THE EMPIRICAL ANALYSIS OF ECONOMIC INTEGRATION:
A LITERATURE SURVEY

Muhittin KAPLAN*
Muhsin KAR**

Abstract:

Regional economic integration, which could take in different forms ranging from a free trade agreement to an economic and monetary union, could provide a number of opportunities to a member country and, therefore, there has been an increasing interest in the regional economic integration in the last half century among the developed and the developing countries. The aim of this paper is to critically review and evaluate the applied literature on the regional integration in the framework of the gravity model.

Keywords: Welfare effects, Economic integration, Trade creation and trade diversion.

1. Introduction

In recent years, an interest in the study of economic integration among nations has re-emerged in line with developments in the global economy and the geo-political environment. Such developments include the
THE EMPIRICAL ANALYSIS OF ECONOMIC INTEGRATION

underlying dynamics of globalisation encompassing corporations, markets and capital as well as the move towards closer international co-operation through the establishment of trade blocs and policy co-ordination. Economic integration occurs when two or more countries link their economies through the removal of trade barriers under specific integration agreements - commonly known as preferential trading arrangements (PTAs). Such arrangements represent various levels of economic integration by eliminating, to varying degrees, barriers among member countries and applying a common-tariff structure to trade with other non-member countries.

PTAs take a variety of forms depending on their purpose, scope and ambition. For analytical purposes, four types of the PTAs corresponding to four (ascending) levels of integration are distinguished (El-Agraa, 1998:1-30; Ertürk, 1993; Dartan and Akman, 1998; Dura and Atik, 2003). The first level is a free trade area (FTA) which is a type of the PTA that involves removing tariffs and quotas on all trade flows between members while leaving free for each country to retain tariffs on trade with non-member countries. The next level, a customs union (CU), arises when the member countries collectively apply a common tariff structure to all trade flows from non-member nations. A common (or single) market represents a still higher level of integration by allowing the free flow of goods, services, people and capital among the member countries. Finally, an economic and monetary union (EMU) exists when the member countries adopt a single currency in conjunction with monetary and economic policy co-ordination.

The research literature distinguishes between four main channels through which economic integration affects the welfare of the member countries (Baldwin, 1994:5-10). These are: (1) static-allocation effects arising from specialisation based on comparative advantages, enlargement of the market, increased competition, and changes in the terms of trade; (2) dynamic-growth effects which lead to increased output as a result of economies of scale, technology spillover effects, and external economies of scale; (3) macroeconomic effects stemming from the effects of integration on growth, inflation, employment, and the balance of payments and (4) location effects arising from the impact of economic integration on the spatial allocation of resources.

The theory of economic integration depends upon the theory of customs union. Since the impacts of the customs union on the economy could be assessed and easily modelled. The development of the theory of customs unions has tended to focus almost exclusively on the static effects of closer economic integration (Hine, 1994). The early development of the theory of
customs unions, which perceived custom unions as a movement towards free trade, concluded that customs unions would always raise welfare (Lipsey, 1960; Baldwin and Venables, 1995; Karakaya and Cooke, 2002). Later, Viner (1950) demonstrated that whether economic integration leads to welfare losses or gains depends on the relative strength of “trade creation” and “trade diversion” effects of customs unions. Trade creation takes place when relatively high cost domestic production in one country is replaced with lower-cost imports from a partner country, leading to welfare gains for both trading countries. Trade diversion occurs when a country switches its imports from a more efficient non-partner country to a less efficient partner country, leading to welfare losses. Consequently, there is no single, unambiguous, answer to the question of whether or not economic integration increases welfare. Given this, researchers have tended to concentrate more on identifying the conditions under which economic integration is more likely to lead to trade creation (Sodersten and Reed, 1994). Thus, it has been concluded that economic integration between countries will be more likely to result in trade creation if: (1) the degree of overlap between the bundles of goods produced by each member country is high; (2) the differences in production costs among countries in their common industries are large and (3) the level of tariffs is higher than before the creation of the union (Dartan and Akman, 1998:673-674). In the same way, it is more likely that economic integration will result in less trade diversion if: (1) the number of union members is big; (2) the initial level of trade relative to domestic activity is low and (3) competitive but potentially complementary countries form a union (Hine, 1994; Dartan and Akman, 1998:673-674).

Despite recent developments in the literature on the theoretical front, the empirical work on economic integration has, almost exclusively, been concerned with identifying the extent of trade creation and trade diversion effects of custom unions rather than actually quantifying the static-allocation, dynamic-growth and macroeconomic effects. The sizes of the trade creation and trade diversion effects are used to determine the welfare consequences of custom unions. There are two deficiencies in this approach to quantify the effects of integration. Firstly, the effect of integration is estimated without any reference to the channels through which the welfare is affected. Second, it is implicitly assumed that the pattern of trade is a good proxy for assessing the total gains from economic integration without paying attention to the effects on total output.

The objective of this paper is to review and evaluate the applied literature on the economic effects of the PTAs. In Section 2, we provide a
review of the theoretical underpinnings of the so-called gravity model approach to assessing integration. Section 3 focuses on the empirical literature dealing with the effects of economic integration and related issues. In Section 4, we turn to consider some of the key econometric issues which arise from the implementation of the gravity model in the empirical literature. Finally, Section 5 concludes with the suggestions concerning the different aspects of the subject of economic integration which require further attention.

2. The Theoretical Foundations of the Economic Integration: The Gravity Model

The importance of the gravity model stems from the fact that it has been useful in helping to explain bilateral trade flows as well as the welfare effects of economic integration. In addition, this model has also been widely used to shed light on various other aspects of inter-regional and international flows such as the trade flow effects arising from exchange rate mechanisms, labour migration, hospital patients, ethnic ties, linguistic identity, and international borders.

The gravity model of international trade was independently developed by Tinbergen (1962) and Pöyhönen (1963) to explain bilateral trade flows. In its simplest form, the model explains the amount of trade flow between two countries as an increasing function of their sizes and a decreasing function of transportation costs (distance) between them. In other words, it explains trade flows on the basis of the physical characteristics of the importing and exporting countries and the existence of any impediments to trade. Therefore, a so-called ‘gravity equation’ can be developed as a reduced form equation that is simply derived from export (or import) supply and demand functions representing the equilibrium of bilateral export supply to export demand.

In its simplest form the gravity equation can be written as:

$$x_{ij} = \alpha_0 + \alpha_1 y_i + \alpha_2 y_j + \alpha_3 n_i + \alpha_4 n_j + \alpha_5 d_{ij} + \alpha_6 p_{ij} + e_{ij}$$

where $x_{ij}$ represents the exports from country $i$ to country $j$; $y_i$ is the income of the exporting country and can be thought as the production capacity of the exporting country (i.e. a proxy for the capital stock of the exporting country); $y_j$ is the income of the importing country and acts as a proxy for the consumption capacity of the importing country; $n$ represents
the population of countries $i$ and $j$ respectively and proxies the size of the economy; $d_{ij}$ is the physical distance between countries $i$ and $j$ and $p_{ij}$ is a dummy variable that takes a value of one if both countries are members of a particular PTA and zero otherwise. Lower case letters show logarithm of variables throughout the text.

Although the gravity model has been successful in explaining a considerable amount of change in bilateral trade in empirical studies, it had been criticised because of its lack of theoretical foundations (Leamer and Levinsohn, 1995). However, more recent researches have indicated that a gravity type equation can be derived from any trade model (Deardorff, 1998).

In general terms, developments of the theoretical foundations of the gravity model can be divided into two groups: (1) those that assume full specialisation in production and (2) those which allow for incomplete specialisation. Although empirical models stemming from the first group assume full specialisation, the source of specialisation is different for different models. For example, specialisation arises from national product differentiation in the approaches followed by Anderson (1979), Bergstrand (1985) and Deardorff (1998) while economies of scale is the cause of specialisation in Helpman (1987) and Bergstrand (1989). Others researchers such as Davis (1995) and Eaton and Kortum (1997) have derived the gravity model by assuming there are technological differences across countries, similar to Deardorff (1998)'s factor endowment differences assumption. Indeed, in his comment on Deardorff (1998), Grossman (1998:29) states that 'specialisation - and not new theory or old theory - generates the force of gravity'.

In contrast, Evenett and Keller (1998) have shown that a gravity equation can arise with incomplete specialisation as well if there are just two countries in the trading alliance. Later, Keller (1998) extended this analysis to many countries.

Since the gravity model approach is derived from an eclectic theoretical perspective, interpretation of the coefficients will always, inevitably, be subjected to heated debate. However, Feenstra et. al. (2001), using the eclecticism of the gravity model, manages to differentiate between alternative theories of trade. In their study, they use the gravity model to test different trade theories based on the size and direction of the estimated coefficients and to demonstrate the significance of this approach in the testing of various theoretical propositions.
3. The Empirical Literature on Economic Integration

Recently, interest in PTAs has re-emerged as a major international policy issue. This has followed as a result of the formation of a number of new trade alliances (such as NAFTA, the North American Free Trade Area) and the importance of the full membership of several Central and Eastern European (CEE) countries into the European Union (EU) after 1 May, 2004. Consequently, there has been a rapid growth in the empirical literature during the past decade focusing on the estimation of the effects of the PTAs on the different aspects of the national economies. While some studies are concerned with the extent to which the PTAs affect the welfare of member and non-member countries, others attempt to measure the size of the potential trade between the CEE countries and the EU. Parallel to the re-emergence of regionalism in the world, some other studies concentrate on searching for viable trading blocs. As the barriers to trade on the basis of a PTA disappear, the economic theory suggests that trade among member countries should intensify as integration advances. However, the literature on the border effects illustrates that the existence of the national borders draws a limiting line to the level of integration.

3.1. The welfare effects of the PTAs: Trade creation or trade diversion?

There is still no consensus concerning the welfare effects of customs-unions in the empirical literature. Empirical results continue to give mixed messages concerning questions such as whether economic integration increases total trade and hence the welfare of members or whether integration harms non-member countries. There is no overall agreement even concerning the most accurate way of arriving at estimates of the size of these welfare effects. While some studies report positive and significant results (Linnemann, 1966; Bergstrand, 1985), others conclude that the welfare effects of regional trade blocs are insignificant (Frankel et.al, 1995; Aitken, 1973).

As noted earlier, the empirical analyses of the welfare effects of the PTAs have almost exclusively been limited to the estimation of the trade creation and trade diversion effects of economic integration. To this end, extended forms of the gravity equation which include PTA dummies have been utilised to quantify and analyse both the size and the source of increased trade. In this setting, all variables except the PTA dummies are assumed to explain normal trade that occurs in the absence of PTAs between countries and thus the PTA dummy variables pick up abnormal levels of trade that could be attributed to a bloc membership effect. The first
question that empirical studies have been concerned with in this context is whether regional integration increases international trade. The second and more important question concerns the identification of the source of an increase in trade flows between members of a PTA – i.e., whether increased trade is the result of trade creation or trade diversion effects.

Early studies that attempted to estimate the trade creation and trade diversion effects of integration have incorporated a PTA dummy variable that takes the value one if both countries are members of the same bloc, zero otherwise (Aitken, 1973; Pelzman, 1977, Primo-Braga et al., 1994). In this way, the measurement of the magnitude of the gross trade creation effect becomes possible. To calculate the trade creation and trade diversion effects, the gravity equation is first estimated for the pre-integration period and uses as the basis for arriving at trade projections. The projected values are then subtracted from the actual trade values with the outcome representing the size of the trade diversion effect. Trade creation is then the difference between the scale of the gross trade creation and trade diversion.

Based on the above methodology, Thursby and Thursby (1987) estimate a gravity equation extended by a set of PTA dummies, each dummy representing one of a sample of 17 European Economic Community (EEC) and the European Free Trade Area (EFTA) countries for the period 1974-1982. They conclude that the trade creation effect of PTAs is significant and positive for the EFTA countries but this is not true for the other European countries with the exception of Denmark and the UK.

More recently, researchers have added a second set of dummy variables to the gravity equation in an effort to capture the effects of the PTAs on the trade between members and non-members (Bayoumi and Erichengreen, 1997; Frankel and Wei, 1998). In this way, it is possible to separate out the trade creation and trade diversion effects of the PTAs.

As mentioned above, trade diversion due to the existence of a PTA occurs when a home country switches trade from a non-member to a member country. However, a home country can divert not only imports but also exports from non-member countries. Endoh (1999) has suggested a refinement of the PTA dummies to estimate these import and export diversion effects for the EEC, the Latin American Free Trade Association (LAFTA), and the Council for Mutual Economic Assistance (CMEA). Furthermore, he examines the effects of these regional formations on Japan’s trade. To this end, import and export diversion effects dummies replace with the dummy for the trade diversion effects in estimations. Formally, the following model is estimated using the ordinary least squares
(OLS) regression technique based on cross-section data covering 80 countries recorded at intervals of five years from 1960 to 1994:

\[ x_{ij} = \alpha_0 + \alpha_1 z_{ij} + \alpha_1 EEC^1_{ij} + \alpha_2 EEC^2_{ij} + \alpha_3 EEC^3_{ij} + \ldots \ldots + \alpha_9 CMEA_{ij} + e_{ij} \]

where \( x_{ij} \) represents the exports from country \( i \) to country \( j \), \( z_{ij} \) represents all the standard variables of the gravity equation namely incomes and populations of both exporting and importing countries, distance, adjacency, and language. Three sets of dummy variables are included, representing the EEC, LAFTA, and CMEA areas. For example, \( EEC^1_{ij} \) is the import diversion dummy taking the value one when a country \( i \) is a non-EEC country and country \( j \) is an EEC country and otherwise zero. \( EEC^2_{ij} \) is the trade creation dummy variable taking the value one when both \( i \) and \( j \) are members of the EEC, zero otherwise. \( EEC^3_{ij} \) is the export diversion dummy and is one when an exporting country \( i \) is a member of the PTA and country \( j \) is a non-member country, otherwise zero. The results from the estimated equations indicate that while the coefficients relating to income variables tend to increase, the coefficients on the population variables decrease prior to 1970. Endoh (1999) argues that this finding arises as a result of the expansion of world trade during this period. In addition, it is argued that the reversing trend after 1970 is due to the global recession and the sharp downturn in the world trade following the oil crisis in the early 1970s. As far as the regional dummy variables are concerned, Endoh’s analysis reveals positive and significant \( EEC^1 \) and \( EEC^3 \) coefficients indicating that the EEC members and non-members show a trend towards higher trade levels than expected over the sample period, contrary to the theoretical expectations. But it would appear that this upward trend is decreasing. Endoh also reports an insignificant trade creation effect arising from the formation of the EEC. For the LAFTA,

---

1 Each dummies have good interpretations as well: If \( EEC^1 \) (import diversion) dummy has a negative and significant coefficient, it means that the member of EEC switched their imports from non-member to member countries. In the same way, positive and significant coefficient on \( EEC^2 \) (trade creation) variable shows that the level of trade among members of the EEC higher than the pre-integration period. Negative and significant values of \( EEC^3 \) (export diversion) coefficient means that members prefer to export to members rather than non-member countries.
estimation results show that this arrangement has negative trade creation and diversion effects suggesting stagnant trade in the Latin American region. The coefficients on the CMEA dummies indicate that while this bloc has a positive and significant trade creation effect, it has also a negative trade-diversion effect. On the basis of these various findings, Endoh (1999) concludes that the trade effects of the preferential trade arrangements across all three regions have been weakening, especially during the 1990s. Finally, Endoh points out that neither the existence of the EEC nor the LAFTA have significantly affected Japan’s overall trade. Although the existence of the CMEA has a negative effect on Japan’s trade with the CMEA member countries, this effect seems to be weakening over time.

In a similar cross-section study, Soloaga and Winters (2001) apply a gravity equation approach to nine PTAs from 58 countries over the period of 1980-96. Using a Tobit model, two sets of results are derived. The first is based on a set of 17 separate regressions—one for each year—over the total time period 1980-1996. The second set is based on averaged data for the sub-periods 1980-82, 1986-88, 1995-96 which are then pooled and estimated as a single equation with three time dummies included and allowing for all coefficients to be different over the three periods. A test is then conducted to see whether or not the estimated coefficients obtained for the period 1995-1996 (a period considered by Soloaga and Winters as embracing post integration years) are different from those obtained for the 1980-82 and 1986-88 periods. A real exchange rate variable is also added into this single equation. It is argued in this approach that the coefficients on the import and export trade diversion dummies can be interpreted as openness proxies. In general, the estimation results enable two key findings to be drawn from this study. Firstly, welfare effects tend to differ widely among the different PTAs. While the Latin American trade arrangements have significantly positive trade creation effects and significantly negative import-export trade diversion effects, intra-bloc trade is negative for the EU, EFTA and ASEAN regions. This suggests that a trade creation effect only exists in the Latin American blocs. Second, there is some evidence of the export diversion in the EU and EFTA regions. Considering the general pattern of the results concerning the EU, Soloaga and Winters (2001)

---

2 These are ANDEAN, CACM, LAIA, MERCOSUR, NAFTA, ASEAN, GCC, EFTA, and EU.

3 They have used Tobit model because trade values are bounded from below zero. However, considering that only 6% of observations were zero in their sample, Tobit estimates will be equal to OLS estimates.

4 CACM, LAIA, ANDEAN, and MERCOSUR.
conclude that the closer economic integration that has taken place between the European countries imposes costs on excluded countries.

3.2. The potential for trade between the CEE and the EU

The opening up of Central and Eastern Europe in 1989 and the expressed interest of the CEE countries to join the EU has rapidly led to the development of an area of research that seeks to arrive at an estimation of the future potential growth in trade between East and West Europe. Naturally, the gravity equation is used as a modelling framework and many researchers have employed a basic or/and an augmented gravity equation to calculate the potential trade. The heart of these discussions have mainly concentrated on two questions: (1) whether the EU’s trade potential with the East European countries is high as the early studies projected or it has already exhausted as suggested by most recent analyses; and (2) whether the data and methods used in empirical analysis are relevant enough to reach reliable conclusions.

Wang and Winters (1992), Baldwin (1993,1994), Hamilton and Winters (1992) argue that the estimation of the gravity model by using the CEE countries’ data provides biased coefficients since these countries are in transition and the gravity model assumes equilibrium relationship. Therefore, they use, in their analysis, the coefficients of the gravity model for the middle income developing countries with the assumption that the CEE countries’ trades will behave exactly like the middle-income countries once they have completed their transition. Then using these estimates and the actual trade data from the CEE countries, they have projected the level of trade for the CEE countries. Then the difference between the actual and the projected trade of the CEE with the EU is taken as potential trade.

Wang and Winters (1992) investigate the potential volume and direction of the CEE countries’ exports and imports. They make use of the data from 76 countries (19 industrial and 57 developing) over the period of 1984-86. Since the purpose of this exercise is to characterise the market economies’ trade pattern, the CEE countries and China are excluded from the sample. They use the data averaged over 1984-86 to reduce the effects of temporary disequilibrium and other temporary shocks and estimate the basic gravity model extended with the PTA dummies. They omit all flows recorded as

5 While the authors include Yugoslavia arguing that Yugoslavia is relatively well integrated into the West, they exclude oil exporting countries.
6 Additional independent variables are the PTA dummies for the EEC, the EFTA, the ECOWAS, the SADCC, the CACM, the AG, the LAIA, the ADEAN; dummy for the EC preference to Africa - Caribbean –Pacific countries, dummy for
zero. The estimation results reveal the theoretically predicted signs for all variables except for the EFTA. Although several of the preference dummies were not statistically significant from zero, they retain them in the equation to ensure that their absence does not bias the estimation of the main parameters. Using the coefficients from the extended gravity model, they calculate the potential trade of each CEEs with the EC, the EFTA, the East European countries and other industrial countries. The results indicate that the CEE countries have only achieved 20% of their exports potential and 17% of their import potential with the EC. However, it seems that the actual trade between the CEE countries is 1.69 times higher than the level of trade predicted by the gravity model. The authors conclude that the potential trade between the CEE countries and the EC is huge and refer to trade in both directions.

Contrary to the predictions of the previous studies, Gros and Gonciarz (1996) argue that opening of the CEE countries has occurred very rapidly and they have already reached and even become highly open economies compared to some market economies. They have challenged the results of the previous studies on the grounds that previous studies are carried out their analysis using pre-regime change data of the CEEs. The previous studies predict higher level of trade for the CEE countries for 1992 because they are based on 1989 GDP figures, which are very much higher than 1992 GDP figures. Furthermore, they have argued that when the same analysis is repeated with using actual 1992 data, it will be seen that the CEE potential trade with the EU has already exhausted as illustrated in Kendall (1994). They have also compared the openness figures (imports over GDP and export over GDP) from the CEE and the EU countries and showed that the CEE countries seem more open than similar EU countries. They conclude that CEE countries should not be expected to open more, so the trade potential with CEE has already exhausted.

However, Breuss and Egger (1999) have shown an inherent weakness of these studies in estimating and predicting the potential trade as far as the unilateral preferences from industrial countries to developing countries, dummy for the ex-colonies of the UK and France, and dummy for adjacency.

With the Fall of Berlin Wall in 1989, the countries in the socialist bloc entered into new era and these economies called the transition economies which represent the transition from plan to market. Following the collapse of the old economic structure, these countries initially witnessed a dramatic decrease in output, shrink in capital, movement in the labour market, change in the institutional structure (Campos and Coricelli, 2002; Facchini and Segnana, 2003; Eren and Bildirici, 2001).
econometrics is concerned. They have argued that econometric estimations and well specified models are necessary as it is the case for cross-section gravity equations, but this is not a sufficient condition for a sensible projection of bilateral trade flows. After estimating commonly used three different versions of the gravity model in the literature, they have concluded that the previous studies are not capable of forecasting the potential trade because of the huge confidence intervals around the potential trade estimates. Thus, the argument of whether or not the potential trade is exhausted should be interpreted carefully, as there is no statistical evidence for it.

Other objections against the use of the gravity models for an analysis of potential trade are due to Brabant (2001). These objections based on a number of issues: (1) Data: The gravity equation assumes that the long-run general equilibrium condition is satisfied. Application of the gravity equation would be irrelevant for the transitional economies because they experienced a strong break in their economic structure since 1989. That is, one cannot predict bilateral trade flows in the post-reform period making use of pre-reform period data. (2) Theory: The theory underlying the model is not solid. Thereby, it is difficult to interpret the estimated coefficients. Furthermore, the equation is definitely a reduced form equation and because we do not have a firm theoretical model, we would never know the structural parameters or structural equations. In this sense, the whole idea seems like a black box. (3) Parameter estimates are not stable over time, and (4) Related to the criticisms given above, it is not possible to use the gravity model for forecasting or predicting the level of future trade.

Brulhart and Kelly (1999) estimate the magnitude of potential trade flows between Ireland and the five CEE countries. Using the data from 24 countries (Ireland and the CEE countries are not included in the sample to provide unbiased estimates) for 1994, they estimate the basic gravity model extended with remoteness index language and a dummy for the EU. The remoteness is defined as the average of a country's distance to its trading partners weighted by those partners' GDPs. The results suggest that the trade effects of the EU enlargement on the Irish economy will be relatively modest.

3.3. The importance of national borders

Growing trade and capital mobility through globalisation and changing significance of national borders with an increasing number of regional trading blocs may have created the impression that national boundaries no longer matter much for trade and capital movements. In the same way, the
idea of "borders effectively disappeared" has found a wide-separate acceptance and become a dominant view among many people. These are very exiting developments because once the barriers to trade between the members of a PTA are removed, the theory suggests that trade between economies will increase and hence income. Furthermore, law of one price asserts that commodity prices will eventually equalise if free movements of goods are allowed. Recently, this type of evaluation of the effects of globalisation and regionalism has been challenged by a very new and growing literature on 'national borders matter'. Findings of this literature have far reaching implications: even the countries achieve full economic integration, the level of activity among member countries will not be the same as happens to be within a national economy. This has also important implications of an assessment of formation of trading blocs because the gains from an economic integration seem to be directly related to the strength of border effect.

McCallum (1995) has studied the impact of the Canada-U.S. border on the pattern of Canada's continental trade pattern. The hypothesis he tested is that whether the borders separating these two very similar countries exerts a decisive impact on continental trade patterns. He has made use of the data set that involves imports and exports for each pair of provinces, as well as imports and exports between each of the 10 provinces and the 30 US states for the year 1988\(^8\). The author has estimated the basic gravity model augmented with a border dummy that takes one for interprovincial trade and zero otherwise. The estimated basic model is as follows:

\[
x_{ij} = \beta_0 + 1.21 y_i + 1.06 y_j - 1.42 \text{dist}_{ij} + 3.09 \text{Border}_{ij}
\]

The elasticity of exports with respect to own GDP, importing region income and distance are respectively 1.21, 1.06 and -1.42. The author explains substantially large coefficient on distance variable\(^9\) because of cost differences between air and water transport. Since the trade between Canada and the US use air and land, which are more expensive than using water, it is sensible that the coefficient on distance variable is higher than one in absolute value. The coefficient on the border dummy is 3.09, which implies that the trade between two provinces is more than 20 times larger than the trade between a province and its neighbour province in the US state [exp

\(^8\) Total numbers of observations are 690: 10*9=90 observations for interprovincial trade, plus, 10*30*2=600 observations for province-state trade. In seven cases there was no recorded trade, leaving 683 nonzero observations.

\(^9\) International studies report less than one in absolute value.
This finding has come as a surprise and latter called as the border puzzle because these two countries are very similar in terms of culture, institutions, language, and trade restrictions between them have long been lower than anywhere in the world (Helliwell, 1996).

The author has also estimated 10 alternative regressions (closely related to the one given above) to deal with econometric and specification problems and to test the robustness of the estimates. The estimations for alternative specifications seem stable and provide the own GDP coefficients ranging between 1.15 to 1.36, importing country coefficients from 0.96 to 1.19 and distance coefficients around -1.4. The border dummy coefficients range from 3.07 to 3.30. The author foresees the possible criticism that the findings of the paper is the result of using 1988 data which is the year Canada-US free trade agreement is signed and once the integration take place the border effect disappears. But he dismiss this idea arguing that the NAFTA will not bring about more tariff reductions because tariffs are already low and observations of long-run trend in tariff-trade relationship also predicts an insignificant increase in trade following the NAFTA. He seems he is convinced that “whatever the reason may be and whatever the future may hold, national borders continue to matter”. The later studies on the border effect are carried out to uncover the determinants of it or to solve it.

Helliwell (1996) has extended the McCallum (1995)’s analysis to 1990. He has aimed to answer the following two questions: Is it possible that McCallum’s finding is the result of the year that he has estimated his model; what are the implications of border effects if exists for Quebec separation and its interprovincial and the US trade patterns? To answer these questions, he has estimated three sets of gravity models using data for 1988, 89 and 1990 separately for each set with the Ordinary Least Squares method (OLS). Then, combining all three years, he has estimated a system of three equations model using the seemingly unrelated regression (SUR) with coefficients restricted to be the same (except intercepts) for all years. In his first set of regressions, estimation results are slightly different from McCallum (1995) for 1988 because of the revision of the data. Border effect is 19.9 for 1988 compared to McCallum’s estimate of 22, 18.7 for 1989 and 25 for 1990. When the data for all years are combined and estimated with the SUR, the border dummy coefficient is 21.1. This result indicates that McCallum’s original estimations are stable over time as well.

In the second set of regressions, Helliwell searches for the possible implications of the border effects for Quebec separation and for international trade. If national borders do not matter for Quebec, then
Quebec does not lose anything from separation. To this end, he has extended the McCallum’s model with two more dummies (first dummy takes one for Quebec’s trade with the US states and zero otherwise and the second dummy takes one for Quebec’s trade with provinces) and estimated the gravity model using the whole sample. He has found that Quebec’s trading links with the rest of Canada, relative to those with the US, are at least as strong as they are for other provinces (with border effect of 26.8).

In the third set of equations, he has estimated the first gravity model using only shipments to and from Quebec. Results confirmed that the border effect for Quebec is factor 21.1 implying that national borders matter at least as much as for Quebec as they do for the rest of Canada.

In a similar study, Anderson and Smith (1999a) extended McCallum’s work for all Canadian provinces and the US states and the world. They have searched for an answer to the question of whether the border effect was the result of the specification of early studies. To eliminate a number of other mis-specification, they have tested the following questions: (1) “Is the border effect the same for the US exporters as it is for Canadian exporters?; (2) Is the bias towards interprovincial trade uniform across Canadian provinces?; (3) How large is the border between Canada and its other trading partners?; (4) Is the border effect the same for imports as for exports?”. The authors have made use of 1988 data for the conformity with McCallum (1995) and Helliwell (1996) (hereafter M&H).

First, there is no reason to believe that the size of the border is the same for Canadian firms and the US firms. So, authors have modified the McCallum’s model by replacing the border dummy with D2 dummy (takes one when the US exports to Canada) and D3 dummy (takes one when exports are from Canada to the US). The estimation results have provided that the coefficients of D2 and D3 dummies are 0.04 [=exp (-3.21)] and 0.051 [=exp (-2.97)] and they are highly significant. These results have suggested that borders are different for the US and Canadian firms and that the US firms do not care as much about benefits from Canadian markets as Canadian firms may be interested in the large US market (because the D2 dummy is bigger in absolute terms than D3 dummy). That is, the borders differ depending on which direction trade flows.
Second, they have tested the hypothesis that M&H results might be sensitive to the exact specification of the gravity model employed and therefore they have used the following model in the estimation:

$$\ln(X_{ij} + X_{ji}) = \beta_0 + \beta_1 \ln(Y_i Y_j) + \beta_2 \ln([Y_i / P_i][Y_j / P_j]) + \beta_3 \ln(Dist_{ij}) + \beta_4 D_{1ij} + \beta_5 Contig_{ij} + u_{ij}$$

The results suggest that a border effect of 15.2 \([=\exp (2.72)]\). Although it is 31 per cent smaller than that of McCallum, it is still a substantial effect suggesting that the border effect is not the result of a particular specification.

Third, to answer the question is that whether the border effect is uniform across the Canadian provinces, they have estimated the basic gravity model for each province using a new data set which includes exports and imports from a province to all other provinces. Furthermore, they have introduced import and export border dummies for the interprovincial trade and for province-US trade to see whether the provinces have different borders for imports and exports. The findings have indicated wide divergences in the coefficients of provincial border effects raging from 2.3 to 3.9.

Fourth, they have tested whether the border between Canada and US is the same as the border with the rest of the world. To this end, they have estimated the basic gravity model with the US and non-US border dummy variables. They have found that the coefficients on both dummies are very similar (3.14 for the US and 3.03 for non-US). They have interpreted this small difference among these two dummies as a result of a specialised pattern of trade. That is, border with the US is higher because the Canada-US trade is subject to similar products and Canada-non-US trade takes place with goods that is not easy to find in somewhere else. Therefore, complementary goods trade led to lower border effect and substitute goods trade cause higher border effects.

They have concluded that first, there exists a strong trade-reducing border between Canada and the US. Second, there is no single border but there are numerous borders. Third, despite the belief that the US is relatively well integrated with Canada, results have shown that US-Canada border

---

10 Because this formulation of the gravity model is due to Frankel, we will call it Frankel type gravity model from now on.
remains as a substantial obstacle between Canada and the US as well as Canada and the rest of the world. Simply, puzzle remains as a puzzle.

In a similar study, Anderson and Smith (1999b) have tested whether protection is the source of large border effects by making use of a new data set for 1990 Canadian provinces – the US trade. Since Canada- the US trade in the transportation equipment has been completely liberalised for several decades, they have estimated the border effects in the transportation equipment trade to test whether the protection is the cause of the large border effect. Furthermore, they have tested whether the border effect is uniform over exports and imports.

They have estimated the basic gravity model and then the extended gravity model with export and import border dummies as well as its Frankel type gravity model. The estimated border effect from the basic model is 12.5 compared with Helliwell (1995)’s estimate of 24.8 for 1990. The border effect for transport equipment is estimated to be 45.7 which is substantially higher than the overall border. Further analysis of export and import borders shows that the border effect is more pronounced on the import side than the export side. Frankel specification has provided similar results. Authors have concluded that a large border effect cannot be explained by the errors of measurement in the data and/or formal or non-tariff barriers. Simply, the empirical findings of this paper rule out official trade protection as the source of the border effect.

Ceglowski (2000) investigates whether or not border effects change over time and the impact of the US-Canada border on international trade has fallen since the Free Trade Area (FTA) went into effect. Actually, there are a number of reasons to believe that the border between the US and Canada is narrowing: (1) Canada- the US FTA established in 1988 may have narrowed the border; (2) Effective Tariff Rates (ETR) for Canadian imports fell from 2.5% to 0.7% in 1995; (3) ETR for Canadian exports dropped from 0.9% in 1989 to 0.3% in 1995. Using yearly and the pooled data for 10 Canadian provinces and 30 US states over the five years (1988, 90, 92, 94, 96), first, the author has estimated the basic gravity model extended with the border dummy for each year with the OLS. The estimation results have indicated no evidence of sustained decline in the border effect relative to its pre-FTA level in 1989. The border effect for the McCallum’s specification is 22.2 for 1988, 23.3 for 1990, 23.3 for 1992, 21.8 for 1994 and 20.9 for 1996.

Wei (1996) has extended the M&H analysis to a broader set of countries over 1982-92 and discussed the welfare implications of the border effect. He
has argued that once a microfoundation of the gravity model is considered, the basic gravity model requires to be modified as follows:

\[ x_{ij} = \beta_0 + \beta_1 \text{Border}_{ij} + \beta_2 y_i + \beta_3 y_j + \beta_4 \text{dist}_{ij} + \beta_5 R_i + \beta_6 R_j + \beta_7 \text{Lang}_{ij} + \beta_8 \text{Adj}_{ij} + u_{ij} \]

where \( \text{Lang}_{ij} \) and \( \text{Adj}_{ij} \) are dummies for country pairs that share a common language or a common land border, respectively. \( \text{Border}_{ij} \) is a dummy that takes the value of one if \( i = j \) and zero otherwise. \( R_i \) and \( R_j \) are remoteness indices that show the geographic position of a country relative to all other countries. More importantly, the author has shown that the coefficient of border dummy is a product of two things: degree of substitutability of goods produced by different countries and log of the tariff equivalent.

The main difficulty in extending the border effect analysis for the OECD countries is the lack of direct data on a country's trade with itself and on intra-national distance. He ingeniously overcomes this problem by defining a country's imports from itself as just the difference between its total production and its total exports to foreign countries\(^\text{11}\).

The author utilise the data set of 19 OECD countries for four years (1982, 86, 90, 94) and estimate the above extended gravity model employing the method of the SUR which allows for correlation across years. The findings of this study can be summarised as follows: First, the estimation of the M&H model for OECD countries has provided an estimate of 9.7 \([=\exp (2.27)]\) for border effect. Once the M&H model is extended with a measure of remoteness, language and adjacency dummies, the estimated home bias has fallen sharply to a factor of 2.3 \([=\exp (0.84)]\). However, the difficulties surrounding the measurement of intra-national distance cast a serious doubt about the extent of the border coefficient. The author himself admits and simulates that if intra-national distance were 25% larger than be used in estimation, the resulting home bias coefficient is about 25% larger. Third, to test whether the degree of home bias is different for trade bloc members, he replace the OECD border dummy by five dummies that takes one when trade takes place between: EC-EC, EC from itself, EFTA-EFTA, EFTA from itself, non-EC and non-EFTA. Estimation

\(^{11}\) To get a measure of a country's imports from itself, first, he subtracted service and transport sectors and called this GGDP. Second, he multiplied GGDP by production to value added ratio to convert value added GGDP figures into gross production figures because trade takes place in products.
results show that intra-EC dummy has a negative coefficient of (-0.25) implying that trade among the EC countries were actually lower than a random pair of the OECD countries. He also estimates the same equation for each year and find out that intra-EC coefficient is rising over time. The coefficient on the EC border dummy has turned out to be 1.7 and lower than the OECD average of 2.5. For the EFTA on the other hand, the average intra-EFTA trade bias is positive (0.27) but not significantly different from zero. The EFTA member countries’ relative degree of home bias is substantial (4.3) showing a member imported from itself 4.3 times as much as from other member countries. For the remaining non-EC/EFTA countries, the degree of home bias is very close to the average of all countries. Third, he tests the hypothesis that home bias stems from the exchange rate uncertainty. He defines exchange rate volatility as the standard deviation of the first difference in the log of the monthly exchange rate in the current and past years (24 months). However, volatility measures provide no evidence of trade depressing effect of exchange rate volatility. Fifth, the reason why one care about border effect because we think that it may have welfare cost. Nevertheless, the coefficient of the border effect is a product of degree of substitutability of goods produced by different countries and log of the tariff equivalent. Therefore, the observed border effect for the OECD countries (0.91) will be equal to a tariff rate of 4.7% \[=\exp (0.91/20)-1\] if elasticity of substitution is 20. Therefore, the author conclude that if goods are perfect substitutes, then any minor barrier may generate infinite amount of the border effect even though this trade pattern does not lead to any welfare loss. Simply, small tariffs and non-tariff barriers can explain high border effects found in empirical literature once it is recognised that the estimated coefficient of the border effect is composed of substitutability and tariff equivalent as the theory suggests. Sixth, the author has questioned the argument that the world is increasingly integrated over time. To this end, he takes a first difference of the gravity model which eliminates all country fixed effects and estimates this equation for 1982-86, 86-90, and 90-94. Although the coefficient of home bias has a negative sign and showing a decline over time, they are insignificant for the sample except 1986-90. The author has repeated the same exercise to examine the evolution of the border effects for trade blocs, the EU and the EFTA, extending the previous differenced gravity model with five border dummies. Estimation results illustrate that the border effect for an average EC member declines dramatically at the rate of about 5% per year over 1982-1994. In contrast with the findings for the EC, border effect for the EFTA increase in every year and intra-EFTA trade intensity does not seem to increase at all.
For other OECD countries, it is observed that the border effect declines over time.

Anderson and Wincoop (2001) has carried out a similar exercise for the US-Canada trade and find that the borders between the US and Canada reduce trade by about 44%. And they argued that M&H estimates are the result of a combination of omitted variables bias and the small size of the Canadian economy.

Contrary to Wei (1996), Nitsch (2000) provides a high border effect for a highly integrated region of the world, the EU. The author estimates the basic gravity model extended with remoteness indices, language and adjacency and the border effect dummies making use of the data set for ten EU countries over 1979-1990 and for 11 EU countries (Spain and Portugal are added) for the period 1983-1990. Different from the previous studies, the author suggests a new measure of within-country distance variable as a function of area size of a country. The SUR estimation method has been employed whereby allowing year specific intercepts and restricting the coefficients on the variables to be the same for all years. Estimation results have provided the home bias of factor 16 in the EU compared with factor 20 of H&M result for Canada. Then he estimates the first difference of the gravity model for different periods. The estimation results show that the coefficients on the home bias are statistically significant and negative implying a declining home bias in the sample. Furthermore, he tests the hypothesis that richer countries have a smaller home bias than poorer countries. The results reveal that the richer countries in the EU have indeed a smaller home bias. The author concludes that national borders still matter, even within the EU.

Wall (2000) has stressed the fact that the existence of heterogeneity among sample units provides biased estimates of the border effect unless it has been dealt with properly. Wall (2000) estimates the gravity model employing the fixed-effect panel data model making use of the post-NAFTA data for 1994-96. The home bias ratios from the fixed effect model are, on average, 43 per cent higher than those from the standard model (factor 21.6 against 15.1). Surprisingly, the fixed-effect model has provided the border effect, which is greater for trade from Canada to the US than on trade from the US to Canada. Specifically, the home bias ratio on Canadian exports to the US is about 55 per cent larger than the home bias ratio on Canadian imports from the US. This is in contrast with the 40 per cent difference in the opposite direction that the standard model yields.
Chen (2001) examines the border effect among EU countries using the data for 14 European countries and for 78 industries for 1996 data. In estimation, the author employs the Tobit model including industry and country fixed effects. The estimation results have indicated that the border effects differ across countries, but especially across manufacturing industries.

Evans (2001a) has argued that the number of goods being traded internationally is very limited compared to the set of goods that are available in the domestic market. She has estimated the basic gravity model extended with remoteness indices and border dummy employing Instrumental Variable (IV) method. The only difference between this formulation and the previous ones is that own income variable (own country GDP) in the model is modified by multiplying it with the proportion of goods that are subject to trade. Estimation results show that while the border effect is factor 14.95 for the standard model, it is only factor 8.27 for the amended model. She concluded that a portion of the border effect is indeed due to difference between the set of goods available domestically and internationally.

Evans (2001b) asks whether home bias arise from pure locational factors such as barriers to imports or access to a local distribution networks or an inherent preference for domestic goods. Using a unique data set of the activities of the US majority owned non-bank foreign affiliates in 9 OECD countries between 1985 and 1994 for seven industries. She utilised the data on total sales and on local sales by these foreign affiliates as measures of production and sales to domestic consumer respectively. She formulated the gravity model as a function of home and partner country income, distance between them, remoteness indices and location effect dummy that takes one for foreign affiliates’ sales and zero otherwise. Estimation results have shown that the apparent tendency to purchase domestic goods rather than imports arises almost entirely from pure locational factors. First, the estimated home bias ranges from factor 20.39 to factor 294.65 for different industries. Second, if a firm establishes and sells from a subsidiary located in the foreign country, its local sales are nearly on a par with those domestic firms in that market. Foreign-ness itself does not appear to impede purchases of imported goods.

3.4. The search for natural trade blocs

It is believed that the certain groups of countries are eligible to form a PTA in statistical sense although there is no formal agreement among them.
Frankel, Stein and Wei (1995) demonstrated that the APEC\textsuperscript{12} region countries already form a regional trading bloc in the statistical sense. They have added a regional dummy variable for a number of potential regions to the gravity model and found out that the APEC dummy has a significant coefficient suggesting that the trade enhancing power of the APEC be considerably high. The size of the coefficient on the APEC dummy reveals that the trade enhancing power of the APEC is even higher than the EU. Hassan (2001) carries out a similar study for the SAARC\textsuperscript{13} and estimation results indicate that there is a vast amount of trade creation effect in this economic bloc. Therefore, he concludes that efforts need to be made to liberalise the border trade among the member countries.

Polak (1996) simply attributes these results to an inherent mis-specification of the standard gravity equation. He regresses the location index due to Linnemann (1966) on actual over estimated imports variable and find a highly significant and negative coefficient. He interprets this result, as a clear indication of the bias inherent in gravity equation and the bias is downward for far-away countries and upward for close-in countries. To overcome this problem, he suggests to use location index as another variable in estimation or to use relative distances instead of absolute ones. And he warns that if the gravity model is applied without these corrections, then the upward and downward biases goes to error term and the PTA dummies catches all these effects and create phantom PTAs like the APEC. Furthermore, it creates anti-phantom regions like insignificant trade creation effect in EU they found.

3.5. Econometric issues related to the gravity equation: Different estimation methods

In most of the cases, it has been argued that the gravity equation is very successful in explaining bilateral trade flows on the basis of goodness of fit (high R-sqr) without paying any attention to empirical properties of the model. It is well known fact that strength of a model lies in the accuracy of its estimates. However, there are a number of studies carried out to demonstrate the bias that previous studies were subject to (Cheng and Wall, 2001; Matyas, 1997; Bayoumi and Eichengreen, 1997).

\textsuperscript{12} APEC (Asian Pacific Economic Cooperation) includes 14 countries: the USA, Canada, Australia, New Zealand, Japan, Indonesia, Taiwan, Hong Kong, South Korea, Malaysia, the Philippines, Singapore, Thailand and China.

\textsuperscript{13} SAARC stands for the South Asian Association for Regional Cooperation
In its very general form, the gravity equation shows the volume of trade flow between countries $i$ and $j$ in year $t$ and can be written as:

$$x_{ijt} = \alpha_0 + \alpha_t + \alpha_{ij} + \beta_{ij} z_{ijt} + \varepsilon_{ijt}$$

$x_{ijt}$ is exports from country $i$ to country $j$ in year $t$, and $z_{ijt}$ is a row vector of all other variables. The intercept has three parts: $\alpha_0$ is common to all years and countries; $\alpha_t$ is specific to a year $t$ and common to all countries; $\alpha_{ij}$ is specific to the country pairs and common to all years. The various forms of the gravity model that have been estimated in the literature are the restricted version of this general model and they can be classified as follows (Cheng and Wall, 2001): (1) cross-section model: Using the cross-section of country data at one year, this model imposes the restriction that the slopes and intercepts are the same for all country pairs; (2) pooled data model: uses the pooled data over cross-section and time series and imposes zero restriction on country pair constants. Obviously, estimates from cross-section and pooled data (because of restrictions they put on estimated coefficients as mentioned above) will be subject to bias because of heterogeneity among country pairs.; (3) fixed effect model (panel data): this model uses the data over cross-section and time series and removes the restrictions over time and country specific constants. Since this model allows the heterogeneity among countries by allowing the country specific effects to differ in both directions, it provides consistent results as far as statistics is concerned. Briefly, the main econometric problem in using gravity model is to overcome the heterogeneity problem.

Cheng and Wall (2001) argued that because standard methods (cross-section-OLS estimations) failed to account the heterogeneity among bilateral trade relationships, they produced biased estimates in estimation of the gravity equation by tending to overestimate trade between low-trade countries, and to underestimate it between high-trade countries. To overcome these problems, they have used panel data methods to allow for the intercepts of the equation to be specific to each country pair. Furthermore, they found out that standard methods tend to overestimate the effects of membership in trade blocks on intra-block trade volume.

14 See Cheng and Wall (2001) for the restrictions on this model in application and details of other formulations.
4. Conclusion

In this study, we attempted to survey the empirical literature on the welfare affects of economic integration. The interest in the subject definitely has an increasing trend, thereby the empirical literature. It is also true that the findings of the studies are not easy to compare because each uses different sample, sample periods, and methodologies. This study attempted to present the general findings of the literature in a critical manner and showed the deficiencies in the present literature.

The originality of this study may be its objectivity in presenting results and recommends possible research topics in this area.

1- In the literature, it has been argued that even countries accomplish to form a PTA, the borders act like an obstacle to free trade among them. In this respect, the EU can be envisaged as one country that was supposed to be completed its integration until now. In other words, if the EU completed economic integration process, the trade among members of the EU should be as high as the trade among the US states or Canadian provinces. Following this line of reasoning, we can achieve two things: (i) we can determine the level of integration reached in the EU and find main obstacles to it; (ii) we can compare dynamics of two PTAs and find out which achieved integration and why.

2- Considering the literature on foreign direct investment (FDI), FDI lead to increased exports or imports depending on the source country. This argument is important for a number of reasons: (i) The gravity model needs to be augmented with FDI variable because FDI is very closely related to the level of trade among countries; (ii) It is unlikely that the statement is true within the EU. It could be possible that low trade creation effect of the EU as found in some studies was the result of high FDI among member countries and FDI replaced with the trade. The reverse is also true that FDI increased trade. That is, we find high trade-creation effect not because of increasing efficiency following integration, but only the result of trade creation effect of FDI.

3- There are vast amount of evidence that, to benefit from export promotion policies, a minimum level of development is necessary. For us, this means that there is a threshold effect in trade, poor countries trade less, and rich countries trade more. Furthermore, we know that PTA dummy variable represents the effect of integration and, at the same time, this dummy is included only for the rich countries. Therefore, it is highly probable that the integration dummy used in the empirical analysis presents
the threshold effect rather than the PTA effect. Is there a threshold level
that guarantees the success of an integration?

4- It is important to know whether gains from enlargement of the EU
will be evenly distributed. Even within the Europe, changes in the centre of
the gravity would have different affects on different countries when new
comers change the centre. At least, the distribution of the shares may not be
equal among partner countries. Therefore, relative gains from the Union will
be different for different former member countries. Depending on the type
of industrial structure new comers had, former members, which have similar
economic structures would be the one that will be affected more.

If the distance was only variable that explains gains from integration,
Britain will be an unlucky country once East-Europeans joined. Then two
more questions arise: How close are the new comer economies to their
neighbours? That is, if new comer and its neighbour was an agricultural
country, what will be the prospect of trade among them? Further, if the
distance is everything, then integration will bring nothing except the gains
from removal of tariffs. Second one, since these new countries are all
neighbours, it is expected (gravity theory) that much of the increased trade
takes place among new comers.

5- Polak (1996) argue that Frankel et. al. (1994)'s natural trading bloc is
phantom because the gravity model is subject to downward bias for far-
away countries and upward bias for close in countries and he suggest to use
distance index to correct this. Cheng and Wall (2001) also argue that the
gravity model is subject to downward bias for the rich and upward bias for
the poor countries because of heterogeneity and they suggest panel data
estimation. If it can be shown that panel data solves the problem that Polak
mention, then we can start searching for natural trading blocs. Otherwise,
we construct the index suggested by Polak and see whether there is really a
natural trading bloc.

6- In a number of papers, trade creation effect of the EU seems to be
very poor, even insignificant. Polak (1996) argues that this is due to the bias
that the gravity equation suffers and this bias stems from an absolute
distance variable used in the equation. The same thing that was suggested in
point five seems relevant for this one as well.

7- Until now, the empirical literature on the economic integration
focused on estimation of static effects of integration. This is very closely
related to the availability of the data and it is easy to estimate cross-section
equations. However, as we have mentioned in our survey, the dynamic
effects of economic integration is expected to be more important than its
static effects. Therefore, there is urgency for time series analysis of economic integration.

References:


