

CHAPTER THREE: METHODOLOGY

3.1 Introduction

This chapter presents the theoretical framework and empirical models used in this thesis. Given that there are three issues, i.e. structure of exports, determinants of exports and export destinations with unrealized potential, being investigated in this research, the chapter is also divided into three main sections. Section 3.2 discusses the theoretical framework as well as the empirical model employed in investigating both inter-industry trade (INT) and intra-industry trade (IIT). The theoretical framework and empirical models used in investigating the determinants of export trade are presented in Section 3.3, while those dealing with export destinations with unrealized trade potential are the subject matter of Section 3.4.

3.2 Analyzing INT and IIT

This section analyses both INT and IIT. The significance of INT and IIT as well as the analytical framework used to analyze these two phenomena are presented in this section.

3.2.1 Introduction

Total trade is composed of both INT and IIT (Al-Mawali, 2005). The former is a result of different factor endowments across trading countries and the consequent specialization as envisioned by both absolute advantage and comparative advantage trade theories. INT theory is considered as the main reason behind trade between developed and developing countries. On the other hand, IIT is underpinned by both increasing returns and product differentiation trade theories, and is known to dominate trade among developed countries (Al-Mawali, 2005; Jayanthakumaran, 2004; and McCorrison and Sheldon, 1991).

For nearly 200 years, and since Adam Smith's "Wealth of Nations" book in 1776, theories of absolute advantage and Ricardo's comparative advantages (or factor endowments) have been the main reasons explaining international trade. These theories postulate that countries with different resources or factor endowments will trade with each other and they are behind what has been termed INT in trade literature, which is defined as the export and import of products from different industries.

On the other hand, increased observations of simultaneous export and import of similar products from the same industries, witnessed since the late 1950s indicated that traditional factor-endowment theories such as those of Smith, Ricardo, and H-O were not able to provide a comprehensive and sufficient explanation of the volume of international trade which took place between same, rather than dissimilar, products. This gave rise to a new phenomenon called IIT. In the contemporary world, statistical trade data shows that approximately one-quarter of world trade is IIT in nature (Kalbasi, 2003a).

3.2.2 Significance of INT and IIT in international trade

The study of both INT and IIT is important in that it provides a better platform for formulating policies which deal with trade adjustments costs caused by increased trade liberalization (Abd-el-Rahman, 1991; and Greenaway *et al.* 1995). Given that Botswana continues to open its trade regime, from bilateral to multilateral trade levels as detailed in Chapter 4, increased trade, either in the form of INT or IIT will result in cost adjustments. Through increased trade, the sizes of most economic sectors will change; some will experience increased (or decreased) exports, other increased (or decreased) imports while others will experience an increase (and/or decrease) in exports and imports, simultaneously.

The changes in exports and imports described above will involve a shift in resources between sectors (Al-Mawali, 2005). In the case of INT driven sectors, increased imports of competing or similar goods will hurt these sectors, and may result in closure of some production entities in these sectors. The assumption is that increased imports of competing products especially from developed countries will be of better

quality, due to these countries' better production technologies, and relatively cheaper, due to developed countries' efficiencies gained through economies of scale and mass production (Kalbasi, 2003b). Thus, these imports will likely dislodge local producers and this will result in loss of employment in a developing economy like Botswana (mainly as a result of structural changes brought by trade liberalization) and re-allocation of capital to other surviving sectors. For labour, it may mean that those formally employed in the closing sectors will have to find jobs in other industries where their work (occupational) experiences may not match the job requirements and this may involve re-training, which in turn will be done at a cost. At the same time, relocated capital, particularly specialized capital such as machinery and equipment, may end up being redundant. Overall, in the case of INT sectors, increased imports will result in relatively large adjustment costs.

In the case of IIT, simultaneous increases in exports and imports is likely to cause marginal, if any, shift of resources between sectors. According to Grimwade (1989), "Intra-industry specialization is likely to give rise to fewer adjustment problems than Inter-industry specialization necessitates a movement of resources from import-competing to export competing industries". Resources adjustment difficulties, especially for labour, happen where resources are geographically and occupationally immobile in the short term, resulting in large-scale structural employment. In the case of intra-industry specialization, adjustment is normally accomplished without the necessity for workers to leave a particular industry or region. As such, the possibility of structural unemployment is reduced (Grimwade, 1989).

3.2.3 Analytical framework

The Grubel and Lloyed (G-L) Index is the mostly used framework in analyzing IIT and INT. Thus, this framework is presented in this sub-section and will be the one adopted by this study to analyze both INT and IIT.

3.2.3.1 The G-L Index

Balassa (1966) was the first to provide the numeric measurement of IIT when he proposed that IIT be measured by the extent to which exports of a given product are offset by imports of the same product. The following is his formulation:

$$INT_j = \frac{|x_j - m_j|}{x_j + m_j} \quad (1)$$

where x_j and m_j are the values of the export and import of a commodity j by a given country. In this formulation, INT_j is considered as inter – industry trade (INT) and shows trade in different products. The drawback of this equation is that in the absence of IIT, the index will be one and in the case where there is “perfectly matching” IIT, the index value takes a value of zero. As a result of this problem, this index has not been used in empirical research. This resulted in Grubel and Lloyd (1975) proposing a new method of measurement whose application has dominated empirical literature on IIT/INT.

The G-L index has been the workhorse on empirical literature dealing with IIT measurement. The index computes the proportion of IIT as a component of balanced trade which shows the overlap of export and import trade between country i and its j partners for a given sector k . The formulation used by Grubel and Lloyd’s (1975) study takes the following algebraic format:

$$IIT = \frac{(x_i + m_i) - |x_i - m_i|}{(x_i + m_i)} = 1 - \frac{|x_i - m_i|}{(x_i + m_i)} \quad (2)$$

where IIT is the intra-industry trade index, and the other variables, x_i and m_i , are as defined earlier.

The above G-L index has however been criticised mainly on two grounds. Firstly, the index has been criticised as having bias due to the fact that it ignores trade imbalances, and secondly, the bias problem of selecting the ‘correct’ data disaggregating level (Isemonger, 2000).

To overcome the above drawbacks levelled against the G-L index, a number of adjusted G-L indexes have been devised, and among them is one by Aturupane *et al.* (1999). The adjusted G-L index, according to Aturupane *et al.* (1997) and modified to represent sectoral flows is represented by the equation:

$$IIT_{ijk} = \left\{ 1 - \left[\frac{\sum_i |x_{ijk} - m_{ijk}|}{\sum_i (x_{ijk} + m_{ijk})} \right] \right\} \times 100, \quad (3)$$

where,

IIT_{ijk} = intra-industry measure of trade between country i and country j in sector k ;

x_{ijk} = exports from country i to country j in sector k ;

m_{ijk} = imports to country i from j countries in sector k

The index for IIT in the above formula varies from 0 and 100, where complete INT is indicated by 0, and complete IIT indicated by 100⁹. Since extreme values (that is, 0 and 100) are rare to find in actual computations of the index, values of 50 and below will be interpreted as indicating INT and values above 50 indicate IIT.

Once a given sector is known to be IIT driven, then the next step will be to further decompose it into either horizontal IIT (HIIT) or vertical IIT (VIIT). Although there are a variety of formulae used to define both HIIT and VIIT, the formula by Greenaway *et al.* (1995) will be the one adopted in this study. Thus, HIIT will be

⁹ In the case where equation (3) is not multiplied by 100, complete INT will be indicated by 0, and complete IIT will be indicated by 1.

assumed to exist for the simultaneous export and import of a given product in sector k between country i and its trading j partners when the following formula criterion is satisfied:

$$1 - \alpha \leq \frac{UVx_{ijk}}{UVm_{ijk}} \leq 1 + \alpha \quad (4)$$

VIIT exists when the following criterion is met:

$$\frac{UVx_{ijk}}{UVm_{ijk}} < 1 - \alpha \quad \text{or} \quad \frac{UVx_{ijk}}{UVm_{ijk}} > 1 + \alpha \quad (5)$$

where, UV = unit value; and x_{ijk} and m_{ijk} are as defined before.

The distinction into either HIIT or VIIT is made possible by the use of relative unit values (UV) of exports and imports. According to Aturupane *et al.* (1997), the assumption underpinning the use of UV is that relative prices are able to reflect quality differences, implying that “a variety sold at a higher price must be of higher quality than a variety sold more cheaply” (Greenaway, 1995).

In terms of interpretation, trade in any sector’s products is considered to follow HIIT (variety difference) in the case where relative unit values fall within the range defined by Equation (4). By contrast, goods are considered to follow VIIT (quality difference) when Equation (5) is satisfied. Although in literature there are mainly two values of α , i.e. 0.15 and 0.25, which have been used, this study will follow Abde-el-Rahman (1991), Greenaway *et al.* (1995) and Aturupane *et al.* (1997) by employing a dispersion value of 15% (that is, $\alpha = 0.15$) for the analysis, and use $\alpha = 0.25$ for robustness check. Over and above the robustness check using this value of $\alpha = 0.25$ as done in the above studies, this thesis will introduce three other values of α , which are 0.10, 0.20 and 0.30 to increase robustness.

3.2.4 Empirical model

Following from the framework presented above, the study will empirically investigate whether the sectoral exports are either IIT or INT driven, by computing the index in Equation (3). If a given sector's exports are discovered to be IIT, then Equations (4) and (5) will be employed to further ascertain whether they follow HIIT or VIIT models.

3.3 Analyzing the determinants of export trade

3.3.1 Introduction

One of the most successful and popular empirical trade device used for more than four decades is the gravity trade model. Applied to a wide range of variety of goods and factors moving over regional and national borders under different circumstances, it has been used to study a whole range of spatial interactions in economics (Ghosh and Yamarik, 2004). In particular, the model has been applied to study the determinants of bilateral trade flows and to assess the impact of various forms of regional economic integration, such as the creation of a customs union as well as the adoption of a common currency (Cieslik, 2007 and Marques, 2008).

Gravity models utilize the gravitational force concept as an analogy to explain the volume of trade, capital flows, and migration among different countries of the world. The basic tenet of the gravity trade model is derived from the gravity theory in physics. Ideally, a flow is regarded as the resultant of the interaction between two objects, and these two objects in international trade are the exporting and importing countries. The 'masses' of the countries are their economic sizes from which potential trade takes place. Thus, the larger the economies of the relevant countries, the larger will be the trade among these countries. However, distance between the trading partners causes resistance to trade due to factors such as transport costs and time spent during shipment. Other potentially relevant factors which may impede trade are import duties, border controls, quantity restrictions, language, and in the contemporary world, subtle non-tariff barriers (NTBs) such as sanitary and

phytosanitary standards (SPS) are also common trade hindrances. Overall, in many empirical applications, gravity models have proved to have significant explanatory power, leading Deardorff (1998) to refer to these models as the “fact of life.”

The popularity of the gravity equation as one of international trade ‘toolkit’ used in examining foreign trade patterns among countries stems from its advantages. Paas (2000) pointed out two advantages of using this model. Firstly, the data needed for the model is in most cases easily accessible and reliable. This advantage is especially favorable when modeling developing (e.g. African) countries given that lack of reliable and internationally comparable statistical data is the most significant obstacle in modeling such economies. Secondly, the theoretical considerations for using these models to explore international trade flows are well discussed and developed in literature (Tinbergen, 1962; Linnemann, 1966; Bergstrand, 1985, 1989 and 1990; Deardorff, 1995 and 1998; Evenett and Keller, 1998; Evenett and Keller, 2002; Mathur, 1999).

Another advantageous dimension of the gravity trade model is its ability to explain more than trade theory. Trade theory, as a canon of trade, explains why countries may trade in different products but does not explain why some countries’ trade links are stronger than others and why the level of trade between countries tends to increase over time. This emphasizes the limited applicability of trade theory in explaining the size of trade flows. Consequently, while trade theory can explain why trade occurs, it cannot explain the extent of trade. However, the gravity model, by allowing more factors to be taken into account, can explain the extent of trade as an aspect of international trade flows.

3.3.2 The gravity model analytical framework

Whilst the gravity model has been used in a number of fields of studies such as human migration and investment flows across countries, its application in international trade seems to dominate its overall use. The gravity trade model borrows from Isaac Newton’s (1687) “Law of Universal Gravitation” which postulates that the force of attraction, F_{ij} , between two separate entities i and j is a positive function of the

entities' respective masses, m_i and m_j , and inversely related to the squared distance, d_{ij}^2 , between the objects. This law is formalized as:

$$F_{ij} = G \frac{M_i M_j}{D_{ij}^2} \quad (6)$$

where:

- F_{ij} = the force of attraction,
- M_i and M_j = are the respective two countries' masses,
- D_{ij}^2 = the distance between the two objects, and
- G = a gravitational constant depending on the units of measurement for mass and force.

In the mid 19th century, Carey (Principles of Social Science, 1858 -1859, as quoted in Pass (2000)) observed the presence of the gravitational force in social phenomena, alluding to the fact that the force was in direct ratio to masses and inverse to the distance. The application of the gravity models in social sciences considered the mass as an economic region or country.

In the second quarter of the 20th century, Reilly (1929) proposed a law of retail gravitation, which among other things, stated that a given city attracts retail trade from customers in its vicinity in proportion to the hinterland's size (population) and in inverse proportion to the square of the distance separating the customer from the centre of the city. The frontier separating the market areas of two cities i and j competing for customers in the hinterland is thus defined as the locus of points for which:

$$\frac{P_i}{d_{xi}^2} = \frac{P_j}{d_{xj}^2} \quad (7)$$

where:

d_{xi}^2 and d_{xj}^2 = distance of cities i and j , respectively, from any point x on the boundary,

P_i = population of city “ i ”, and

P_j = population of city “ j ”.

Nearly twenty years latter, Stewart (1948) pointed out that similar forces might underlie the interaction between social units, such as people. Following the above mentioned Newtons’ gravitational force formula, Stewart defines the demographic force as the result of a constant times the product of two masses divided by the squared distance (d_{ij}) separating the masses. Taking P_i and P_j as representing the respective populations of the cities (similar to Reilly’s approach), the demographic force, DM_{ij} , is mathematically presented as follows:

$$DM_{ij} = G \frac{P_i P_j}{d_{ij}^2} \quad (8)$$

where,

G = a constant corresponding to the gravitational constant.

In empirical studies mass has been measured in a variety of ways, depending on the subject matter being studied. For instance, Isard and Freutel (1954) used income as a measure of mass and developed the concept of income potential to parallel Stewart’s concept of population potential. On the other hand, if migration is to be studied, it may be appropriate to use employment or income as a more significant index of mass than population. Likewise, when the marketing problem for manufactured products is being assessed, the total volume of retail and wholesale sales tends to be the most appropriate and significant measure of mass. Thus, it can be argued that the measure of mass depends on the problem to be studied, available data, and related considerations. According to Isard (1960), the array of possible measures range from total investment in facilities, number of car registrations, hospital beds, investment in tractors and farm equipment, commodity output, value added in manufacture, etc (Disdier, 2009). Therefore economic mass is weighted differently.

In a like manner, distance has been and can be measured in a number of ways related to various aspects of the problem and the state of transport technology. For example, in a metropolitan traffic study, travel time is the most appropriate measure of distance. The other possible measures of distance which have been used in empirical literature are mileage along a specific transport route (waterway, highway, airline, and railway), fuel (energy) in transportation, number of gear shifts, stops, etc. (Isard, 1960).

Summarizing the historical development of the gravity theory, it can be noted that the theory has primarily been centered on fields where distance plays a significant role. To this end, the gravity theory has proven to be useful in describing social phenomena in space such as population migration, flow of goods, money, and information, traffic movement and tourist travel.

One can therefore specify a gravity theory to describe the above social phenomena following Nijkamp's (1975) formulation:

$$t_{ij} = K o_i^{b1} d_j^{b2} f(s_{ij}) \quad (9)$$

where,

- t_{ij} = the volume of flows between two points,
- K = a constant,
- o_i = volume of flows from the points of origin,
- d_j = volume of flows at the points of destination,
- $b1, b2$ = weighted geometric averages of o_i and d_j respectively,
- $f(s_{ij})$ = distance friction, a decreasing function of s_{ij} .

In analyzing trade, the basic gravity trade model which has been used in empirical work over the years was originally specified by Tinbergen (1962) and Poyhonem (1963) as follows:

$$Trade_{ij} = \alpha \frac{GDP_i^{\beta_1} GDP_j^{\beta_2}}{(D_{ij})^{\beta_3}} \quad (10)$$

where $Trade_{ij}$ represents bilateral trade between countries i and j , while GDP_i and GDP_j denote countries i and j 's respective gross domestic products. D_{ij} is used as a proxy of bilateral distance between the two trading countries. In the formula above, the α and β 's are parameters and the signs of β_1 and β_2 are expected to be positive, while that for β_3 will have *a priori* negative sign. Thus, comparing Equations (6) and (10), it can be seen that in analyzing trade using the same gravity principle, the entities in Equation (6) are replaced by a pair of countries in Equation (10), while the countries' masses in Equation (6) are proxied by the respective GDP in Equation (10) with distance replaced by a variable representing resistance, which in most cases is the actual distance between the trading countries.

Rewriting Equation (10) in logarithmic format, a linear version of the model can be represented as follows (Batra, 2004; and Ghosh and Yamarik, 2004):

$$\ln(Trade_{ij}) = \alpha + \beta_1 \ln GDP_i + \beta_2 \ln GDP_j - \beta_3 \ln(D_{ij}) + \mu_{ij} \quad (11)$$

where α , β_1 , β_2 and β_3 are coefficients to be estimated. The disturbance error term (μ_{ij}) captures random events which may have an impact on bilateral trade between the two trading countries and is assumed to be stationary, with a mean of zero and a constant variance. Thus Equation (11) is the core gravity equation which has been used in all empirical studies, albeit with added right hand side (RHS) variables, with each RHS variable added depending on the particular facet of trade being analyzed, the objectives to be achieved and availability of data. In this basic gravity theory, a positive correlation is expected between trade and GDPs, while a negative relationship will be expected between trade and distance.

3.3.2.1 Theoretical foundations of gravity models

Despite the extensive use of gravity equations in economics until recently they have tended to lack strong theoretical foundations. Currently, due to theoretical

developments that have been done since Anderson's (1979) work, the gravity equation in its basic form can be derived from a variety of competing theoretical models, based on either neoclassical or monopolistic competition assumptions. As pointed by Frankel (1998:2) the gravity equation has recently "gone from an embarrassment of poverty of theoretical foundations to embarrassment of riches".

There are broadly two competing models of international trade that provide theoretical justification for the gravity model. They are the differentiated products model and the Heckscher-Ohlin (H-O) model. Anderson (1979) popularized the differentiated product model. His point of departure was the use of the Armington assumption that products were differentiated by country-of-origin. Thus, Anderson (1979) demonstrated how to derive a gravity equation by employing the properties of a Cobb-Douglas expenditure system in a case where each commodity was produced by one country. Helpman's (1987) theoretical exposure also assumed monopolistic competition and product differentiation among firms in all industries rather than countries. The monopolistic competition approach was viewed as a stylish way of endogenizing product differentiation and explaining formally the basis for the Armington assumption. The main purpose of monopolistic competition in Helpman's (1987) model was to ensure that different countries specialize in production of different varieties of differentiated products due to existence of economies of scale at the firm level which enhance the incentives for foreign trade.

Empirically, Krugman and Helpman (1985) identified the relationship between the bilateral trade flows and the product of two trading countries' GDPs by utilizing the differentiated products trade model. According to Krugman and Helpman, under the imperfect substitute model, where each firm produces a product that is an imperfect substitute for another product and has monopoly power in its own product, consumers show preference for variety. When the size of the domestic economy (or population) doubles, consumers increase their utility, not only in the form of greater quantity but also in the form of greater variety. International trade can provide the same effect by increasing consumers' opportunities for even greater variety. According to Linder (1961), when two countries have similar technologies and preferences, they will naturally trade more with each other in order to expand the number of choices available for consumption. The association between the gravity equation and the

differentiated products model was empirically proven by Helpman (1987) when he applied his test on OECD countries' trade data. His results supported the argument that the gravity equation can be applied to the trade flows among industrialized countries where IIT and monopolistic competition are well developed.

In contrast, Deardorff (1995) has shown that the gravity trade model can be derived from several variants of the H-O model based on comparative advantage and perfect competition if it is properly considered. He found out that the absence of barriers to trade in homogeneous products causes producers and consumers to be indifferent to the trading partners, both domestic and foreign, so long as they buy or sell the desired goods. Based on this assumption, he derived expected trade flows that correspond exactly to the simple frictionless gravity equation whenever preferences are identical and homothetic. Hummel and Levinsohn (1995) conducted an empirical test with a set of non-OECD countries where monopolistic competition was not so plausible. Their results proved that the gravity equation was also efficient in explaining the trade flows among developing countries where inter-industry trade was dominant with scarce monopolistic competition. Their findings questioned the uniqueness of the product differentiation model in explaining the success of the gravity equation and proved that a variety of other models, including the H-O model, can serve as alternatives.

Evenett and Keller (1998) also emphasized that gravity prediction constitutes the most important result regarding the volume of international trade. They argued that little production is perfectly specialized due to factor endowment differences and that as long as the production is not perfectly specialized across countries, both models of the H-O and differentiated products are likely to account for the empirical success of the gravity equation. Deardorff (1998) and Feenstra (2004) have also noted the compatibility of the gravity equation with some structures of the H-O model, as well as the need for empirical evidence to distinguish among potential theoretical bases: product differentiation by country of origin; product differentiation by firm; and particular forms of the H-O-based comparative advantage. In each of these cases, the common denominator is complete specialization by countries in a particular commodity. Without this feature, Deardorf (1998) and Feenstra (2004) argued that bilateral trade tends to become indeterminate.

Following the aforementioned theoretical developments, it is generally accepted that a number of trade models are responsible for the empirical success of the gravity equation. While the H-O theory would account for the success of the gravity equation in explaining bilateral trade flows among countries with large factor proportion differences and high shares of inter-industry (the so-called ‘North-South’ trade), the differentiated product model would serve well in explaining bilateral trade flows among countries with high shares of intra-industry trade (the so called ‘North-North’ trade) in increasing returns with monopolistic competition. Thus, there is currently no single theoretical model which claims monopoly in explaining the success of the gravity model, as the exact theoretical model which best describes the empirical findings of the gravity model is still a matter of contention (Ram and Prasad, 2007).

Therefore, given that no single trade theory can claim monopoly in explaining the success of the gravity trade models, some of the possible theories are presented in Table 2 (See also Table A26 in Section 12 of the Appendices).

Table 2: Theories behind the gravity trade model

Type of model		Technology Differences (Ricardian)	Increasing Returns to Scale	Heckscher-Ohlin
1	Structural assumption: Identical homothetic demand, free trade etc	Technology differences with industry classes across countries	Increasing returns at the firm level, monopolistic competition, product differentiation	Homogenous goods and multi-country (large factor proportion differences)
	Consistent with absence of factor proportions differences?	Yes	Yes	No

2	Implication for Nature of Trade	Intra-industry trade (IIT) in goods with alternating technological superiority	Intra-industry trade (IIT) in product varieties with potentially identical technologies across countries	Inter-industry trade (INT)
	Consistent with trade in goods with identical factor requirements?	Yes	Yes	No

Source: Evenett and Keller (1998)

3.3.3 Empirical model

Although the basic gravity trade model used in most empirical studies has been developed from Equation (11), theoretical developments for sectoral gravity models are still scarce. One notable study which developed a theoretical sectoral gravity trade model is by Marques (2004) and is presented in Section 7 of the Appendices. From the model presented in the Appendices, sectoral gravity trade models for each of the three sectors, diamond, textile, and meat and meat products, can be represented by simplification of the formulations in Equations (A14) and (A15), as shown in Section 7 of the Appendices. To appropriately achieve this thesis' objective number 2, the sectoral gravity models estimated are further augmented and modified as shown in Equation (12). The model is "augmented" in that it includes a number of variables, besides the importer GDP, exporter GDP and distance, as presented in the basic gravity Equation (11), which account for other possible factors which may have an impact on bilateral trade. The model is also "modified" in the sense that it is formulated to take into account sectoral, as opposed to aggregate or total trade factors.

$$\begin{aligned} \ln X_{ijk} = & \alpha_0 + \alpha_1 \ln Y_{ik} + \alpha_2 \ln \left(\frac{K_{ik}}{L_{ik}} \right) + \alpha_3 PTA_{ijk} + \alpha_4 \ln Y_j + \alpha_5 \ln D_{ij} + \\ & \alpha_6 \ln P_j + \alpha_7 \ln \Pi_j + \alpha_8 \ln \text{exr}_{i,usa} + \varepsilon_{ij} \end{aligned} \quad (12)^{10}$$

where,

X_{ijk} = value of sector k export trade flow from Botswana (country i) to country j ,

Y_{ik} = is the GDP of sector k in country i (i.e., Botswana),

(K_{ik}/L_{ik}) = capital per unit of labour (capital labour ratio) in sector k in country i (i.e., Botswana),

PTA_{ijk} = Preferential trade arrangement (PTA) dummy variable between Botswana (i.e., country i) and partner j in sector k ,

Y_j = GDPs of importing countries,

D_{ij} = physical distance from the economic centre of country i (i.e., Gaborone, Botswana) to that of country j ,

Π_j = inflation of importing country,

$\text{Exr}_{i,usa}$ = exchange between Botswana Pula and US dollar

α_0 = a constant.

Sector specific preferential trade arrangements (PTA) dummies are:

- Diamond sectoral exports– UK De Beers Diamond Trading Company dummy
- Textile sectoral exports – AGOA dummy
- Meat and meat products exports – Cotonou trade agreement dummy¹¹

¹⁰ Where Variables Y_{ik} , K_{ik}/L_{ik} and PTA_{ijk} are sector specific

Given the empirical evidence to be presented in Chapter 5, which shows that the diamond sector is IIT driven while both textile, and meat and meat products are INT, the study is now at a position to extend Equation (12) and add new explanatory variables to capture both IIT and INT in the sectoral gravity models. The resulting equation thus becomes:

$$\ln X_{ijk} = \alpha_0 + \alpha_1 \ln Y_{ik} + \alpha_2 \ln \left(\frac{K_{ik}}{L_{ik}} \right) + \alpha_3 PTA_{ijk} + \alpha_4 \ln Y_j + \alpha_5 \ln D_{ij} + \alpha_6 \ln P_j + \alpha_7 \ln \Pi_j + \alpha_8 \ln exr_{i,usa} + \alpha_9 IIT(INT)_{ik} + \varepsilon_{ij} \quad (13)$$

The rationale for including either an IIT or INT variable in the gravity trade model is to see whether this new variable in Equation (13) can lead the study in concluding that the IIT/INT variable is also an important determinant of Botswana's sectoral exports or not. For the diamond sector where exports are IIT driven, a representative variable to capture product differentiation trade theory will be put where there is IIT (INT) in Equation (13). Literature, for instance, Hafbauer (1970) suggests the use of either the number of traded Harmonised Commodity Description and Coding System (HS) lines within the sector or the Hafbauer index, which is presented in Section 8 of the Appendices, as possible variables to represent IIT, or product differentiation theories. In the case of the textile, and meat and meat products sectors which are INT driven, as will be empirically tested in Chapter 5, an INT variable will replace IIT (INT) in Equation (13). The revealed comparative advantage index (RCAI), presented in Section 9 of the Appendices, has been one of the possible variables suggested in empirical literature (for instance Kandogan, 2003a) to capture INT or H-O trade theories in gravity equations. Thus, two gravity model equations for each of the three sectors will be estimated based on Equations (12) and (13). Following IIT/INT explanations in Chapter 5, the coefficient of IIT/INT is expected to be positive.

¹¹ In each of the three sectoral gravity models, there will be only one PTA dummy, hence no problem of dummy traps.

3.3.3.1 Selected factors that determine exports

With regards to factors that determine exports, as outlined in Chapter 2, there is a pool of potential variables, besides the three core variables of importer GDP, exporter GDP and distance, which explain direction of exports. In fact, Ghosh and Yamarik (2004) indicate that there are around 48 factors that have been used in gravity trade model literature which seek to explain the direction of exports¹². This section therefore discusses some of the determinants found in literature and especially those that will be used to explain Botswana's sectoral exports.

3.3.3.1.1 GDP or economic mass

Overall, in the gravity models, trade is assumed to occur when domestic production is not equivalent to domestic demand. The sectoral GDP (Y_{ik}) of Botswana measures productive capacity in a given sector and can also be considered as a proxy for the range of product varieties available, which increase the availability of exports in that sector. The GDP values for Botswana's mining, manufacturing and agriculture sectors will approximate the sectoral GDPs used in the gravity trade model equation for diamond, textile, and meat and meat products, respectively. The GDP (Y_j) of the importing country measures absorptive capacity and represents potential demand for imports from Botswana. Thus, α_1 and α_4 in both Equations (12) and (13) are expected to have positive signs.

3.3.3.1.2 Distance between trade partners

Head (2003) alluded to the fact that distance in gravity models acts as a sort of tax "wedge," imposing trade costs, and resulting in lower equilibrium trade flows. Ram and Prasad (2007) consider the following five factors as the reasons for inclusion of distance as an explanatory variable in gravity trade models:

- distance acts as a better measure of transport costs,

¹² This thesis has experimented with a number of possible variables in each of the three sectoral gravity models and the ones presented in this section are the most significant ones.

- the time which elapsed due to shipment can be indicated by distance, with the probability of surviving intact for perishable goods diminishing with time in transit,
- distance is normally correlated with transaction costs in the form of searching for trading opportunities and the establishment of trust between potential trading partners,
- synchronization costs are positively related to increasing distance, i.e., the synchronization costs incurred in cases where production factories combine multiple inputs from different countries in order to prevent delays or emergence of bottlenecks; and
- distance captures cultural diversity, which can retard trade by inhibiting communication, differences in negotiating styles, etc.

Thus, as distance between trading partners increases, export flows are expected to decline. In this case, theory predicts a negative relationship between export trade and distance.

3.3.3.1.3 Factor endowment

The capital-labour ratio (K_{ik}/L_{ik}) characterises the sectoral technologies and their relative factor requirements. This variable captures the relevant supply-side characteristics of production which normally impacts on exports. A higher capital-labour ratio is expected to affect sectoral exports positively.

3.3.3.1.4 Preferential trade arrangement (PTA) dummy

The PTA_{ijk} is a vector of specific PTA dummies between Botswana and a trading partner in sector k. For the three sectors under study, the specific PTA dummies, as will be explained in Chapter 4, are the diamond's UK De Beers Diamond Trading

Company dummy, the textile African Growth Opportunities Act (AGOA) dummy, and the Cotonou Agreement's meat and meat products dummy. In all three cases, the dummy will take a value of 1 if a trading partner is a member described by the dummy and 0 otherwise. According to Carrere (2006); Silva and Tenreyro (2006); and Jakab *et al.* (2001), inclusion of PTA variables in the gravity model is useful in assessing whether the PTA among member states has been trade-creating or trade-diverting. In this thesis, membership to a trade preferential arrangement is expected to generate a significant increase in trade given that countries enter into such arrangements mostly with the aim of increasing trade among them. Thus, coefficients of all the trade related arrangement dummies are expected to be positive.

3.3.3.1.5 Population

Population is used as a measure of country size, and larger countries, as measured by population, are assumed to have more diversified production, a large domestic market, a high probability of self-sufficient and less need to trade (Nilsson, 2000). A negative correlation will be expected between population and export trade in such a scenario. However, Bergstrand (1985) pointed out that there is an inconsistency in this argument, as larger populations allows for economies of scale which are translated into higher exports resulting in a positive relationship between population and export trade. Therefore, the sign of the coefficient of the exporting country would be indeterminate. At the same time, a large population in the importing country can affect imports either negatively or positively due to the same reasons as given for the exporting country.

3.3.3.1.6 Inflation

Inflation measures the purchasing power of the importing countries. The sign for this variable is indeterminate. Both negative and positive signs are supported by theories from international finance. When an importing country is in an inflationary period, it means that citizens will try to avoid domestic inflation by importing (with the assumption that world imports prices will be relatively lower compared to domestic

prices). In this case, a positive relationship between inflation and imports will be expected. On the other hand, inflation means that most consumers will scale down their purchases, including imports, as their real purchasing power falls, thus resulting in a negative relationship between imports and inflation.

3.3.3.1.7 Exchange rate

To capture the impact of depreciation of Botswana pula against the US dollar, the gravity trade models for each of the three sectors were augmented by including exchange rate of dollar in terms of Botswana pula¹³ following Bergstrand (1985)¹⁴, and Soloaga and Winters (2001), among others. The relationship between exchange rate and export trade is expected to be positive. That is, an increase in the exchange rate between US dollar and Botswana pula implies depreciation (weakening) of the latter currency. This will encourage exports from the latter country as importers will be able to import more products with the same US dollars than if there was no depreciation.

3.3.4 Estimation Procedure

This subsection provides the procedure that is followed in estimating panel data econometrics in general and the gravity trade model in particular.

3.3.4.1 Pooled versus individual fixed effects

Generally, panel data regression differs from time-series or cross-section regressions in that its econometric representation contains double subscript on its variables. An illustrative representation can take the following form:

$$y_{it} = \alpha + X'_{it} \beta + \mu_{it}, \quad i = 1, \dots, N; t = 1, \dots, T \quad (14)$$

¹³ Direct exchange rate is used, that is, the rate at which US\$1 is converted into Botswana pula (BWP)

¹⁴ An exchange rate variable has first been formally introduced into the gravity equation by Bergstrand (1985).

with i denoting households, individuals, firms, countries, etc., and t denoting time. The i subscript, therefore, represents the cross-section dimension, whereas t denotes the time-series dimension. In Equation (14), α is a scalar, β is a $k \times 1$ matrix vector and X_{it} is the i^{th} observation on K explanatory variables. In empirical literature, most of the panel data estimations (e.g. Jayasinghe and Sarker, 2007; Marques *et al*, 2005; and Rojid 2006) utilize a one-way error component model for the disturbances, with

$$\mu_{it} = \psi_i + \varepsilon_{it} \quad (15)$$

where ψ_i denotes the unobservable bilateral country-level individual specific effects and ε_{it} denotes the remainder disturbance which is assumed to be independently and identically distributed over i and t , with a mean of zero and constant variance, σ_{ε}^2 . In the sectoral gravity equation, y_{it} (in Equation (14)) and represented by X_{ijk} in the actual empirical model Equations (12) and (13), will measure the value of exports from a given sector, whereas X_{it} (in Equation (14)), represented by applicable right-hand variables in Equations (12) and (13), contains a set of variables such as the respective GDPs and distance. On the other hand, ψ_i (in Equation (15)) (or ε_{ij} in Equations (12) and (13)) is time-invariant and accounts for any individual country-specific effects that are not included in the regression, such things as race, language, etc. The remainder disturbance, ε_{it} in Equation (15) is considered as a well-behaved white noise error term, with mean of zero and a constant variance.

In terms of the estimation procedure, one may assume that there are no individual country-specific effects present in the panel, thus assuming all the countries in the panel to be the same, and that the estimation will have one coefficient for μ_i . If such a route is chosen, a pooled estimation will be the model implemented. The second possible option is to estimate an equation where individual country-specific effects are assumed to be present in the panel, in which case a fixed effects model will be estimated. Nevertheless, in order to decide whether to estimate a pooled model or a fixed model, an econometric test for the joint significance of the individual effects to be estimated has been done. This is the so-called F-test and the results are presented in Section 5 of the Appendices. Given that the F-tests for all the three sectoral equations have rejected the null hypothesis of no individual effects in favour of individual fixed

effects, and also that Botswana's trading partners in all the three sectors investigated are heterogeneous, the analysis of this thesis will therefore be based on the fixed effects model (FEM) estimation procedure.

3.3.4.2 Fixed Effects Model (FEM) versus Random Effects Model (REM)

There are two different models which can be used to estimate the individual country specific effects, i.e., the FEM or the REM. Given these two possible estimation models, a decision regarding the treatment of these effects, either as fixed or random has to be made. Following Baltagi (2005:15), "the random effects model (REM) is an appropriate specification if we are drawing N individuals from a large population". On the other hand, Egger (2000:26) argues that the FEM model is appropriate when estimating trade flows between a predetermined selection of nations. Since this research is concerned with trade flows between Botswana and its main selected trading partners as opposed from drawing N trade partners from a large population, the FEM will be more appropriate than the REM specification for capturing the country-specific effects. As such, the FEM will be used in this thesis estimation.

3.3.4.3 Possible Endogeneity

Possible issues of endogeneity might involve the use of both revealed comparative advantage index (RCAI), in place of IIT/INT, and GDP on the right hand side of both Equations (12) and (13). This section addresses these issues. There should be no endogeneity problem emanating from the use of the RCAI since, although the RCAI is based in part on exports, all the information in it is based on the composition and not on the level of exports. Thus, the dependent variable has information on the level of exports which is different, with the information contained in the RCAI.

Although economic size (GDP) is treated as an exogenous variable in Equations (12) and (13), there is however theoretical and empirical support for the impact that exports can have on income. The possibilities of endogeneity of these variables, therefore, cannot be denied. To resolve this potential problem, Cyrus (2002), among

others, suggested the use of instrumental variables (IV), such as factor accumulation variables physical capital, human capital, and labor accumulation rates and population, as instruments for income since these instrumental variables are assumed to be uncorrelated with the error term in a gravity regression. Despite suggestions to use instrumental variables, most authors (including Cyrus 2002) have found no greater improvements in terms of results when one uses instrumental variables. This thesis has experimented with IV such as physical capital and arrived at the same conclusion, that is, the results do not change when IV are used in place of GDP.

3.3.4.3.1 Hausman test for exogeneity and misspecification

For this thesis, the Hausman test for both exogeneity and misspecification was conducted for each of the three sectoral gravity model formulations and the results are presented in Section 6 of the Appendices. Suffice to say that in each of the three gravity formulations, the null hypothesis of no misspecification (or no correlation between the dependent variable and regressors) is not rejected. The conclusion therefore is that the X-regressor in each of these three gravity models is exogenous, thus suggesting that there is no misspecification problem

3.3.4.4 Univariate characteristics of variables

The general procedure to be followed, according to Van Eyden (2007) and Baltagi (2005), in a case where a panel has enough time-series length, i.e., a panel with time length (T) of at least 10 years, is that the variables should be tested for stationarity before estimation.

Investigation of the univariate characteristics of the data which entail panel unit root tests is generally important given the fact that the unit root test is the first step encouraged in the determination of a potentially cointegrated relationship between variables. Normally, in the case where all the variables used in the estimation are stationary, then traditional estimation methods can be used to estimate the relationship between variables. On the other hand, in the case where variables are non-stationary, a

test for cointegration will be required. There are basically six potential panel unit root tests that can be employed for a panel stationarity test and these are summarized in Table A19 of Section 10 of the Appendices.

Given that the study's panel covers 1999 to 2006, which is 8 years, and less than the minimum panel length of 10 years required for unit root tests (Eyden, 2007 and Baltagi, 2005), this study will not perform panel root tests.

3.3.5 Limitations of the gravity model

Despite its celebrated empirical success, the gravity trade model has not been immune to criticism. Although initially the model's theoretical foundations were not strongly grounded, currently, the model's main regressors, namely income (mass) and distance, are actually supported by a wide range of theoretical models (Deardorff, 1998). Nevertheless, due to theoretical developments and using Frankel's (1998, p. 2) words, the gravity equation has recently "gone from an embarrassment of poverty of theoretical foundations to embarrassment of riches", thus settling the theoretical doubt of the model.

On the modeling side, Cyrus (2007) criticized the specification of the gravity models, especially when using Ordinary Least Squares (OLS) estimation techniques. The source of this specification problem is that 'the causality between income and trade is not clear-cut'. For instance, in the estimation, the gravity equation suggests that high income causes high trade, but he argued that perhaps it is trade that instead causes income to be high. At the same time the use of such policy measures as free-markets may push up both income and trade. In that case, as Cyrus argued, the gravity equation is mis-specified, for income will be correlated with the error term in the regression, so that the ordinary least squares will not provide consistent estimates. To this end, Cyrus (2007) suggested the use of instrumental variables such as factor accumulation variables physical capital, human capital, and labor accumulation rates, population and land area for size as instruments for income since these instrumental variables are assumed to be uncorrelated with the error term in a gravity regression. Despite suggestions to use instrumental variables, a number of authors (e.g., Ram and

Prasad, 2007; and Cyrus, 2007) have found no greater improvements in terms of results when one uses instrumental variables.

Another drawback which is important especially when using the model for analyzing potential export destinations emanates from the fact that historical data is used. It follows that should any structural break happens, for instance in the form of new innovations which result in other countries becoming cheaper competitors to a country's exports, discoveries of say large diamond reserves in another country¹⁵, an import ban (by importing countries), establishment of import substitution industries (by importing countries) and an unexpected change in demand (against an exporting country's product) in particular export destination countries, will render the use of the gravity model's ability to predict export destination potential off trajectory. In these cases, simulated export potential destinations from the gravity model will be misleading.

Despite the above criticisms leveled against the gravity theory, overall, the advantages of using it have outweighed its weaknesses, thus cementing its continued widespread application in empirical trade work.

3.4 Analyzing export destinations with unrealized potential

This section of the thesis relies entirely on the gravity equations to be estimated in Chapter 6, whereby those estimations are considered as a first step, with some of the results further employed in analyzing potential export destinations in the second stage of analyzing the unrealized export trade potential. In terms of methodology and procedure, the estimated coefficients from gravity trade model estimations in Chapter 6 will be used to investigate Botswana's trade patterns in general and with some regional groupings in particular.

¹⁵ New discoveries can happen. For instance, although South Africa used to be the country with largest gold reserves for more than 100 years, this has changed following large discoveries of gold reserves in China in 2007.

3.5 Data sources and description

3.5.1 Data sources

All export series in US dollars (USD) used for the empirical analysis in this study are obtained from South Africa's Trade and Industrial Policy Strategies (TIPS) database. This database provides a comprehensive interactive database for all trade categories in the Harmonized Commodity Description and Coding System (HS) from 2 digit level (HS-2) to 6 digit level (HS-6). The database shows the export amounts in USD as well as in the local currency (Pula), the various export destinations, and the export partners' shares in the country's total exports of a given product line. The data on distance are from the following website: www.timeanddate.com. Population data and national GDP series are from the International Monetary Fund's (IMF) World Economic Outlook (WEO). Botswana's Central Statistical Office (CSO) is the source for sectoral employment, sectoral GDP and sectoral capital investment data.

3.5.2 Descriptive statistics

Through provision of simple summaries of any information sample or population data, descriptive statistics are mainly used to describe the basic features of the data in any given research or study. By nature, descriptive statistics simply describe what is or what the data shows and no conclusion or inference can be made beyond the immediate data alone. In other words, descriptive statistics are employed to present quantitative descriptions in a manageable form. The descriptive statistics for all the variables used in each of the three sectoral analyses are shown in Table A20 through to Table A22 of the Appendices.

3.5.3 Correlation matrix

The correlation matrix, or table, shows the strength or degree of linear association of one variable with itself and/or between two variables (Gujarati, 2003). As shown in

Tables A23 through to A25, the entries in the main diagonal, i.e. those running from the upper left-hand corner to the lower right-hand corner, give the correlation of one variable with itself, which is (and should be) always 1 by definition. On the other hand, the entries off the main diagonal are the pair-wise correlations among X variables. In terms of interpreting the off-diagonal pair-wise correlation, high correlation figures suggest severe collinearity problems, while low correlation values indicate absence of the collinearity problem.

The correlation matrix for the data used in the three sectoral equations as presented in Table A23 through to Table A25 of the Appendices indicates absence of collinearity among the variables used in each of the three gravity equations. Absence of collinearity problems is evidenced by low off-diagonal pair-wise variables which are generally below 0.5 in each of the three sectoral correlation matrix tables.

3.6 Conclusion

In this chapter the methodologies used to achieve the objectives of this study have been discussed. The Grubel and Lloyd (G-L) Index has been presented as the analytical model that will be used to investigate the extent to which Botswana's sectoral exports are either inter-industry trade (INT) or intra-industry trade (IIT) driven. In analysing the determinants of these sectoral exports as well as export destinations with untapped potential, the gravity trade model will be used to achieve this type of investigation.