_{ijt}\beta)$ as given in equation (20).

$$V[y_{ijt} \mid x] = \exp(x_{ijt}\beta) + \alpha^{-1}\left[\exp(x_{ijt}\beta)\right]^2 > E[y_{ijt} \mid x] \qquad \text{…...........(19)}$$

In both variants, the NBPML expected value is given as that of the PPML, however, the variance is specified to include the mean $\exp(x_{ijt}\beta)$ and an unobserved heterogeneity given as a dispersion parameter α [29], which allows unobserved heterogeneity to be incorporated into the model. In addition, the dispersion parameter is allowed to take on other values than 1, thereby explicitly taking care of overdispersion.

Zero Inflated (ZI)Estimators

The zero inflated estimators consider two different groups within the population – the non-poissoneness group with a strictly zero probability of trading, i.e., those who do not trade at all and the poissioness group who has a non-zero probability of trading, some of which are actually trading and others are not. These two underlying processes of the ZI model are estimated in 2 stages (equations 21 and 22). The first stage equation[30] specifies a logit (or probit) regression to estimate the non-Poisson zeros i.e. the probability of no bilateral trade. The second stage is given as a Poisson regression model given that country-pairs have a non-zero probability to trade (Poission zeros). We distinguish between two types of ZI estimators – the Zero Inflated Poisson (ZIP) and the Zero Inflated Negative Binomial (ZINB) estimators.

More specifically, the ZIP model takes the form:

$$P(y_{ijt} = 0 \mid x_{ijt}) = \omega + (1-\omega)\exp(-\exp(x_{ijt}\beta) \qquad \text{……………….(20)}$$

$$P(y_{ijt}) = (1-\omega)\frac{\exp(-\exp(x_{ijt}\beta)(\exp x_{ijt}\beta)}{y_{jt}!} \qquad \text{...……………(21)}$$

where ω is the proportion observations with a strictly zero count $(0 > \omega > 1)$, which reduces the zero inflated Poisson model to the Poisson model when it is 0.

---

[29] This dispersion parameter serves as a formal test of overdispersion in the dependent variable.

[30] The first stage equation is similar to Heckman's first stage except that the YINPB is less stringent and less restrictive as it neither rely on the normality assumption nor exclusion restriction in the second stage of the model. In addition, it is said to account for unobserved heterogeneity in the population with a zero count.

The ZINB regression model is defined likewise. The first stage equation[31] specifies a logit regression while the second stage is given as a Negative Binomial Poisson Maximum Likelihood regression model (e.g. NBPML II). More formal

$$\Pr\left[y_{ijt} = 0\right] = \omega + (1-\omega)\left(\frac{\alpha^{-1}}{\alpha^{-1} + \exp(x_{ijt}\beta)}\right)^{\alpha^{-1}} \qquad \text{........(22)}$$

$$\Pr\left[y_{ijt}\right] = (1-\omega)\frac{\Gamma(\alpha^{-1} + y_{ijt})}{\Gamma(\alpha^{-1})\Gamma(y_{ijt}+1)}\left(\frac{\alpha^{-1}}{\alpha^{-1} + \exp(x_{ijt}\beta)}\right)^{\alpha^{-1}}\left(\frac{\exp(x_{ijt}\beta)}{\alpha^{-1} + \exp(x_{ijt}\beta)}\right)^{y_{ijt}} \qquad \text{..........(23)}$$

---

[31] The first stage equation is similar to Heckman’s first stage except that the ZINPB is less stringent and less restrictive as it neither relies on the normality assumption nor exclusion restriction in the second stage of the model. In addition, it is said to account for unobserved heterogeneity in the population with a zero count.

Table 1: The Zero Trade and Logarithmic Transformation in Gravity Modeling – A Summary of the Debate (GMED)

Model/Estimator	Scholar	Characteristics/Merit	Criticism/Demerit	Response to Critics
Tobit	Anderson and Marcoiller (2002), Rose (2004), Martin and Pham (2008).	- To deal with the zero trade problem due to unobservable trade flows or measurement error from rounding up. - Applied to fit dataset that is only observable over some range. - Applicable there is difference between actual outcomes and desired outcomes.	- Linder and de Groot (2006) opined that zero trade occur due to binary decision making on the profitability of trade and not from censoring that the model posited, which makes it inappropriate to take care of the zero trade. - Frankel (1979) argued that the estimator is liable to measurement errors, which will impact on the result due to the artificial censoring of positive small trade values. - In response to the position of Martin and Pham (2008), Santoa Silva and Tenreyro (2011) find the threshold Tobit model to have large bias that rise with sample size, which makes it an inconsistence estimator in a simulation exercise.	- Martin and Pham (2008) suggested the use of Eaton and Tamura (1994) threshold Tobit model that gives the lowest bias and outperform all other estimators in a simulation exercise.
Poisson Pseudo Maximum Likelihood (PPML)	Santos Silva and Tenreyro (2006, 2008, 2009, 2011), Staub and Winkelmann (2012).	- It is used to deal with the zero trade and logarithm transformation. - The gravity equation is specified at levels in order to avoid the problem that arose using OLS under logarithm transformation. - It takes into consideration observed heterogeneity; zero trade dealt with through the multiplicative form of the fixed effects in PPML and avoid under-prediction of large trade volume by generating estimates of trade flows rather than the log of trade flows. - Gives the lowest bias among estimators. - Proponents suggest the estimator as the workhorse for the gravity model.	- Burger et al. (2009) argued that the model is vulnerable to over-dispersion in the dependent variable and excess zero flows. This only takes care of observed heterogeneity and unobserved ones. - The assumption of equidispersion in the dependent variable leads to overdispersion due to unobserved heterogeneity. - The overdispersion generates consistent but inefficient estimates of trade flows (Burger, et al. 2009; Turkson, 2010) - Martinez-Zarzoso (2013) opined that PPML is not always the best estimator as its estimates are outperformed by both OLS and	- Santo Silva and Tenreyro (2011) opined that despite the identified overdispersion and excessive zero trade problems, PPML is consistent and generally well-behaved in the presence of overdispersion in the dependent variable and large zero trade will not affect its performance. - Soren and Bruemmer (2012) argued that PPML performs quite well under overdispersion, and show that the PPML is well-behaved under bimodal distributed trade data. - Santo Silva and Tenreyro (2008) responded by justifying the use of PPML as the best estimator in gravity model, but acknowledged that

			FGLS estimates in out of sample forecast, so, it is not always the best estimator. - The PPML assumption regarding the pattern of heteroscedasticity is rejected by the data in most cases (Martinez-Zarzoso, 2013). - Martin and Pham (2008) argue that PPML is not robust to the joint problems of zero trade and heteroscedasticity.	PPML estimator can be outperformed by other estimators in some cases. - PPML consistent in the presence of excessive trade zero (Staub and Winkelman, 2012). - Santo Silva and Tenreyro (2011) responded to the critics of PPML arguing that the studies of the critics of PPML did not generate its data through a constant elasticity model, with which their study did. - Also, Santo Silva and Tenreyro (2011) re-investigate the performance of PPML in the presence of large zero trade data in a constant elasticity model. The results show that PPML estimator is consistent, well-behaved with large zero trade and not affected by overdispersion in the dependent variable.
Negative Binomial Pseudo Maximum Likelihood (NBPML) and Zero Inflated Models e.g. Zero Inflated Pseudo Maximum Likelihood (ZIPML) technique, Zero Inflated Binomial Pseudo Maximum Likelihood (ZINBPML).	Burger et al. (2009)	- To correct for the overdispersion in the dependent variable and the vulnerability of the PPML to excessive trade zero. - It incorporates unobserved heterogeneity into the condition mean and thus, takes care of unobserved heterogeneity.	- One of the drawbacks of NBPML and PPML is the excessive number of zero trade that is derived from non-Poissoness of the model (Johnson and Kotz, 1969). - Turkson (2011) argued that these estimation techniques cannot handle excessive zero. - Staub and Winkelmann (2012) posit that both ZIPML and ZINBPML are inconsistent if the models are misspecified.	- Burger et al. (2009) opined that even though the Poisson model and NBPML model can technically handle zero trade, however, both are not well positioned in the case where the number of observed zeros trade value is greater than the number of zero predicted by the model. - The Zero Inflated Models perform better as they corrected excessive zeros and overdispersion in the dependent variables. The models theoretically well situated in Poisson and non-Poisson estimation.
Zero Inflated Poisson Quasi Likelihood (ZINPQL)	Staub and Winelmann (2012)	- Consistent in the presence of excessive zero trade. - Unaffected by unobserved heterogeneity. - It is robust to misspecification as it	- ZINPQL can be less efficient compared to zero inflated estimators when the zero inflated models are correctly specified.	

		consistently estimate the regression coefficients irrespective of the true distribution of the counts, while ZIPML and ZINBPML demonstrate considerable bias in the medium sample.		
FGLS and other generalized least square (GLM) e.g. Gamma Pseudo Maximum Likelihood (GPML), Non-Linear Least Square (NLS).	Martinez-Zarzoso et al (2007), Martinez-Zarzoso (2013) -FGLS, Manny and Mullay (2001) – GPML, Frankel and Wei (1993) – NLS.	- FGLS can be applied in the presence of unknown form of heteroscedasticity. - It is an efficient estimator among the class of least square estimators. - Variance of the disturbances needs to be re-estimated to correct for heteroscedasticity errors. - The comparison of the best estimators should be between FGLS and other generalized least models (GLMs) such as; Non-linear least square (NLS), Gamma Poisson Maximum Likelihood (GPML) and PPML. - Gamma Psuedo Maximum Likelihood (GPML) techniques is more efficient under the assumption that the conditional variance depends on higher power of the conditional mean, thus, given more weight to conditional mean. - NLS assigns more weight to noisier observations. - NLS consistent in the modeling of zero. - NLS gives more weight to observations with large vaeiance.	- Santos Silva and Tenreyro (2008) debunked the claim of FGLS proponents and provided justification for the PPML estimator in the context of log-linear gravity model. - Santos Silva and Tenreyro (2011) found GMPL to be consistent and well-behaved under Monte Carlo simulation with excessive zero trade values in a constant elasticity model, but has a larger bias than the PPML. - Martine-Zarzoso (2013) argued that the GMPL may suffer from substantial loss of precision whenever the variance function is misspecified or when the log-scale residuals have high kurtosis. - NLS efficiency is reduced due to its allocation of more weight to noisier observation (Santos Silva and Tenreyro, 2006). Also, NLS is inefficient because it generally ignores heteroscedasticity in the data.	Martinez-Zarzoso (2013) argued that the choice of the best estimator is a function of the dataset and there is no absolute best estimator for all typology of dataset. Thus, the most appropriate estimator is data specific and could be determined by model selection tests.
Heckman Selection Model	Heckman (1979), Linder and de Groot (2006), Munasib and Roy (2011).	- This model corrects for sample selection bias and specification error when zero trade do not occur randomly. - It is a two-step approach under the normality assumption: first, estimation of the probability of trade at the firm levels (probit regression), finally, using the first	- Burger et al. (2009) argued that in both Heckman and HMR models, it is difficult to satisfy the exclusion restriction because the instrumental variable is often difficult to find. - The transformation of these models into logarithmic form before	

		approach to estimate the volume of trade. - It has theoretically sound method and offers econometrically elegant solution. - Providing avenue of using information from zero trade observation.	estimation might cause biased coefficient (Haworth and Vincent, 1979; Santos Silva and Tenreyro, 2006). - Flam and Nordstrom (2011) and Santos Silva and Tenreyro (2009) posited that these models did not control for heteroscedasticity that are pervasive in trade data.	Linder and de Groot (2006) and Heqetal (2010) included the excluded variables and imposed the normality of the error term.
Extensive and Intensive Trade Margins Model	Helpman, Melitz and Rubinstein – HMR (2008)	- It extended the Heckman model by controlling for both sample selection bias and firm heteroscedasticity. - It solves the zero trade problem with a two-step estimation procedure. - It measures the effects of the number of exporting firms and volume of trade. - First, it estimates the probit regression for probability of trading at the firm's levels (extensive margin). - Using the first stage estimation result to estimate the intensive trade margin. - It assumes homoscedasticity.		

5.0 Concluding Remarks

The theoretical framework underpinning the use of gravity model is no longer in doubt among international economists. Gravity model is very useful in modeling bilateral, regional, plurilateral and multilateral economic relations. The equation can arise from a wide range of trade models; the standard, new and ‘new’ new trade theories. These theoretical options in the application of the models and specification of the equation would depend on the preferred set of assumptions and models. Differences are noticed in the underlying assumptions and models in gravity modeling, which could be due to the various specifications in the empirical studies. These often resulted in difference outcomes and inferences for these studies.

To this end, this review, although do not claim to have exhausted all theoretical and empirical studies, has shown that the current emphasis in the theoretical literature is to ensuring that empirical applications of gravity models is well rooted on its theoretical ground and that it can be linked to anyone of the available and appropriate theoretical frameworks. However, we opined that irrespective of the theoretical framework adopted, most of the subsequent justifications of the gravity equations are variants of the initial theoretical foundation.

The bottom line of this review is that each technique has its pros and cons as enunciated in this paper. Thus, the best performing estimator for the estimation of the gravity equation still remains an empirical issue as the consensus on a commonly accepted solution has not yet been reached. Therefore, given these merit and demerit of each estimator, the gravity model should be used as a workhorse, cookbook or toolkit in the modeling of international economic relations.

References

Anderson, J. E. & van Wincoop, E. (2004) “Trade costs”, Journal of Economic Literature, (17):691-751.

Anderson, J.E., Van Wincoop, E. (2003) “Gravity with gravitas: A solution to the border puzzle”, American Economic Review (93):170–192.

Anderson, J. E. (1979) ‘’A theoretical foundation for the gravity equation’’, American Economic Review, vol.69, March, pp. io6-i6.

Anderson, J.and D., Marcouiller (2002) “Insecurity and the pattern of trade: An empirical investigation”, Review of Economics and Statistics, 84 (2), 342–352.

Armington, P. S. (I969) “A theory of demand for products distinguished by place of production”, IMF Staff Papers, vol. i6, pp. I59-77.

Baldwin, R. and D., Taglioni (2007) “Trade Effects of the Euro: A Comparison of Estimators”, Journal of Economic Integration, 22 (4), 780–818.

Baldwin, R. (2006) “The euro’s trade Effects”. Tech. rep., European Central Bank.

Baldwin R, and V., DiNino (2006) “Euros and Zeros: The Common Currency Effect on Trade in New Goods”, Center for Economic Policy Research Discussion Papers no 5973

Bernard, Andrew B., Jonathan Eaton, J. Bradford Jensen, and Samuel Kortum (2003) “Plants and Productivity in International Trade,” American Economic Review, 93:1268–1290.

Bergstrand, J.H. (1985) “The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence”,Review of Economics and Statistics 67(3):474–81.

Bergstrand, J. H. (1990) “The Heckscher–Ohlin–Samuelson model, the Linder hypothesis, and the determinants of bilateral intra-industry trade”, Economic Journal, 100 (4), 1216–1229

Bergstrand, J.H. (1989) “The Generalized Gravity Equation, Monopolistic Competition, and the Factor Proportions Theory in International Trade”,Review of Economics and Statistics 71(1):143– 53.

Baier, S.L., Bergstrand, J.H. (2001) “The growth of world trade: Tariffs, transport costs, and income Similarity”, Journal of International Economics, (53) 1–27.

Burger M. F. van Oort and G.J. Linders (2009)“On the Specification of the Gravity Model of Trade: Zeros, Excess Zeros and Zero-Inflated Estimation”, Spatial Economic Analysis, 4(2): 167-90

Cameron, A.C. and P.K. Trivedi (1986) “Econometric Models Based on Count Data: Comparisons and Applications of Some Estimators and Tests”, Journal of Applied Econometrics, 1, 29-53.

Chaney, T. (2008) “Distorted Gravity: The Intensive and Extensive Margins of International Trade", American Economic Review (98):1707-21.

Chaney, T. (2014) “The Gravity Equation in International Trade: An Explanation” http://www8.gsb.columbia.edu/programs-admissions/sites/programs-admissions/files/finance/Macro%20Workshop/spring%202014/Thomas%20Chaney.pdf

Chen, N. and D., Novy (2011) “Gravity, Trade Integration and Heterogeneity across Industries”, Journal of International Economics, Vol.85 (No.2), 206-221.

Deardorff, A.(1998) “Determinants of bilateral trade: Does gravity work in a neoclassical world? In: Frankel, J.A. (Ed.)”, The Regionalization of the World Economy, (Chapter 1).

Deardorff, A. (1984), “Testing Trade Theories and Predicting Trade Flows”, In: Handbook of international economics. Vol. 1. Elsevier, pp. 467–517.

De Benedictis, L. and D., Taglioni (2011) “The Gravity Model in International Trade”, In: De Benedictis, L., Salvatici, L. (Eds.), “The Trade Impact of European Union Preferential Policies: An Analysis Through Gravity Models”. Springer, Ch. 4, pp. 55–90.

Eaton, Jonathan, and Samuel S. Kortum, (2002), “Technology, Geography, and Trade,” Econometrica, (70):1741–1779.

Egger, P. and M., Pfaffermayr (2003) “The proper panel econometric specification for the gravity equation: A

three-way model with bilateral interaction effects”, Empirical Economics 28 (3):571–580.

Egger, P. (2000) “A note on the proper econometric specification of the gravity equation”, Economics Letters, 66, 25–31

Egger P, and M., Pfaffermayr (2004) “Distance, Trade and FDI: A SUR Hausman-Taylor Approach”, Journal of Applied Economics, 19: 227-246

Egger P, and M., Pfaffermayr (2003) “The Proper Panel Econometric Specification of the Gravity Equation: A Three Way Model with bilateral interaction effects. Empirical Economics, 28: 571-580

Egger, P. (2004) “On the Problem of Endogenous Unobserved Effects in the Estimation of Gravity Models“, Journal of Economic Integration, vol.9, pp. 182-191

Eaton, J., and A. Tamura (1994): "Bilateralism and Regionalism in Japanese and U.S. Trade and Direct

Foreign Investment Patterns," Journal of the Japanese and International Economies, (8):478-510.

Eichengreen, B. and D. A. Irwin (1998) “The Role of History in Bilateral Trade Flows”, In: Jeffrey A. Frankel, (Eds), The Regionalization of the World Economy (pp.33-62). Chicago, IL and London: University of Chicago Press.

Evenett, S. J. and W., Keller (1998) “On theories explaining the success of the gravity equation”, NBER Working Paper No. 6529, April

Feenstra, R. C., James R. M. and A.K,. Rose (1998) “Understanding the home market

effect and the gravity equation: the role of differentiating goods”, NBER Working Paper No. 6804.

Feenstra, R. C., Markusen, J. A. and A. K., Rose (1998) “Undertstanding the Home Market Effect and the Gravity Equation: The Role of Differentiating Goods", NBER Working Papers 6804, National Bureau of Economic Research, Inc.

Flowerdew, R. and M. Aitkin (1982) “A Method of Fitting the Gravity Model Based on the Poisson Distribution”, Journal of Regional Science, 22, 191-202.

Frankel, Jeffrey A., (1997).”Regional Trading Blocs in the World Economic System”, Institute of International Economics.

Frankel, J. and S., Wei (1993) “Trade Blocs and Currency Blocs”, NBER Working Paper no. 4335. Cambridge, Mass: National Bureau of Economic Research.

Gomez-Herrera, E. (2013) “Comparing Alternative Method to Estimate Gravity Models of Bilateral Trade”, Empirical Economics, Springer, vol. 44(3), pp. 1087-1111.

Gourieroux, C., Montfort, A. and A., Trognon (1984) “Pseudo Maximum Likelihood Methods: Applications to Poisson Models”, Econometrica, 52, 701-20.

Haq Z, Meilke K. and J., Cranfield (2010) “Does the Gravity Model Suffer from Selection Bias? Canadian Agricultural Trade Policy Research Network Working Paper no 90884.

Haveman, J. and D., Hummels (2004) “Alternative Hypotheses and the Volume of Trade: The Gravity Equation and the Extent of Specialization”, The Canadian Journal of Economics / Revue canadienne d’Economique, 37 (1), 199–218.

Haworth, J.M. and P.J., Vincent (1979) “The Stochastic Disturbance Specification and its Implications for Log-Linear Regression”, Environment and Planning A, 11, 781-90.

Head, K. and T., Mayer (2013) “Gravity Equations: Workhorse, Toolkit, and Cookbook", Sciences Po Economics Discussion Papers, 2013-02, Sciences Po, Department of Economicsrs

Heckman, J. (1979) “Sample Selection Bias as a Specification Error”, Econometrica (47): 153-161.

Helpman, E., Melitz, M. and Rubinstein, Y. (2008) “Estimating trade flows: trading partners and trading volumes”, Quarterly Journal of Economics (73): 441–486.

Helpman, E. and P. Krugman(1985)Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition, and the International Economy. Cambridge, MA:MIT Press.

Johnson, N. and S. Kotz (1969) Distributions in Statistics – Discrete Distributions, John Wiley and Sons, New York.

Leamer, E.E., and J., Levinsohn (1995) “International Trade Theory: The Evidence.” In Handbook of International Economics, vol. 3.

Leamer E and R., Stern (1971) “Quantitative International Economics”, Journal of International Economics, 1: 359-361

Linders, G. J. M. and H. L. F. de Groot (2006) “Estimation of the Gravity Equation in the Presence of Zero Flows”, Tinbergen Institute Discussion Paper, No. 06-072/3.

Linnemann, Hans. (1966) An Econometric Study of International Trade Flows.

Mayer, T. and G., Ottaviano (2008). The Happy Few: The Internationalisation of European firms. Bruegel Blueprint Series.

Manning, W. and J., Mullahy (2001) “Estimating log models: to transform or not to transform?”, Journal of Health Economics, 20 (4), 461–494.

Mayer , T. And G., and Ottaviano (2008) “The Happy Few: The Internationalisation of European Firms," Intereconomics: Review of European Economic Policy, Springer, vol. 43(3), pp.135-148.

Melitz, M. J. (2003) “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity”, Econometrica, (71):1695-1725.

Martinez-Zarzoso, I. D. Nowak-Lehman and S. Vollmer (2007) “The log of Gravity Revisited”, CeGE Discussion Paper 64, University of Gottingen.

Martinez-Zarzoso, I. (2013) “The log of Gravity Revisited”, Applied Economics, Volume 45, Issue 3. University of Gottingen.

Martínez-Zarzoso I. and F., Nowak-Lehman (2003) “Augmented Gravity Model: An Empirical Application to Mercosur-European Union Trade Flows”,. Journal of Applied Economics, 6: 291-316

Martin, W. and C. S. Pham (2008) “Estimating the Gravity Model when Zero Trade Flows are Frequent. World Bank manuscript, available at:

http://www.deakin.edu.au/buslaw/aef/workingpapers/papers/2008_03eco.pdf.

McCallum, John (1995) "National Borders Matter: Canada-U.S. Regional Trade Patterns," American Economic Review 85(3):615-623.

Novy, D. (2012) "International Trade without CES: Estimating Translog Gravity", Journal of International Economics, vol. 89, issue 2, pp. 271 – 82. URL http://www.sciencedirect.com/science/article/pii/S0022199612001584?v=s5

Novy, D. (2010) "International Trade Without CES: Estimating Translog Gravity", CEP Discussion Papers dp1031, Centre for Economic Performance, LSE.

Obstfeld, M and K., Rogoff (2001) The six major puzzles in international macroeconomics. Is there common cause? NBER Working paper 7777.

Pöyhönen, P. 1963)"A Tentative Model for the Volume of Trade between Countries", WeltwirtschaftlichesArchiv90(1): 93-99.

Ravenstein, E.G. (1885) "The Laws of Migration", Journal of Royal Statistics Soc 52:241–310.

Rose, J. and E., Van Wincoop (2001) "National money as a barrier to international trade: The real case for currency union", American Economic Review: Papers and Proceedings 91 (2):386–390.

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

Rose, Andrew K., (2000). "One Money One Market: Estimating the Effect of Common Currencies on Trade," Economic Policy, (15): 7–46.

Feenstra, R.C. (2002) "Border Effects and the Gravity Equation: Consistent methods for estimation", Scottish Journal of Political Economy, Vol. 49, No. 5, November,

Santos Silva, J.M.C. and Tenreyro, S. (2006) "The log of gravity", The Review of Economics and Statistics, (88):641-658.

Santos Silva, J.M.C. and Tenreyro, S. (2009) "Trading Partners and Trading Volumes: Implementing the Helpman-Melitz-Rubinstein Model Empirically", CEP Discussion Paper No 935.

Santos-Silva, J. and S., Tenreyro (2011) "Further Simulation Evidence on the Performance of the Poisson-PML Estimator", Economics Letters, 112 (2), 220–222.

Shepotylo O. (2009) "Gravity with Zeros: Estimating Trade Potential of CIS Countries", Kyiv School of Economics Discussion Papers 16.

Sören, P. and B. Bruemmer (2012) "Bimodality & the Performance of PPML", Institute forAgriceconomics Discussion paper 1202, Georg-August Universität Göttingen, Germany.

Staub, K. E. and R., Winkelmann (2013) "Consistent Estimation of Zero-Inflated Count Models", Health Economics, vol. 22(6), pp 673-686.

Tinbergen, J. (1962) Shaping the World Economy: Suggestions for an International Economic Policy, New York: The Twentieth Century Funds.

White, H. (1980) "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct

Test for Heteroskedasticity", Econometrica 48 (4): 817–838.

Verbeek, M. (2004) A Guide to Modern Econometrics." John Wiley & Sons Ltd. Chichester.

Zipf, G.K. (1946) "The P1P2/D Hypothesis: On the Inter-City Movement of Persons. American Sociology Review 11:677–686.

Author contacts:

Fatima Olanike Kareem
GlobalFood Research Training Group 1666 and Department of Economics
Georg-August-University,
Gottingen, Germany
Email: fkareem@uni-goettingen.de

Olayinka Idowu Kareem
Global Economic – GGP
Robert Schuman Centre for Advanced Studies
European University Institute
Badia Fiesolana, Via dei Roccettini 9,
50014 San Domenico di Fiesole
Italy
Email: Olayinka.Kareem@EUI.eu