An Econometric Analysis of Trade Creation and Trade Diversion in Mercosur and Paraguay

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Victor Federico Gauto Gunther

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Terry Roe, Advisor

August 2012

Acknowledgements

There are many people without whom this achievement would not have been possible. First, I would like to thank the person who motivated, encouraged, and gave me unconditional support and confidence to enter the Ph.D. program, the late Professor G. Edward Schuh. I was never sure this day would come and I know he is looking down on us with a smile. I am also grateful to my academic advisor, Professor Terry Roe, for taking the baton, and wisely guiding me through graduate school. I am thankful he encouraged and supported my interests as well as immensely contributing to the shape of my dissertation. I am also grateful for the professional opportunities he presented to me.

I am grateful to my committee members. I would like to thank Professor Pamela Smith whose guidance greatly enriched the content of my dissertation. Furthermore, I found her support during my job search to be invaluable. I am extremely grateful to Professor Quiqiong Huang for carefully reading and quickly discussing all my drafts with me. Her suggestions made a great difference in my work and her encouragement was motivating. Finally, I would also like to thank Professor Robert Kudrle for kindly reading my drafts and pushing me to revise and improve the content of my work. I am grateful for the touching generosity my advisor and mentors have extended to me.

I am also grateful to the faculty and staff who gave me numerous job opportunities and made it financially possible for me to survive graduate school. I greatly enjoyed working for Professor Steven Taff, with whom I learned valuable skills and realized achievements I never imagined. I am also grateful to Professor Kent Olson who gave me the opportunity to do empirical work, something I have always wanted to become proficient in, even before entering the program. Professor Olson has supported me with research opportunities to the very end of my graduate student days. I would also like to thank Professors Edward Goetz and Thomas Fisher of the University Metropolitan Consortium. My work for them one summer led to my first publication. I would also like to thank Professors Jeffrey Apland and Robert King for giving me the opportunity to be their teaching assistants during my final year in the program. These were invaluable learning experiences which undoubtedly played a great role in obtaining my first teaching job. I am also thankful to Mr. Gary Cooper for his support during the TA selection process.

I am truly indebted to the Fulbright Association, the U.S. State Department, and the Humphrey School of Public Affairs. It is thanks to their support I was able to arrive to the U.S. for graduate school and slowly find my way into the Ph.D. program. I am also grateful to Ms. Gabriele Schmiegle from the International Student & Scholar services for her help, patience, and unwavering support during all my time at the University of Minnesota.

I would like to thank my generous friends in the Twin Cities, Paraguayans & spouses, who became my immediate support group in the U.S. and often gave me something to look forward to on weekends. I am also grateful to my Peruvian friends for the good times.

I am grateful to my parents, Armando and Elisa, for their love and encouragement during the trying times I endured in Minneapolis. My sisters, Veronica and Nadia, also touched me with words of support, and celebrated with me from a distance, the small victories along the way.

Finally, I would like to thank my loving wife Natalia, whom I married while in the program. She has done everything possible to make our lives as graduate students at universities in different continents possible. I will always cherish her heart-warming support.

I will always be grateful to the University of Minnesota for this life-changing experience and I will always be proud of being a Gopher.

Victor Gauto
St. Paul, Minnesota
August 2, 2012

Dedication

To my parents, Armando & Elisa

Abstract

The 1990s marked the beginning of a period of proliferating regional trade agreements (RTAs) around the world. This dissertation analyzes the effects a RTA in South America, Mercosur, has had on Paraguayan import flows. The analyses are carried out at both aggregate and disaggregate levels of detail in order to identify patterns of trade creation and trade diversion on the Paraguayan economy as a whole, as well as more detailed commodity levels. Mercosur (Common Market of the Southern Cone) was signed in 1991 between Argentina, Brazil, Paraguay, and Uruguay. It is well know that the share of foreign trade with respect to GDP is larger for small countries. Consequently, the effects of a trade agreement between large and small countries are likely to be larger in small economies.

In the first chapter, I use a variant of the gravity model employing a re-parameterization of the difference-in-difference estimator to analyze Paraguayan import values of merchandise from 155 countries over a 41 year period (1970-2010), at the single digit level of detail (SITC Rev. 1). Additionally, I explicitly include zero trade flows and implement a Heckman sample selection bias correction with country fixed effects. I find the creation of Mercosur has increased average regional imports by 266% since 1995 without decreasing imports from the rest of the world, which is evidence of trade creation. The greatest import expansions have been in *Beverages & Tobacco* and *Animal and Vegetable Oils & Fats*.

In the second chapter, the sensitivity of the estimates are tested by comparing the results of the traditional OLS estimator and the Heckman sample selection approach to a recently recommended technique in the literature, the Poisson Pseudo-Maximum-Likelihood (PPML) estimator. The specification of the three estimators is implemented using country fixed effects panel analysis of the same data and period. I include additional refinements such as real exchange rate data and controls for major macroeconomic and political shocks. I also test the sensitivity of the estimates to varying sample sizes. Finally, the dissertation analyzes the dynamic time path effects of Mercosur on Paraguayan import flows in an interval analysis. Qualitatively, the results of great trade creation and no trade diversion are consistent across the estimators and the sample sizes. Quantitatively, the Poisson measures of trade creation are well below those of the other two techniques and even reveal some evidence of small trade diversion when the specification is estimated on Paraguay's 19 major trading partners only.

Table of Contents

List of Tables

List of Figures

1 An Econometric Analysis of Trade Creation and Trade Diversion in Mercosur: the Case of Paraguay

1.1 Introduction

Regional and bilateral trade agreements have proliferated around the world over the last thirty years. The World Trade Organization[1] recognizes 217 trade agreements currently in force, out of which 195 have been consummated since 1990. Mercosur and NAFTA are among the older of the more recent trade agreements having been created in 1991 and 1994, respectively. A well known phenomenon in international trade is that smaller, relatively narrow-based economies, the extreme of which are island economies and city states, tend to trade a larger share of their GDP than do larger more diversified economies (Frankel, 1997). Trade agreements between a small economy and larger economies are thus likely to have a larger effect on the share of the small economies' foreign trade in GDP. The net effect on small country welfare may be positive even if the agreement results in some trade diversion provided the gains from trade expansion that would not otherwise have occurred dominate the loss from diversion. Moreover, these gains or losses are likely to be far larger for a small economy. While there have been a plethora of studies investigating the effects of trade agreements, few have focused on small economies.

This paper presents an analysis of the effects Mercosur has had on Paraguayan import flows. Created in 1991, Mercosur is a regional trade agreement (RTA), functioning as a customs union, in South America made up of Argentina, Brazil, Paraguay, and Uruguay. A recent study on Mercosur's common external tariff (CET) reported member tariff rates have changed as each country has gradually converged to the common external tariff. According to this study, Argentina, Brazil, and Uruguay have converged down to the CET and reduced average tariff rates by 50%, 33%, and 25%, respectively. On the contrary, Paraguay has converged up to the CET and increased average tariff rates by 50% (Berlinksi, 2005). This finding suggests Mercosur has raised Paraguay's barriers to international trade. Yet, hundreds of exemptions to the CET remain and negotiations on such product categories as 'capital

[1] http://rtais.wto.org/UI/PublicAllRTAList.aspx

goods' and 'information and communications technology' have continued[2]. Viner (1950) synthesized the benefits and losses of custom unions in the concepts of trade creation and trade diversion, respectively. Trade creation occurs when it is more efficient to import goods from an RTA member country than to produce it domestically, while trade diversion takes place when imports from efficient sources in the rest of the world are shifted to inefficient sources, RTA member countries, benefitting from tariff preferences. Consequently, the overall impact of Mercosur on Paraguay's import flows is the empirical question I attempt to answer.

The creation of Mercosur can be considered a natural policy experiment where the underlying variation in member and non-member tariff rates can be used to identify its impact on import flows from member and non-member countries. Although I do not use the variation in the actual tariff rate, this impact is approximated using dummy variables for member and non-member countries at the time Mercosur's key policies went into effect. I use a rich data set, spanning over forty years, made up of trade values at the commodity level. One of the main objectives of this paper will be to identify patterns of trade creation and trade diversion at aggregate and disaggregate commodity levels. Trade creation occurs when it is less costly to import a good from an RTA member in a trade agreement than to produce it domestically, while trade diversion takes place when a least cost supplier from the rest of the world is replaced by RTA member countries benefitting from tariff-free trade.

Unlike similar studies relying on the gravity model of international trade, I explicitly include 'zero trade flow' observations in the data to account for sample selection bias. My contribution to this literature is the employment of a re-parameterization of the difference-in-difference estimator which I use in combination with a Heckman sample selection correction similar to that proposed in Helpman et al. (2008). Additionally, I also control for unobserved time invariant country characteristics through fixed effects. I find the creation of Mercosur has increased average regional imports by 266% since 1995. Also, I do not find statistically significant evidence of trade diversion in any of the ten commodity categories in

---

[2] The Paraguayan Ministry of Finance reports extended exemptions for 1,200 capital goods and 390 information and communications technology items (http://www.hacienda.gov.py/web-hacienda/index.php?c=96&n=3335). Some examples of exempted goods are 'mobile cellular phones' (HS-2006 85171100) and 'personal computers' (HS-2006 84715010) which have a CET of 20% and 16% respectively. On the other hand, electrical appliances are an example of durables without exemptions, which have a CET of 20%.

the data set. As a result, the expansion of imports attributable to Mercosur reflects gains which could have hardly been realized otherwise. Finally, I find heterogeneity bias from time-invariant country characteristics to be slightly greater than sample selection bias.

Section 1.2 of this paper discusses the context of the study and related literature. Section 1.3 discusses the underlying conceptual framework motivating the empirical strategy. In section 1.4 the data and methodology are reviewed and results are presented in section 1.5. The paper ends with a brief conclusion.

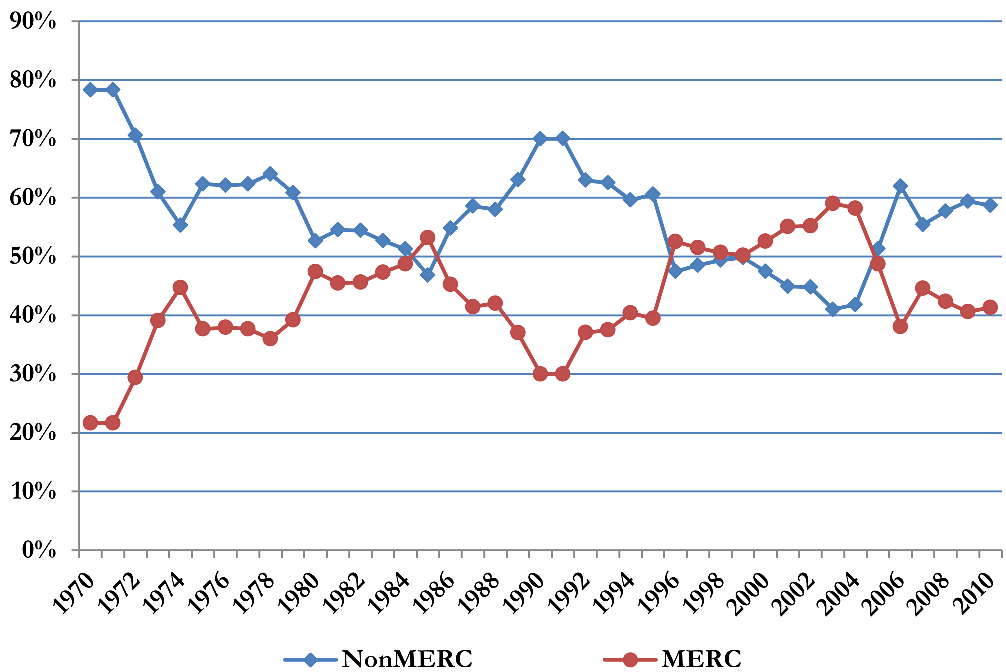
1.2 Context of a study of Mercosur in the Literature

1.2.1 Context

Examining the Mercosur share of Paraguayan imports is illustrative of the history of Paraguay's trading relationship with its Mercosur partners. Figure 1 exhibits import shares from Mercosur and non-member countries from 1970-2010. There are several distinct periods in this figure which may intuitively be explained by major economic shocks to the region. For example, the increasing trend in regional import shares between 1972 and 1986 is consistent with the deeper economic ties Paraguay and Brazil developed during the construction of the Itaipú hydroelectric dam. Similarly, the declining trend in the regional trade share after 1985 is attributable to the inflationary crises that afflicted both Brazil and Argentina in the late 1980s. Argentina suffered inflation rates between 672% and 3,079% from 1985 to 1989. Brazil suffered an inflation rate of 2,900% in 1990 (World Bank, 2011). Paraguay and Uruguay also had peak inflation rates in 1990 reaching 37% and 112% respectively. Given these conditions, it is not surprising Mercosur import shares dropped from 1986 to 1991.

The literature on regional trade agreements refers to anticipatory effect of regional trade agreements (Frankel, 1997; Magee, 2008). Anticipatory effects consist in a sudden increase in regional trade prior to the implementation of a regional trade agreement as firms presumably prepare to take full advantage of tariff preferences. Yet, considering Mercosur began in 1991, notice how these anticipatory effects are not evident in Figure 1. However, these trade shares are unadjusted values and the macroeconomics shocks just described are confounded with the policy intervention of Mercosur. Econometric methods based on regression analysis are useful to cleanse the data of unwanted influences.

Figure 1 Import shares from Mercosur & Non-Mercosur countries, 1970-2010.



Source: UN Comtrade database.

1.2.2 Literature

In a 'clean' natural experiment, one could analyze an outcome of interest before and after a policy intervention and attribute any changes in the outcome to the policy intervention. However, given the existence of any number of confounding factors, such a natural experiment is quite a departure from traditional controlled scientific methods. Nevertheless, by making use of certain refinements, unadjusted trade shares may be cleansed of unwanted influences. In studies where a regional trade agreement is the policy intervention, confounding factors include expansions in regional trade due to anticipation by regional firms as well as macroeconomic shocks. Nevertheless, numerous studies have attempted to evaluate the impact of such policy interventions using the gravity model of international trade. These studies fall into two general categories, those simultaneously analyzing multiple trade agreements, and those analyzing the impact of single trade agreement, the former generally outnumbering the latter.

Often referred to as the 'workhorse' for empirical analyses of international trade, the gravity model's most basic form predicts the volume of trade between countries is a function the partner's GDPs, populations or GPD per capita, and distance. Introduced by Tinbergen (1962), it gained popularity after Anderson (1979), Bergstrand (1985), Helpman (1987), and Deardorff (1997) developed theoretical foundations for its implementation. More recently, the theoretical foundation developed by Anderson & van Wincoop (2003) removes GDP per capita from its specification and incorporates a multilateral resistance term reflecting country price indexes and trade barriers. Anderson & van Wincoop's (2003) was a cross-sectional analysis and Baldwin & Taglioni (2006) extended this theory to a dynamic environment. In the same spirit as the gravity model's abundance of theoretical foundations, the literature has been stirred with discussions about the 'correct' method of estimating gravity models (Matyas, 1997 & 1998; Egger, 2000; Anderson & van Wincoop, 2003; Baldwin & Taglioni, 2006; Linders & de Groot, 2006; and Martin & Pham, 2008), most of which recommend the inclusion of country fixed effects accounting for unobserved time-invariant factors affecting levels of trade between countries.

Traditionally, studies have estimated the gravity equation on countries having positive trade (Frankel, 1997; Clausing, 2001; Feenstra, 2002; Fukao, Okuba, & Stern, 2003; Anderson & van Wincoop, 2003; and Baldwin & Taglioni, 2006). Consequently, any occurrence of 'zero trade flows' are discarded from the sample. One of the main criticisms of this approach is that the absence of trade between countries holds important information, which if ignored, produces biased estimates. One alternative to this method of ignoring 'missing' observations is to consider these observations as evidence of an absence of trade and impute zero-valued trade observations to the data set. Still, this creates additional problems since the gravity model is usually estimated in log-linear form and the natural log of zero is undefined. Out of the several approaches suggested to deal with this problem, Silva & Tenreyro (2006) and Helpman et al. (2008) have been among the most influential. Silva & Tenreyro (2006) estimate the model using a Poisson Pseudo Maximum-Likelihood estimator (PPML), which in addition to producing unbiased estimates, is robust to heteroskedasticity in the error terms. Helpman et al. (2008) develop a theory predicting positive as well as zero trade flows between countries and propose a two-stage estimation procedure, accounting for firm heterogeneity, similar to a Heckman sample selection model.

Studies simultaneously analyzing multiple trade agreements can be further distinguished by whether they incorporate the time dimension to the analysis. Frankel (1997), which has been the basis of this research, reported regression results of yearly cross-sections at five year intervals. He also pooled several years of data and treated RTAs as fixed over time. This treatment of RTAs does not necessarily capture trade created since the policy intervention but it does give an indication of whether regional trade was disproportionately high prior to the implementation of an RTA. Similarly, Soloaga and Winters (2001) also evaluated multiple regional trade agreements in a re-parameterized gravity model. They treated zero trade flows as a data censoring problem and implemented their empirical strategy using a Tobit model. They also estimated their model on single and pooled three year cross-sectional data and studied the parameters of RTAs over time to identify any major expansions or contractions of trade flow before and after the policy intervention. These types of studies are not only used to analyze individual regional trade flows but also to address the broader question of whether RTAs are undermining the multilateral trade liberalization process[3].

Bayoumi & Eichengreen (1997) and Magee (2008) also analyzed multiple trade agreements simultaneously. Both incorporated interactions of RTA with time dummy variables capturing periodical changes in the level of regional trade before and after the creation of RTAs. Bayoumi & Eichengreen (1997) observed several European RTAs over a forty year period and found that the strongest evidence of trade creation was during the 1960s. They also found evidence of a reduction in European trade with the rest of the world. Magee (2008) accounted for zero trade flows in one of the three specifications in his study using a fixed effects Poisson Pseudo Maximum Likelihood estimator. Magee's findings suggest RTAs signed in the 1980s and early 1990s were more trade creating than trade diverting, implying the RTAs in his study had a beneficial overall effect.

Studies focusing on a single trade agreement using the gravity model have focused on US trading relations and the effects of NAFTA on trade flows. Clausing (2001) addressed the Canada – United States Free Trade Agreement and its effects on Canadian exports to the

[3] Regionalism is said to undermine multilateral trade liberalization or act as a 'stumbling block' if RTAs hinder trade liberalization with non-member countries. Frankel (1997) finds that by his research countries entering liberal trade agreements also tend to open up with non-member trading partners, suggesting RTAs support multilateral trade liberalization or act as 'building blocks' for international free trade (p. 227).

US. The time span of her study was from 1989 to 1994 and she used detailed tariff rates to identify the impact of the RTA on trade flows. She did not find evidence of trade diversion considering Canada's growth in exports to the US was not at the expense of third countries that did not obtain the same preferences as Canada. Fukao, Okubo, & Stern (2003) focused on the effects of NAFTA on US imports from 1992 to 1998. Their principal specification was a panel country and year fixed effects model with a dummy variable for NAFTA countries since its creation. They did not take into account zero trade flows in their estimation, but used a very detailed data set at the four digit level of commodity disaggregation. Additionally, rather than using trade values as the dependent variable they used shares of imports, indexed by commodity, country, and year. Similar to Magee (2008), their specification did not permit an analysis of import flows from non-member countries. They found evidence of trade diversion in the US in certain commodity categories.

Finally, another group of studies analyzed changes in trade share before and after the implementation of a RTA. Krueger (1999) performed such an analysis of NAFTA while Yeats (1998) did so for Mercosur. Krueger's (1999) study was based on a 'shift in share' analysis. Although she found an increase in U.S. imports from Mexico, seemingly at the expense of East Asian import shares, Krueger argued that Mexican exports had gained shares in other international markets as well and therefore Mexico's gain in the U.S. market was not evidence of trade diversion. Yeats (1998) analyzed several trade indices from 1988 to 1994. These indices revealed that products experiencing the greatest growth in regional trade shares were not competitive in international markets. The implication being that Mercosur's trade expansion reflected a pattern of trade diversion. One of his explanations for this finding was that Mercosur's tariff rates were four to six times higher than those of in European RTAs.

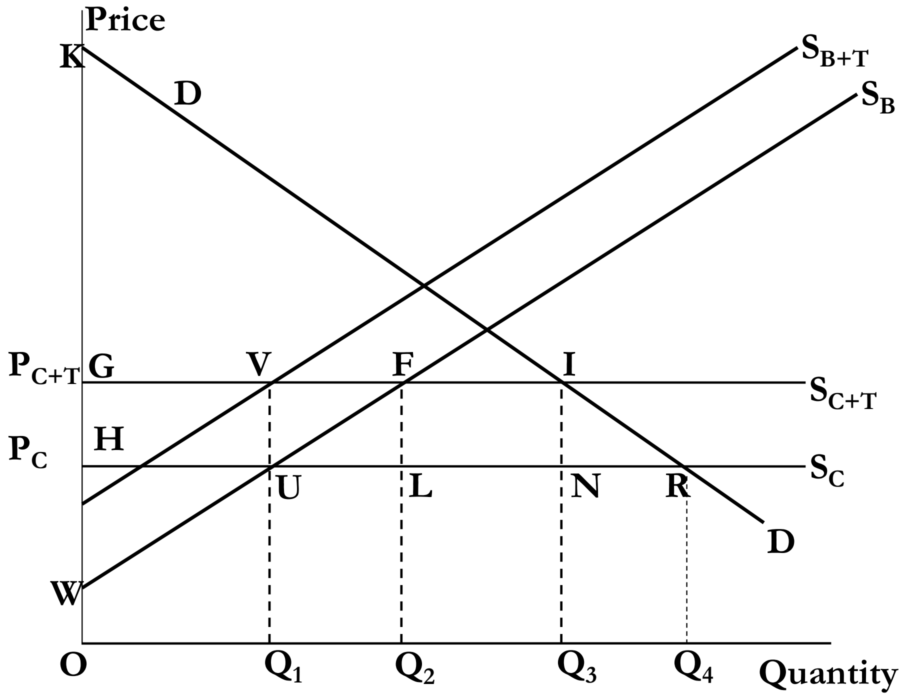
1.3 Theory of Trade Creation and Trade Diversion

One of the often cited gains from trade is that countries can achieve higher levels of economic welfare through trading than under autarky. Simply put, this is because countries can export goods that are more valuable in world markets abroad, and import goods which are relatively costly to produce domestically, from the rest of the world. All countries are better off in such a relationship because they achieve higher levels of consumption than would otherwise be possible. This is the argument for multilateral trade liberalization and it is also the argument that was initially used to promote preferential or regional trade liberalization. Policy-makers believed countries in regional trade agreements could obtain the gains from trade previously discussed.

Viner (1950) was one of the first to argue and demonstrate that under certain conditions, regional trade agreement generated losses rather than gains. He used the example of a customs union to demonstrate how countries might actually be worse off. Viner synthesized the benefits and losses of custom unions into the concepts of trade creation and trade diversion. In the Vinerian sense, trade creation occurs when it is more efficient to import a good from a partner in a trade agreement than to produce it domestically. This is a welfare increasing consequence of an RTA because it improves the importing country's terms of trade by expanding its consumption capacity. Trade diversion takes place when imports from efficient sources are shifted to inefficient sources that are benefitting from tariff preferences associated with membership to an RTA. This is a welfare-reducing consequence of regional trade agreements since the terms of trade of the importing country decrease by the amount of tariff revenue forgone in shifting imports to an RTA partner.

The theory of trade diversion and trade creation is based on a partial-equilibrium analysis of trade. Bhagwati and Panagariya (1996) extend Viner's model to highlight the effects on partner countries and the rest of the world, rather than focusing exclusively on the importing country. Assume there are three countries A, B, and C. C represents the rest of the world while A and B are potential union partners. Figure 2 models import demand of a given product in country A. A's import demand curve is represented by DD. The rest of the world's supply (C) is assumed to be perfectly elastic and represented by the horizontal supply curve S_C, while B's supply curve is upward sloping (S_B). Country B's upward sloping supply curve will permit an analysis of welfare change in the exporting country, a feature that is absent in the original work of Viner (1950). The model below permits the analysis of various scenarios, including the gains from trade under nondiscriminatory tariffs and free trade between A and B. Subsequently, I will present the extension of this analysis for the case of Paraguay, which is a member of the customs union Mercosur.

Figure 2 Partial Equilibrium Model



In a scenario of non-discriminatory tariffs the equilibrium price is P_{C+T} and A imports OQ_3 in total. Country B supplies OQ_1, and Q_1Q_3 is imported from C. The gains from trade for A are given by the area under the demand curve and above the international price-line free of tariff (KHNI). This area is made up of the consumer surplus (KGI) and tariff revenue (GHNI). The gains to B are represented by the area over the supply curve S_B and under the international price-line, P_C, represented by triangle HWU.

Suppose A enters a free trade agreement with B. That is, A discriminates imports from C by waiving tariffs on products from B only. Just as in the case of non-discriminatory tariffs, the equilibrium price is given by P_{C+T} and total import quantities are OQ_3. However, imports from B increase to OQ_2, while imports from C decrease to Q_2Q_3. Country A's consumer surplus is unchanged, but tariff revenues fall from GHNI to FLNI. Hence, A's welfare has decreased by GHLF. Country A has worsened its regional terms of trade by the amount of tariff revenue forgone. The quantity Q_1Q_2 is the extent of trade that is diverted to B.

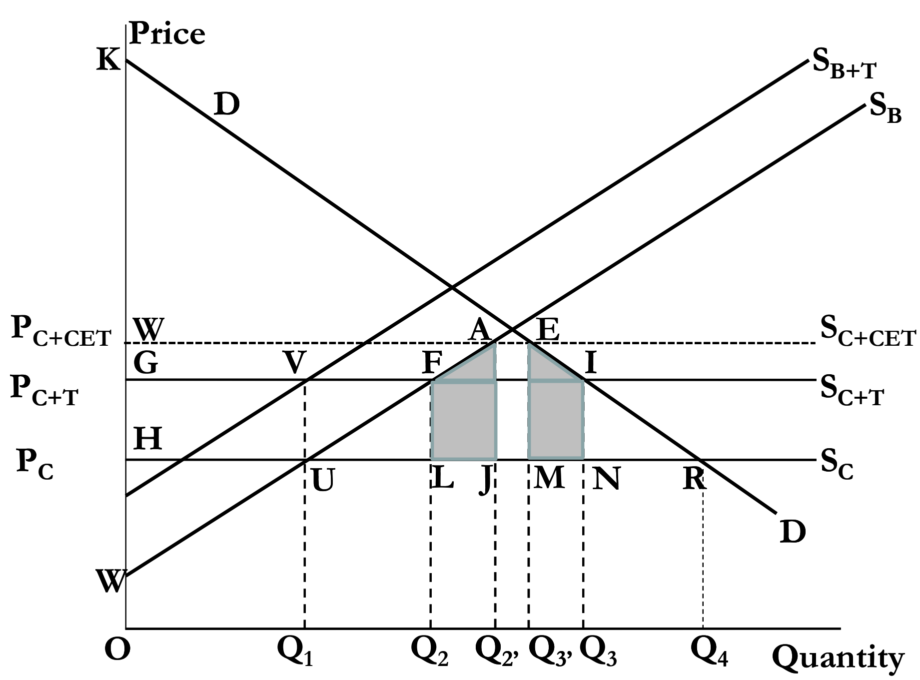
On the other hand, B obtains gains from trade that partially offset A's losses. B's producer surplus increases from HWU to GWF. Notice that Country B's gain (GHUF) is smaller than Country A's loss (GHLF). The excess loss of A is represented by the area FUL. This is an example of an overall welfare-reducing free trade agreement where the burden of the welfare loss falls completely on A.

Based on this analysis, Bhagwati and Panagariya (1996) also cast doubt on the 'natural trading partner' hypothesis. Briefly, the natural trading partner hypothesis claims that to the extent partners in a regional trading agreement trade disproportionately with each other before the creation of a bloc, trade diversion is less likely. However, as demonstrated in the explanation above, the greater the initial imports from the preferential trading partner the greater the loss to the importing country due to forgone tariff revenues.

Mercosur: The Case of Paraguay

It is possible to extend the analysis above to that of a country entering a customs union and converging up to the common external tariff, as I argue is occurring in Paraguay. Figure 2 adds an additional price-line, P_{C+CET}, which is drawn higher than P_{C+T} to reflect the idea of a country converging up to a common external tariff.

Figure 3 Partial Equilibrium Model of Paraguay



Under these conditions, country A reduces its total import demand from OQ_3 to $OQ_{3'}$. Additionally, imports from country B increase from OQ_2, in the case of a free trade agreement, to $OQ_{2'}$ reflecting an increase in the extent of trade diversion in this new environment. A's position is unequivocally worse by the amounts of the shaded areas in Figure 3. Area EMNI reflects welfare loss in country A associated with reducing total imports from OQ_3 to $OQ_{3'}$. It holds consumer surplus losses and government tariff revenue losses. Area AFLJ reflects overall welfare losses in the regional trade agreement due to the fact that country B gains to not completely offset country A's losses. As in the previous example, the burden of these welfare losses is completely absorbed by country A. As drawn, there is a slight tariff revenue gain in country A given by the trade volume $Q_{2'}$ $Q_{3'}$ and the marginal increase in the tariff rate.

In many respects, the model above is a reflection of what has theoretically occurred with Paraguay entrance to Mercosur. A small open economy entered an agreement with the obligation of converging up to a common external tariff. Paraguay's reluctance to converge is evident by the numerous postponements to applying the Mercosur common external tariff rates on numerous commodities, especially high technology products. However, the model does not fairly reflect the benefits to partner countries considering the minute scale of the Paraguayan economy compared to Brazil and Argentina. The model prompts the empirical question of identifying how the flows of trade into Paraguay have been affected by Mercosur. Additionally, if evidence of trade diversion is discovered, the objective will be to quantify this measure in the aggregate and individual sectors.

1.4 Data and Methodology

1.4.1 Data

The data collected to estimate this model are trade values, US CPI, real GDP, real GDP per capita, distance, common language, land area, and religion. Trade values at the one digit level of disaggregation were retrieved from the United Nations Comtrade database (http://comtrade.un.org/db/). The selected commodity classification system was the SITC Rev.1 which has data available for years even before 1970. Import trade values were deflated using the US CPI. Geographic variables were collected from the CEPII (Centre d'Etudes Prospectives et d'Informations Internationales) and GDP related variables were collected from the USDA Economic Research Service. Finally, information on religion was collected from the CIA World Factbook. The data collected on religion was used to create a religion index[4].

1.4.2 The Gravity Model

The most basic form of the gravity model defines trade between two countries as a function of the product of their GDPs, populations (or GDPs per capita), and the distance between them. Other variables that are also sometimes included control for cultural affinity and land area. Land area supplements economic size variables since it incorporates information about natural resources to the model (Frankel, 1997). The dependent variable varies across studies. Some studies use the sum of import and export flows as the dependent variable (Frankel, 1997; Bayoumi and Eichengreen, 1997) while others focus on a single trade flow, usually import flows, when the objective is to analyze trade diversion and trade creation (Soloaga & Winters, 2001; Fukao et al., 2003; Clausing, 2001; Magee, 2008). In this paper I also use import flows as the dependent variable and following Frankel (1997), the basic specification, for cross-sectional data, takes the following form:

$$\ln\left(I_{ij}\right) = \beta_0 + \beta_1 \ln\left(GDP_i * GDP_j\right) + \beta_2 \ln\left[\left(\frac{GDP}{pop}\right)_i \left(\frac{GDP}{pop}\right)_j\right] + \beta_3 ln\left(Dist_{ij}\right) + \beta_4\left(Lang_{ij}\right) + \varepsilon_{ij} \qquad (1)$$

In this model I_{ij} represents the value in current US dollars of country *i*'s imports from country *j*, the product of GDPs captures the size of the economies while the product of GDPs per capita gathers information on income and population. Distance is theorized to be inversely proportional to trade volumes and cultural affinity in the dummy variable for language ($Lang_{ij}$) enters the equation positively.

[4] The religion was calculated following Helpman et al (2008): (% Protestants in Paraguay X % Protestants in country *j* + % Catholics in Paraguay X % Catholics in country *j*). No other religions outside Christianity are reported in Paraguay.

To analyze the effects of regional trade agreements, the specification above is usually augmented to include a set of dummy variables representing intrabloc and extrabloc trade. A positive and significant coefficient on the parameter representing membership is interpreted as trade in excess of what is predicted by the gravity model and is considered evidence of trade creation. A negative and significant coefficient on the dummy variable for extrabloc trade is evidence of less trade than predicted by the basic specification, and interpreted as evidence of trade diversion. Consequently, the magnitudes of these parameters play a significant role in determining what the overall effect of a regional trade agreement has been for a country or group of countries.

1.4.3 The difference-in-difference estimator

A criticism of some RTA studies based on the gravity model is that often, researchers assume any increase to regional trade since the signing of an agreement is mostly attributable to the agreement, ignoring the possibility that cultural and historical links have played a significant role in explaining disproportionately high levels of trade between neighbors and even the formation of regional trade agreements (Magee, 2008; Soloaga & Winters, 2001, Bayoumi & Eichengreen, 1997). Such studies are often cross-sectional, and often treat the existence of an RTA as fixed over time. The implication of this criticism is that the formation of an RTA is correlated with the error term resulting in biased estimates. Incorporating the time dimension into the analysis permits the utilization of the difference-in-difference estimator, which is useful to capture valuations of any increase in regional imports since the creation of an RTA in excess of the trade attributable to historical or cultural ties.

I incorporate a re-parameterization of the difference-in-difference estimator and combine it with the gravity model to obtain parameters which measure the magnitude of the impact of Mercosur on Paraguayan imports from both member and non-member countries, which is an advance over the previous literature which has exclusively captured measures of trade creation. For practical purposes, this methodology treats member countries as part of a treatment group and remaining trading partners as controls.

The application of the difference-in-difference estimator to Paraguayan aggregated import flows takes the following form:

$$\ln(I_{jt}) = \gamma_0 + \gamma_1 Post_t + \gamma_2 Mercosur_j + \gamma_3 (Post \cdot Mercosur)_{jt} + \varepsilon_{jt} \quad (2)$$

The subscript *j* on the dependent variable indexes the exporters to Paraguay, while *t* indexes years. The variable $Post$ is a dummy variable taking a value of one for all years since the creation of Mercosur. The variable $Mercosur$ (M) takes a value of one for all Mercosur member countries for all years. Finally, $Post \cdot Mercosur$ (PM) is the interaction of these two dummy variables and consequently, γ_3 is the average treatment effect of the Mercosur policy variable. The variable *PM* takes on a value of one for all Mercosur countries since the creation of Mercosur. That is, the variable *PM* represents membership to Mercosur, so I will use *Members* and *PM* interchangeably from here on. After exponentiating[5] *PM*, it represents the average percent change in imports from non-member countries subtracted from the average percent change in imports from member countries for two given periods. More formally,

$$\gamma_3 = \left[\overline{\ln(I_{M2})} - \overline{\ln(I_{M1})}\right] - \left[\overline{\ln(I_{NM2})} - \overline{\ln(I_{NM1})}\right] \quad (3)$$

where the subscript *M* refers to Mercosur countries and *NM* to non-Mercosur countries, while the subscripts *1* and *2* refer to the years before and after the creation of Mercosur, respectively.

Intuitively, the average percent change in imports from non-member countries before and after the creation of Mercosur is subtracted from the average percent change in imports from members. This estimator controls for both cultural and historical links of member countries (embedded in '$Mercosur$' of equation (2). '$Post$' in the same expression captures aggregate or time factors, that would have affected trade flows of both members and non-members even if Mercosur had not been created (Wooldridge, 2002).

The concepts of trade creation and trade diversion are implicit in γ_3 of the expression above. To make these concepts explicit I reparameterized the difference-in-difference specification in (2) as follows:

[5] The exact percent change in the predicted value of I when $Post \cdot Mercosur$=1 versus when $Post \cdot Mercosur$=0 is given by $100(e^{\gamma_3} - 1)$.

$$\ln(I_{jt}) = \gamma_0 + \gamma_2 M_j + \gamma_3 (PM)_{jt} + \gamma_4 (PNM)_{jt} + \varepsilon_{jt} \tag{4}$$

where *M* and *PM* have been previously defined and *PNM* ($Post \cdot NonMercosur$) takes on a value of one for non-member countries since Mercosur's creation. This is justified given the previously defined dummy variable *Post* is a linear combination of *PM* and *PNM*

$$Post_t = (PM)_{jt} + (PNM)_{jt}. \tag{5}$$

This decomposition of the dummy variable $Post$ will facilitate the identification of the average percent change in import flows from member and non-member countries since Mercosur's creation. Hence, in the re-parameterized form (4), the coefficients have the following interpretation:

$$\gamma_3 = \overline{\ln(I_{M2})} - \overline{\ln(I_{M1})} \tag{6}$$

$$\gamma_4 = \overline{\ln(I_{NM2})} - \overline{\ln(I_{NM1})} \tag{7}$$

The benefit of this methodology is that difference-in-difference estimators control for unobserved factors determining high or low levels of bilateral trade among members of an RTA. The $\gamma_3$estimate above capture trade creation or the average percent change in regional imports since the creation of Mercosur, while γ_4 captures trade diversion, or the average percent change in imports from the rest of the world since the policy went into effect.

I estimate a benchmark pooled cross-sectional Ordinary Least Squares (OLS) model taking the following form:

$$\begin{aligned} \ln(I_{kjt}) = \beta_0 + \beta_1 t + \beta_2 \ln(GDP_{Pyt} \cdot GDP_{jt}) + \beta_3 \ln\left[\left(\frac{GDP}{pop}\right)_{Pyt} \left(\frac{GDP}{pop}\right)_{jt}\right] \\ + \beta_4 ln(Dist_j) + \beta_5(Area_j) + \beta_6(Lang_j) + \gamma_2(M_j) \\ + \gamma_3(PM_{jt}) + \gamma_4(PNM_{jt}) + \varepsilon_{kjt} \end{aligned} \tag{8}$$

Recall the data are at the one digit level of commodity disaggregation. Hence, the additional subscript *k* is introduced to index commodities[6]. Considering only imports into Paraguay are analyzed, the subscript *j* indexes exporting countries and *Py* reflects Paraguayan values. Parameter $\beta_2 - \beta_6$ reflect the gravity and additional variables used to predict 'natural' flows of trade from the rest of the world to Paraguay. The last three parameters in (8), $\gamma_2 - \gamma_4$, reflect the re-parameterization of the difference-in-difference estimator used to capture measures of trade creation and trade diversion. Recall M_j is time invariant and consequently is dropped from the difference-in-difference specification. I use 1995 as the beginning year of Mercosur considering it was the year regional trade was originally planned to be liberated, as well as when convergence to the common external tariff began.

An additional parameter in (8) is β_1, the linear time trend coefficient. It is included in the specification to capture trending factors such as global inflation and growth. The usual method to do this is by including year fixed effects. However, in order to keep the difference estimators γ_3 and γ_4 in equation (8), it is not possible to include year dummy variables because the year fixed effects starting in 1995 form a linear combination with the dummy variables PM_{jt} and PNM_{jt}. All models are estimated using robust standard errors in order to correct for serial correlation and heteroskedasticity in the error terms (Wooldridge, 2002).

1.4.4 Zero trade flow in trade data

This data set is made of import flows from the rest of the world to Paraguay from 1970-2010. In the data, there are a total of 155 countries that export to Paraguay in at least one year and one commodity during this time. Consequently, after imputing the 'missing' observations into the data as zero trade flows, these represent 79% of the observations. As Frankel (1997) explains, missing values in a data set are attributable to little expected trade between countries because of size of the economies and remoteness from each other. Considering Paraguay is a small landlocked country and the trade values are at the one-digit level of detail, it is highly likely many countries will not export to Paraguay in all commodity codes in all years. I follow the traditional method of treating missing trade flows as the absence of trade between countries and treat these observations as zero trade flows.

[6] There are ten commodity codes ranging from 0 to 9.

Econometrically, zero flows of trade represent a methodological challenge since the natural logarithm of zero is undefined. The literature describes several methods of dealing with zero trade flows. A very common approach is to ignore these missing observations. Frankel (1997), Fukao et al (2002) explain this is what they have done in their study and it is likely Clausing (2001) also followed this approach considering the data in her study was rich in detail. Other techniques include substituting arbitrary small numbers for the zeroes, such as $1,000, or in constructing the dependent variable taking the natural log of the value of imports plus one. A third approach in dealing with this problem is using a Tobit model given trade values are bounded from below by zero. This is the approach followed by Soloaga & Winters (2001). However, the appropriateness of using the Tobit model has been questioned by Linders & de Groot (2006) on the basis the Tobit model would be justified if the censored data reflected negative trade values, or if the dependent variable existed but was unobservable. Frankel (1997) justifies the omission of the zero trade flows arguing the final results are not very sensitive to insertion of zero flows of trade. Soloaga & Winters (2001) make the same argument, claiming their results are robust to either the Tobit approach or the more traditional OLS methods omitting the zero flows of trade.

The main concern with ignoring the zero trade flows is that results based on this approach may be biased. More recently, two widely cited approaches have been developed to handle the problem of zero trade flows. One solution is estimating the model using the import values measured in levels and implementing a Poisson maximum-likelihood estimator, traditionally used in estimating count regression models (Silvia and Tenreyro, 2006). The second approach treats the existence of zero trade flows as a sample selection problem and relies on the implementation of the Heckman two-step procedure for sample selection correction (Helpman et al., 2008; Linders & de Groot, 2006).

I follow the second approach in this paper for one main reason. First, considering Paraguay's relatively remote geographic location and developing country status, it is likely there are numerous countries in similar conditions raising transaction costs. The procedure based on the Heckman procedure models the probability of trading events to be realized conditioned on country characteristics.

1.4.4.1 Heckman Sample Selection Model in a gravity application

In order to correct potentially biased estimates, the Heckman two-step estimator is used to include the zero trade flows in this analysis and correct the sample selection problem. As its name suggests, this approach follows two steps. In the first stage or selection equation, the existence of a trading relationship is modeled using a probit model. Hence, the positive-valued trade observations take a value of one and all missing or zero-valued observations are censored at zero in the probit model. The predicted values of this model are used to construct the inverse Mills ratio, which is augmented to the outcome equation in the second stage, the gravity model in this application. The second stage or outcome equation is estimated on the uncensored observations only.

This approach has several advantages. First, it adds flexibility to the comparable Tobit model, which restricts the censoring mechanism to be part of the outcome equation, the gravity model in this application. The two-step procedure models the processes separately. Using an OLS estimator in the second stage rather than maximum-likelihood methods permits the adoption of weaker distributional assumptions of the joint distribution of the error terms of both models (Cameron and Travedi, 2005). A priori, any potential correlation between the error terms of the first and second stages can be verified and corrected by this procedure. A lack of correlation between the error terms Cov($\varepsilon_{kjt}, u_{kjt}$)=0, is indicative of no sample selection bias. Notice ε_{kjt} refers to the error term of the model prior to the implementation of the correction procedure.

Consequently, I estimate the following model:

Selection Equation:

$$T_{kjt} = \beta_0 + \beta_1 t + \beta_2 \log\left(GDP_{Pyt} \cdot GDP_{jt}\right) + \beta_3 \log\left[\left(\frac{GDP}{pop}\right)_{Pyt}\left(\frac{GDP}{pop}\right)_{jt}\right] + \beta_4 log\left(Dist_j\right) + \beta_5\left(Area_j\right) + \beta_6\left(Lang_j\right) + \beta_7\left(Rel_j\right) + \gamma_2\left(M_j\right) + \gamma_3\left(PM_{jt}\right) + \gamma_4\left(PNM_{jt}\right) + u_{kjt} \qquad (9)$$

$$T_{kjt} = \begin{cases} 1 & Trade\ Volume > 0 \\ 0 & Trade\ Volume = 0 \end{cases}$$

Outcome Equation:

$$\ln(I_{kjt}) = \beta_0 + \beta_1 t + \beta_2 \log(GDP_{Pyt} \cdot GDP_{jt}) + \beta_3 \log\left[\left(\frac{GDP}{pop}\right)_{Pyt}\left(\frac{GDP}{pop}\right)_{jt}\right]$$
$$+ \beta_4 log(Dist_j) + \beta_5(Area_j) + \beta_6(Lang_j) + \gamma_2(M_j)$$
$$+ \gamma_3(PM_{jt}) + \gamma_4(PNM_{jt}) + \lambda(\hat{T}_{kjt}) + \theta_{kjt}$$

All of the variables in (9) have been discussed except Rel_j in the selection equation. This variable represents an index for religious similarity between trading partners. The higher the index the more similar the trading partners. It is included in the selection equation to meet the exclusion restriction recommended for two-part Heckman models such as (9)[7]. The exclusion restriction requires a variable to be correlated with inclusion in the sample (dependent variable in the selection equation), but uncorrelated with the error term in the outcome equation. That is, to meet the exclusion restriction requirements 'religion' should be positively related with the probability of countries trading with each other, but it should be uncorrelated with volume of trade in the outcome equation. This is also the variable suggested and used as an exclusion restriction in Helpman et al. (2008), which is by the author's account, mostly a methodological paper developing an estimation procedure that corrects for potential biases in standard gravity equations.

Additionally, the outcome equation includes the inverse Mills ratio $\lambda(\hat{T}_{jt})$. This is the variable included in the gravity model to correct for sample selection bias. A statistically significant coefficient on this variable is consistent with evidence of sample selection bias.

1.4.4.2 Bias from time invariant unobserved country specific factors

The significance of controlling for unobserved factors which may contribute to explain why a pair of countries might have a high or low levels of bilateral trade was discussed to be one of the benefits of using the difference-in-difference estimator in the gravity model

[7] Both Cameron & Travedi (2005) and Wooldridge (2002) explain that both stages may be run using the same regressors omitting the exclusion restriction, but doing so risks the identification of the model and multicollinearity issues in the outcome equation due to an approximately linear inverse Mills ratio term over a wide range of arguments. However, this problem is attenuated by greater variation in the participation model. In this application the estimates are insensitive to whether an excluded variable is incorporated or not. Interestingly, Linders & de Groot (2006) apply a two-step Heckman selection model with maximum likelihood on the same set of regressors for each stage.

environment. Although the difference-in-difference estimator controls for these unobserved factors of member countries it does not control for unobserved factors explaining excess or limited trade with non-member countries. Introducing country fixed effects to the model will removes country-level heterogeneity bias associated with time invariant unobserved factors affecting the flows with all countries. The Heckman model now takes the following form:

Selection Equation:

$$T_{kjt} = \beta_0 + \beta_1 t + \beta_2 \log\left(GDP_{Pyt} \cdot GDP_{jt}\right) + \beta_3 \log\left[\left(\frac{GDP}{pop}\right)_{Pyt}\left(\frac{GDP}{pop}\right)_{jt}\right] +$$

$$\gamma_3\left(PM_{jt}\right) + \gamma_4\left(PNM_{jt}\right) + C_j + \cdots + C_N + u_{kjt} \qquad \text{for } j = 1, \ldots, N$$

$$T_j = \begin{cases} 1 & Trade\ Volume > 0 \\ 0 & Trade\ Volume = 0 \end{cases} \qquad (10)$$

Outcome Equation:

$$\ln\left(I_{kjt}\right) = \beta_0 + \beta_1 t + \beta_2 \log\left(GDP_{Pyt} \cdot GDP_{jt}\right) + \beta_3 \log\left[\left(\frac{GDP}{pop}\right)_{Pyt}\left(\frac{GDP}{pop}\right)_{jt}\right] +$$

$$\gamma_3\left(PM_{jt}\right) + \gamma_4\left(PNM_{jt}\right) + \lambda\left(\hat{T}_{kjt}\right) + C_j + \cdots + C_N + \theta_{kjt} \qquad \text{for } j = 1, \ldots, N$$

The country fixed effects are reflected in the C_j dummy variables for N countries. This methodology significantly changes the original proposed specification considering time invariant variables can no longer be included in the model. In addition to controlling for time invariant factors affecting trade, including these fixed effects reduces the possibility that the researcher's prior beliefs influence the specification of the gravity model Magee (2008). Several references recommend this technique for gravity model estimation (Matyas, 1997; Anderson & Van Wincoop, 2003; Baldwin, 2006; Helpman et al., 2008). As before, a pooled OLS country fixed effects model is also estimated for comparison.

1.5 Results

The estimation results are based on import values from Paraguay's 155 exporters during the sample period[8]. Before reporting the main regression results I provide a glance at the structure of Paraguayan import values from 1970-2010. Table 1 reports the value and share of zero valued trade flows included in the data set. Clearly, the group of observations that may be categorized as manufactured merchandise has the highest number of positive trade value frequencies with respect to the potential number of observations[9]. 'S1-6 Manufactured goods classified chiefly by material', 'S1-7 Machinery and transport equipment', and 'S1-8 Miscellaneous manufactured articles' all have ratios of censored observations or zero trade flows below 70%. By no coincidence, they are also the categories of greatest import values only clearly surpassed by 'S1-3 Mineral fuels'. This data is consistent with the structure of the Paraguayan economy being an exporter of mainly agricultural products.

Table 1 Total value of Paraguayan imports by commodity in constant 2005 $US, 1970-2010.

Class. Code	Commodity Description	Censored Share	Imports (millions)
S1-0	Food and live animals	81%	4,432
S1-1	Beverages and tobacco	82%	6,477
S1-2	Crude materials, inedible, except fuels	81%	1,140
S1-3	Mineral fuels, lubricants and related materials	87%	14,185
S1-4	Animal and vegetable oils and fats	94%	112
S1-5	Chemicals	71%	11,794
S1-6	Manufactured goods classified chiefly by material	69%	11,584
S1-7	Machinery and transport equipment	68%	33,355
S1-8	Miscellaneous manufactured articles	67%	10,160
S1-9	Commod. & transacts. Not class. Accord. To kind	91%	128

Source: UN Comtrade database.

In the following section two sets of results are reported. The first set omits country fixed effects while the second set includes them, controlling for country-level heterogeneity. Each set holds pooled OLS and Heckman sample selection specifications. The Heckman

[8] The 155 partners are reported in Table A.1 of the appendix.

[9] The total number of potential observations for each commodity category is given by the number of countries by the number of years (41X155)=6355.

procedure and country fixed effects each correct a specific kind of bias, sample selection and heterogeneity, respectively. The results suggest the impact of Mercosur on Paraguayan imports flows is less sensitive to sample selection than heterogeneity bias. The estimates of the average percent increase in imports from members since the creation of Mercosur is consistently 322% in the first set of models accounting for selection bias but ignoring country-level heterogeneity (Table 2). Once heterogeneity is also accounted for, this figure is reduced to 267%. Consequently, omitting country fixed effects from the Heckman model overestimates member imports. Interestingly, it also underestimates non-member imports. Finally, I find evidence of great trade creation in *Beverages and Tobacco* and *Animal and vegetable oils & fat*. Contrary to theory, we find no clear evidence of trade diversion.

1.5.1 Two Step Heckman Procedure

The first set of results presented is a comparison of specifications (7) and (8) described in the previous section. The results are presented in Table 2 below.

As discussed earlier, any correlation between the error terms of the selection and outcome equation prior to correction, suggests estimates ignoring sample selection will be biased. The output suggests the error terms between the selection and outcome equations are positively correlated, Cov($\varepsilon_{kjt}, u_{kjt}$)=0.515. Similarly, a standard t-test on the coefficient of the of the inverse Mills ratio (λ), is a valid test of the null hypothesis of no selection bias (Wooldridge, 2002). Given the coefficient is statistically significant at the 99% confidence level, this hypothesis is rejected.

Inspecting the coefficient estimates on the outcome model and the benchmark, standard gravity model based on pooled OLS regression reveals that the magnitude of all coefficients of the benchmark are lower than those of the outcome equation from the Heckman two-step approach, indicating the benchmark model underestimates the true parameters. The coefficient on *Religion* in the selection model below is positive and statistically significant, which suggests sharing a common religion increases the probability of establishing a trading relationship. The difference in magnitude between the gravity model across the pooled OLS and Heckman model is quite large in all the coefficients. In fact, this difference in magnitude is statistically significant at the 95% confidence level for all coefficients except

GDP per capita and *PM (Member)*. The estimate on PNM (*Non-member*) is zero or almost zero in both models[10].

Almost all estimates are highly statistically significant and there are no contradictions in signs of the coefficients. Additionally, notice the semi-elasticity of imports on *PM (Member)* is 1.44 in both the OLS model and the outcome equation in the Heckman procedure. Hence, to some degree, this is consistent with claims in the earlier literature that ignoring the zero trade flows does not greatly affect regression estimates.

Table 2 Gravity model on Paraguayan import flows from 1970-2010.

Import Flows	Pooled OLS		Heckman two-step			
			Selection		Outcome	
	Est.	P-val.	Est.	P-val.	Est.	P-val.
Time, (β_1)	-0.039	0.000	-0.026	0.000	-0.061	0.000
GDP, (β_2)	1.053	0.000	0.654	0.000	1.569	0.000
GDP per capita, (β_3)	-0.340	0.000	-0.083	0.000	-0.367	0.000
Distance, (β_4)	-1.008	0.000	-0.592	0.000	-1.432	0.000
Area, (β_5)	-0.277	0.000	-0.168	0.000	-0.421	0.000
Language, (β_6)	0.294	0.000	0.377	0.000	0.812	0.000
Mercosur , (γ_2)	1.482	0.000	0.848	0.000	1.949	0.000
PM (*Member*), (γ_3)	**1.438**	**0.000**	**0.311**	**0.014**	**1.439**	**0.000**
PNM (*Non-member*), (γ_4)	**0.054**	**0.497**	**0.337**	**0.000**	**0.327**	**0.002**
Religion (β_7)			0.157	0.000		
Intercept	-20.167	0.000	-22.494	0.000	-40.011	0.000
Mills (λ)					1.357	0.000
$\rho_{\varepsilon u}$			0.515			
σ_{ε}			2.636			
No. Observations	13,260		63,550			
Uncensored Obs.			13,260			
R^2	0.3082					

Note: All coefficients with the exception of 'PNM (*Non-member*)' in the OLS model are significant at the 5% level.

A result that differs from that in the traditional literature is the sign of the coefficient of *GDP per capita*. The basic 'gravity variables' in the model are *Distance*, real *GDP*, and real *GDP per capita*. By the Helpman-Krugman theory of trade, a theoretical foundation of the gravity model, trade is expected to increase in GDP and GDP per capita, and decrease with

[10] Calculated based on a two sample t-test.

distance. They propose countries with similar levels of productivity, as measured by GDP per capita, will trade more than countries with dissimilar levels (Frankel, 1997). Hence, developed countries with high GDPs per capita are more likely to trade with each other, which is usually reflected in a positive coefficient on *GDP per capita*.

On the contrary, by the Heckscher-Ohlin theory of trade, countries with dissimilar levels of output per capita are expected to trade more because countries are expected to export resource-abundant products while they import resource-scarce varieties, depending on their labor and capital endowments (Frankel, 1997). Similarly, countries with high capital endowment ratios are likely to have higher labor productivity levels[11], which are reflected in higher *GDP per capita* levels. The negative sign on *GDP per capita* variable in Table 2 suggests Paraguayan imports decline in partner capital endowment ratios (GDP per capita), implying Paraguay imports more from countries with similar levels of output per capita, which is consistent with the Helpman-Krugman theory of trade. Considering *GDP per capita* also reflects country population levels[12], this implies Paraguayan imports less from largely populated countries. GDP, population, and land area may be thought of as measures of country size which tend to have lower shares of trade to GDP ratios than smaller countries such as Singapore or Luxembourg.

The underestimation of the benchmark model is most evident on the coefficient for real GDP in Table 2. The *GDP* estimate in the outcome equation is greater than the benchmark by 0.51%. Land area, a proxy for resource endowment, has the expected negative sign. Its estimate in the outcome equation suggests imports decrease by 0.42% for every percent increase in partner land area.

A series of binary variables follow in the specification including the difference-in-difference related parameters. First, the estimate on common language increases trade in the outcome equation by 125% [100 *$(e^{0.812} - 1)$ = 125%]. Next, the difference-in-difference related coefficients are informative on how the creation of Mercosur has affected import flows to Paraguay since the initial year regional trade was planned to be liberalized, 1995. The *Mercosur* variable captures historical time invariant factors affecting regional imports which

---

[11] This argument borrows from Carrere (2006), who described *GDP per capita* as a proxy for the capital-endowment ratio and hypothesized a negative sign on such a coefficient.

[12] Gravity models may be estimated using either GDPs per capita or population without other estimates in the specification being affected (Frankel, 1997).

are likely to make regional imports disproportionately high. This estimate is 1.949 in the outcome equation of the Heckman model, quite greater than the estimate on the benchmark 1.482. The estimate on this coefficient in the outcome model implies regional factors increase trade by an average of 602% than predicted by the basic gravity variables. This gives an idea of how disproportionate significance of Mercosur countries in Paraguay's import market.

Interestingly, magnitudes of the difference estimators across the Heckman and the pooled OLS model are very similar, implying these estimates are quite insensitive to sample selection bias. The estimate on *Members* is interpreted as the average percent increase in imports from Mercosur countries since 1995. Hence, by the Heckman model, imports from Mercosur countries have increased by a factor of 1.438 or 322% since 1995. Over the same period of time, imports from the rest of the world increased by 39%, as calculated from the estimate on *Non-Members*. The estimate on *Non-Members* in the benchmark model is lower (5.6%) and not statistically significant.

The intercept in the difference-in-difference model is interpreted as the mean value of the dependent variable for the base group. In this case, the base group is made up of non-member countries in the years prior to Mercosur. However, considering the difference-in-difference estimator is estimated along with the gravity model, its interpretation is not as clear. Bayoumi and Eichengreen (1997) give an explanation of the intercept in the gravity model as representing greater than proportional growth in trade with respect to GDP if the constant is negative, as it is in this model.

Although the estimates on the *Members* and *Non-Member* across the models are quantitatively different, considering the coefficient on *Non-Member* is not statistically significant in the benchmark OLS model, they do not contradict each other. By these results, there is evidence of trade creation in Paraguay as a result of Mercosur. Regional trade has categorically increased since 1995 and the results do not suggest this increased regional trade has come at the expense of imports from the rest of the world. Imports from the rest of the world have also increased by the results in the Heckman model, and they are at least unchanged according to the benchmark model. The coefficients on *Non-Member* would have to be negative and statistically significant to support an argument for trade diversion in Paraguay.

Similarly, in a separate set of unreported estimation results, I calculated the proper difference-in-difference estimator or average treatment effect of Mercosur for both sets of models presented in Table 2. This estimate corresponds to γ_3 in equation (2). Its values are in line with the findings thus far. The estimates in both models are statistically significant at the 95% confidence levels. Their values are 1.38 and 1.11 in the OLS and Heckman outcome equation, respectively. A two sample t-test yields the difference in these coefficients is not statistically significant. This result is also taken as evidence of trade creation since the magnitude of the coefficients indicates imports from regional partners have not been realized at the expense of imports from non-member.

1.5.2 Sensitivity Analysis using Country Fixed Effects

A shortcoming of the difference-in-difference methodology reported previously is that unobserved time invariant factors that may affect levels of trade of non-member countries are not controlled for. These unobserved factors may also introduce bias in the estimates and a common approach used to address this issue is by estimating the model using country fixed effects. Unobserved and observed time invariant variables are differenced out of the model in this approach and consequently, the list of explanatory variables is reduced to only the time-varying variables.

However, the interpretation of the difference estimators *Member* and *Non-Member* remains unchanged. As in the previous exercise, I estimate a benchmark pooled OLS model with country fixed effects along with the Heckman two-step model, also with country fixed effects. Several variables are removed from the specification due to their time invariance including *Distance, Area, Language,* and *Mercosur*. The variable *Religion* is also removed from the selection equation in the two-step procedure since the index for religion is also time-invariant. The results are reported in Table 3.

First, these results are consistent with previous estimates in that once again there is evidence of sample selection bias. The estimate on the Inverse Mills ratio in the two-step procedure is significant at levels higher than the 95% confidence level, justifying the sample selection correction. As before, the correlation between the error terms of the selection and outcome equations in the two-step procedure have a positive correlation of Cov($\varepsilon_{kjt}, u_{kjt}$)=0.632. Moreover, as in the previous set of results, the estimates in the benchmark model are an underestimate of the outcome equation in the Heckman model. The estimates in the benchmark model and the outcome equation in the two-step procedure are significant at the 95% confidence level except the coefficient on *Non-Member* in the benchmark.

Table 3 Fixed effects estimation results, Paraguayan import flows from 1970-2010.

			Heckman two-step			
Import Flows	Pooled OLS		Selection		Outcome	
	β	P-val.	β	P-val.	β	P-val.
Time (β_1)	-0.102***	0.000	-0.002	0.514	-0.119***	0.000
GDP (β_2)	2.484***	0.000	0.362***	0.000	3.172***	0.000
GDP per Cap. (β_3)	-1.272**	0.010	-0.093	0.165	-1.664***	0.000
PM, *Member* (γ_3)	**1.440*****	**0.000**	**0.116**	**0.371**	**1.299*****	**0.000**
PNM, *Non-member* (γ_4)	**0.159***	**0.088**	**0.281*****	**0.000**	**0.410*****	**0.000**
Intercept	-85.202***	0.000	-17.807	0.000	-114.53***	0.000
Mills (λ)					1.624***	0.000
Country Fixed Effects	Yes		Yes		Yes	
$\rho_{\varepsilon u}$			0.632			
σ_{ε}			2.567			
Observations	13,260		63,550			
Uncensored Obs.			13,260			
Overall R^2	0.1529					

Note: *, **, *** imply significance at the 10%, 5%, and 1% levels, respectively.

Gravity models without country fixed effects have been suspected of overestimating the impact of regional trade agreements. The reason being, without country fixed effects the model does not control for historical unobserved factors which might explain unusually high levels of trade between neighboring countries. Although I argued the difference-in-difference approach discussed in the previous section controls for these unobserved member specific factors, including country fixed effects reduces the membership estimates from about 1.440 in both OLS models and the Heckman approach without fixed effects, to 1.299 in the two step procedure with country fixed effects. This figure represents a change in the average increase of Mercosur imports from 322% to 266%. The average percent changes in imports based on the estimates of our main parameters of interest, 'PM - Member (γ3)' and 'PNM Non-Member (γ4)', of all implemented specifications are summarized in Table 4.

Table 4 Percent change in imports from member and non-member countries with and without fixed effects and by estimator.

Parameters	No Fixed Effects		Fixed Effects	
	OLS	Heckman	OLS	Heckman
PM, *Member* (γ_3)	321.2%***	321.6%***	322.1%***	266.5%***
PNM, *Non-Member* (γ_4)	5.6%	38.6%	17.3%*	50.7%***

Note: *, **, *** imply significance of estimates at the 10%, 5%, and 1% levels, respectively.

Additionally, including country fixed effects reveals an increase of greater magnitude in average imports from non-member countries with respect to both, the pooled OLS and Heckman specifications. Hence, omitting controls for country-level heterogeneity underestimates imports from non-member trading partners. Imports from non-members have increased by 50.7%, rather than 38.6% as shown in Table 4. Regardless of this change in magnitude, the original interpretation of these results stands. I find no empirical evidence sustaining that the increase in imports from regional partners has been offset by a decrease in imports from the rest of the world. These empirical results support a finding of trade creation in Paraguay.

Regarding the average treatment effect of Mercosur on Paraguay using the parameterization of the gravity model consistent with equation (2) and country fixed effects, the unreported difference-in-difference estimators are 1.28 and 0.89 in the pooled OLS and Heckman models, respectively. Although the difference in magnitude is not statistically significant at the 95% confidence level, it is once again evidence that the sample selection procedure attenuates the magnitude of the OLS difference-in-difference estimator. The interpretation of the Heckman model is that membership to Mercosur has increased imports from the region net of the increase in imports from non-members an average of 143% since 1995 (260% using pooled OLS with fixed effects). Although this estimate is smaller than the specifications omitting fixed effects, this result continues to be consistent with trade creation.

1.5.3 Sector analysis using the pooled OLS model with fixed effects

To make use of the richness of the data, I have also estimated the Heckman model with country fixed effects at the commodity level. Table 5 introduces the results of ten regressions, each at the commodity level of disaggregation. The columns reported under

Member and *Non-Member* are the coefficient estimates of the specification described by (10). Columns (3) and (4) report the exact percentage interpretations of the coefficient estimates in columns (1) and (2). The number of countries exporting to Paraguay within each commodity category and the number of observations are also reported in columns (5) and (6), respectively. The share of censored data for each regression is reported in column (6).

As discussed in Table 1, Table 5 shows that most import observations take place in the following manufactured commodity categories: *Chemicals*, *Manufactured goods classified by materials*, *Machinery & transport equipment* and *Miscellaneous manufactured articles*. Not only is the number of observations the greatest in these categories, but so is the number of exporting countries, giving an indication of the significance of manufactured goods in imports. Though the coefficient for *Member* is statistically significant at least at the 90% confidence level in all regressions except *Mineral Fuels & Lubricants*, the coefficients for *Non-Member* are statistically significant for fewer classifications. In addition to being not statistically significant in *Mineral Fuels & Lubricants*, *Non-Member* is not significant in *Food and Live Animals* and *Animal and vegetable oils & fats*.

The greatest import expansions are observed in *Beverages and Tobacco*, and *Animal and vegetable oils & fats*[13]. The average percent increases in imports in these categories are, 3,816% and 5,325%, respectively. Although these import expansions reflect examples of great trade creation, they are also commodities in which the common external tariff (CET) is effective. This implies these merchandise categories may be benefitting from certain protection. For example, beginning in 2006 the CET on vegetable oils has been applied at 10%. Additionally, the CET on tobacco and beverages ranges from 14-20% (Chapter 22 HS-2006 – Mercosur Secretariat). Although the greatest expansion of imports is in commodities where the CET is effective, they are also categories in which Paraguay has a production capacity. Additionally all Mercosur countries are recognized producers of primary goods such as those included under *Animal and vegetable oils & fats*. Furthermore, the coefficient on *Non-Member* for *Beverages and Tobacco*, and *Animal and vegetable oils & fats* is positive and statistically significant for the former and negative but not statistically significant for the latter. These results attenuate the claim of trade diversion and support the argument of trade creation. A similar primary product category such as *Food and Live Animals* also exhibited

[13] The category *Commodities & transactions not classified according to kind.* also encountered a great expansion but is not discussed due to the vague description of the merchandise included in this category.

increases in both regional and international imports by 201% and 15%, respectively, results consistent with trade creation.

Table 5 Impacts of Mercosur on Paraguayan imports at the one digit SITC Rev. 1 level.

Merchandise Description	(1) Member (p-value)	(2) Non-Member (p-value)	(3) Member % Change	(4) Non-Member % Change	(5) No. of Ctry.[1]	(6) Obs. (Cens. Ratio)[2]
S1-0 Food and live animals	1.103*** (0.000)	0.136 (0.398)	201%	15%	72	2,952 (59%)
S1-1 Beverages and tobacco	3.668*** (0.000)	0.391** (0.018)	3,816%	48%	73	2, 993 (61%)
S1-2 Crude materials, inedible, except fuels	0.69* (0.079)	0.394** (0.037)	99%	48%	89	3,649 (66%)
S1-3 Mineral fuels and lubricants	0.219 (0.665)	-0.392 (0.235)	25%	-32%	62	2,543 (69%)
S1-4 Animal and vegetable oils & fats	3.994*** (0.000)	-0.279 (0.484)	5,325%	-24%	34	1,394 (73%)
S1-5 Chemicals	1.565*** (0.000)	0.473*** (0.000)	378%	61%	104	4,264 (57%)
S1-6 Manuf. goods classified by material	1.100*** (0.000)	0.517*** (0.000)	111%	68%	110	4,510 (56%)
S1-7 Machinery & transport equipment	0.563* (0.057)	0.481*** (0.000)	76%	62%	125	5,125 (61%)
S1-8 Miscellaneous manuf. articles	1.345*** 0.000	0.471*** 0.000	284%	60%	126	5,166 (60%)
S1-9 Commod. & transacts. Not class.	4.050*** (0.001)	3.115*** (0.006)	5,641%	2,153%	64	2,624 (79%)

Note: ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. 1) Number of countries. 2) Censored Ratio.

Imports in manufactured goods have also expanded significantly, especially *Chemicals* and *Miscellaneous manufactured articles*. Imports in these products have increased by 378% and 284%, respectively. Imports from non-member countries have also increase by 61% and 68%, respectively. This is also evidence of trade creation.

The same is true of *Manufactured goods classified by material* and *Machinery & transportation equipment* which have experienced statistically significant regional and extra-regional import expansions since 1995, denying any claim of overall trade diversion. Recall, it is in the manufactured good categories where there are numerous exemptions to the CET, especially in capital goods and information and communication technology equipment.

1.6 Conclusions

In this paper I used the gravity model of international trade to analyze how the creation of Mercosur has affected import flows into Paraguay. In order to account for zero trade flows I compared a benchmark pooled OLS model to an application of the Heckman model with and without country fixed effects. The Heckman model corrected an underestimation of 'gravity' variables in the pooled OLS estimates attributed to sample selection bias. However, there was no noticeable difference in the impact of membership to Mercosur between these two models. Including country fixed effects in the model decreased this estimate (*PM-Member*) and increased the estimate for imports from the rest of the world (*PNM - Non-member*). This result highlights the significance of correcting for time invariant bilateral trading characteristic across all countries. That is, heterogeneity bias had a greater influence than selection bias in the policy variable in these results.

Although I outlined a conceptual framework where I anticipated finding evidence of trade diversion in Paraguay, the empirical results did not support this claim in the aggregate or commodity level trade values. Quite the opposite, I found evidence of trade creation in all commodity categories. The only commodities where a decrease in imports from non-member countries is observed are *Mineral fuels and lubricants* and *Animal and vegetable oils & fat*, but these estimates are not statistically significant. The most disproportional increases in import levels from member countries were observed in *Beverages and Tobacco* and *Animal and vegetable oils & fat*. Considering the CET is in effect in these categories and is as high as 20% for some products (Mercosur-Secretariat), it is possible some trade diversion may have taken place. However, Paraguay has a production capacity in these categories and their import expansion attributable to Mercosur overwhelmingly outweighs the negligible losses attributable to trade diversion. Such developments are clearly beneficial. The expansion in imports of manufactured goods from Mercosur is less disproportionate and is possibly attenuated by the fact that many products in this sector are still exempt from the CET. Further research may illustrate how imports of non-exempt manufactured goods from the CET have been affected by Mercosur.

Looking forward, it is likely Paraguay's policy will be to continue supporting the benefits obtained in the agricultural sector, but to keep postponing the application of the CET on high technology manufactured goods.

2 Sensitivity Analysis of Gravity Model Estimations of Trade Creation and Trade Diversion: The Case of Paraguay and Mercosur

2.1 Introduction

In this chapter I test the sensitivity of part of the results of the first chapter by comparing them to a third estimator, the Pseudo-Maximum-Likelihood estimator, recommended in the literature to correctly address zero trade flows. Specifically, I compare traditional Ordinary Least Squares estimates to those resulting from the Heckman sample selection model and a Poisson Pseudo Maximum-Likelihood estimator. Given the availability of a commodity-level panel data set over a forty year period, country fixed effects are included in all techniques which control for country-level heterogeneity or multilateral resistance (Matyas, 1997; Egger, 2000; Anderson & van Wincoop, 2003). Moreover, I continue to employ the re-parameterization of the difference-in-difference estimator discussed in the first chapter to quantify the extent of trade creation and trade diversion of Paraguayan imports since Mercosur's creation. The sensitivity of these estimates is not only studied through different methodologies, but also through a reduction in sample size by gradually eliminating countries little to Paraguay.

A secondary objective of this study is to analyze the dynamic time path effects of Mercosur on Paraguayan import flows in an interval analysis. The literature is rich with suggestions that the formation of regional trade agreements between 'natural trading partners', usually neighboring countries trading disproportionately with each other, are likely to be welfare-improving (Wonnacott & Lutz, 1989; Krugman, 1991; and Summers,1991). Moreover, there are likely to be increases in regional trade around the time of the creation of an RTA as businesses position themselves to enter newly opened markets (Frankel, 1997; Magee, 2008). Additionally, the signing of an agreement does not imply immediate regional trade liberalization. There are usually lengthy transition periods from the moment an agreement is signed to its effective date. I analyze the time path of Paraguayan import flows in a 3-year interval analysis to test how levels of imports from Mercosur countries have evolved before and after its creation. Disproportionately high levels of regional imports occurring solely after Mercosur's adoption would be strongly attributable to this policy intervention.

The inclusion of measures for both, trade creation and trade diversion, in studying the impact of RTAs on an individual country in a dynamic time path environment is a contribution to this literature. My study is similar to Magee's (2008) to the extent I include a time path analysis of regional trade and similar to Fukao (2003) to the extent I focus on a single RTA. However, I depart from these studies by including a measure for trade diversion in the analysis of import flows.

2.2 Regional Economic and Political Shocks

In addition to the difficulties of capturing the effects of a trade agreement due to administrative transition periods, there may be any number of confounding factors. Far-reaching economic shocks such as the oil crises of 1973 and 1979, the Mexican 'Tequila' crises of 1994, and the Asian financial crises of 1997 are likely to have affected trade flows worldwide. Additionally, regional events are much more likely to have differential consequences. Table 6 lists several events which may have influenced Paraguayan import flows before, during, and since the implementation of Mercosur. Some of the regional events listed include the Falklands War, the construction of the hydroelectric Itaipú Dam, the beginning of a democratic government in Paraguay, and several regional financial crises.

Table 6 Regional and International Macroeconomic Shocks

Year	Event
1973	1st Oil Crises
1979	2nd Oil Crises
1982	Falklands War
1975-1982	Itaipú Dam construction. Operations begin in 1984.
1989	Argentine hyperinflation. Democratic government in Paraguay.
1990	Brazilian hyperinflation.
1991	Mercosur agreement is signed.
1991-2002	Argentine Currency Board.
1994	Brazil inflation and Mexican financial crises.
1994-1999	Brazilian Plan Real.
1995	Start of Mercosur policies and Paraguayan financial crisis.
1997	Asian financial crises.

Regional inflation and modifications of exchange rate policies have arguably been very influential on regional trade. The inflationary crises listed in Table 6 were the greatest in

Argentina, but were not much less severe in Brazil. Argentina experienced inflation rates between 672% and 3,079% from 1985 to 1989 while Brazil reached an inflation rate of 2,900% in 1990. That year, Paraguay and Uruguay also had reached peak inflation rates of 37% and 112% respectively (World Bank, 2011).

In an effort to stabilize their economies, Argentina and Brazil implemented fixed exchange rate policies named the *Plan Convertibilidad* (1991-2002) and *Plan Real* (1994-1999), respectively. The *Plan Convertibilidad* was a currency board where the local peso was fixed to the US dollar by law (Frankel, 1999), while the *Plan Real* was a more flexible peg to the U.S. dollar. Both of these regimes ended with currency devaluations, though in Argentina this devaluation was accompanied by a banking crisis. Paraguay also experienced a banking crises in 1995, the same year regional tariffs were planned to be completely phased-out in Mercosur. These banking crises consisted in numerous banks and financial institutions going under and affecting the savings of hundreds.

By making use of certain refinements discussed in the methods section, I attempt to cleanse the estimation results of the political and macroeconomic shocks discussed above.

2.3 Data

The data collected for this chapter include import values, US CPI, real GDP, real GDP per capita, and real exchange rates from 1970 to 2010. Additionally, I collected time invariant data such as bilateral distance between Paraguay and the exporter, language, land area, and religious similarities. Import values at the one digit level of disaggregation were retrieved from the UN Comtrade database. The selected commodity classification was the SITC Rev.1 which has data available for the most years, even before 1970. Import trade values were converted into constant values by deflating them using the US CPI. Geographic variables were collected from the CEPII (*Centre d'Etudes Prospectives et d'Informations Internationales*) and GDP related variables were collected from the USDA Economic Research Service.

During 1970-2010, there is sufficient data for variables to include 152 countries in a sample of countries exporting to Paraguay in at least a single year and single commodity[14]. The data set holds 13,209 positive valued observations, out of which less than 1% are less than $100. As explained in Frankel (1997) and mentioned earlier, missing observations in such trade data are attributable to the absence of trade between countries which may be due to the size of the trading economies or remoteness from each other. Considering Paraguay is a small landlocked country and the import values are at the one-digit level of detail, it is likely many countries do not export to Paraguay in all commodity codes in all years. In fact, no country in the data set exported to Paraguay in all commodities during all years. As part of this study, I enter these 'zero trade' observations to the data set which expands the number of observations in the data set to from 13,209 to 60,780[15].

2.4 Methodological Specifications

The models covered in this section describe the three estimators that will be used to compare measures of trade creation and trade diversion in Paraguay. The motivation behind this exercise is to analyze the consistency of the estimates across different estimators. These estimators are the pooled Ordinary Least Squares (OLS), the Heckman sample selection model, and the Poisson Pseudo Maximum Likelihood (PPML) estimator. In addition to country fixed effects, all models are estimated using robust standard errors in order to correct for serial correlation and heteroskedasticity in the error term (Wooldridge, 2002). I also discuss a sample size reduction exercise used in analyzing the sensitivity of the various estimators. The last section discusses an interval analysis aimed at gathering how trade with Mercosur countries has evolved over time.

---

[14] Table A.1 of the appendix lists the 155 countries used for the analysis in Chapter 1. Afghanistan, Rwanda, and Zimbabwe have been removed from the analysis in Chapter 2 due to lack of data on the real exchange rate, a variable included in the current analysis, reducing the list to 152 countries.

[15] Each country can have a maximum of 410 observations (10 commodity codes X 41 years). As an example, if a country had 200 observations of positive exports to Paraguay I added the remaining 210 zero-valued export flows. I have also estimated the models to be presented by including only commodities of each country for which trade occurred at least once. This implies if a country has never exported 'silver' to Paraguay, then silver is removed from the data set for that country. Doing this substantially reduces the number of observations in the data set, but the results are equivalent among 2 of the 3 estimators in the study (OLS & Poisson) and minimally different using the Heckman specification.

2.4.1 Sensitivity Analysis

As stated earlier a very large share of the data set is represented by zero trade flows or censored observations. These occur for mostly two reasons, it may be that trade between certain countries is rounded to zero and consequently goes unrecorded or there is genuinely no trade between two countries in a given year. It is very likely no trade takes place between small, developing, remote countries. Consequently, it is not surprising that in my data set there are very few occurrences of trade in any merchandise category with numerous countries[16]. Such observations can be up to 50% of the observations in studies involving multiple developed countries. In this initial data set, they make up 78% of the observations.

Alternatives to discarding the zero-valued observations include their substitution for a small constant such as $1000. This permits the OLS estimation of the log-linear gravity model but may cause additional difficulties since including a large number of censored values into the model specification may exacerbate heteroskedasticity in the error terms which produces biased and inefficient estimates. Another approach is to estimate the gravity equations using the Tobit estimator given trade values are bounded from below by zero, an approach followed by Soloaga & Winters (2001). However Linders & de Groot (2006) questioned the appropriateness of this method given the Tobit model requires censored data to reflect either 'desired negative trade' or the rounding to zero of very small values of trade, both of which do not occur. Still, some researchers who have followed any of these approaches claim their results are robust to the inclusion of the zero-valued observations or not (Frankel, 1997; Soloaga & Winters 2001), minimizing the need for any correction.

Recently, two techniques to address these problems have been recommended in the influential papers of Silva & Tenreyro (2006) and Helpman et al. (2008). Silva & Tenreyro (2006) propose the use of a Poisson Pseudo Maximum-Likelihood (PPML) estimator for gravity models while Helpman et al. (2008) treat the presence of zero trade flows as a sample selection problem and recommend a two-step estimation procedure similar to the Heckman selection model. Silva & Tenreyro (2006) argue the PPML estimator addresses two challenges of estimating gravity equations. First, if the gravity model takes a multiplicative exponential form and the error term is heteroskedastic, estimating such models using OLS after log-linearization leads to inconsistent estimates of the parameters of interest. The

[16] Burundi, Botswana, Namibia, and even Suriname which is in South America are examples of countries with which Paraguay has had an occurrence of trade in a single merchandise classification during the sample period.

underlying reason for this is the expected value of the logarithm of the error term depends on its mean and variance. If the variance of the error term in the multiplicative exponential form is a function of other regressors in the model, as may be the case under heteroskedasticity, the log of the error term will also depend on these regressors. This correlation violates the condition for consistency of OLS. Consequently, the authors recommend estimating such models in the multiplicative exponential form. Second, given this recommendation, zero-valued observations no longer pose a challenge since the gravity equation is estimated in levels rather than log form. Additionally, the authors assessed the empirical performance of PPML using Monte Carlo simulations. They found that in the presence of heteroskedasticity, estimates based on log transformation procedures such as OLS and Tobit, obtained biased estimates while those based on the PPML method were robust to different patterns of heteroskedasticity[17].

The second technique proposes estimating the gravity model with a correction for sample selection bias (Helpman et al., 2008; Linders & de Groot, 2006. Helpman et al. (2008) develop a theory predicting positive as well as zero trade flows between countries, which their two-stage estimation procedure is based on. In the first stage they model the probability of trade occurring while the second stage reflects the gravity equation with a correction for sample selection as well as firm heterogeneity, making it similar to a Heckman sample selection model. This procedure will be discussed in more detail later, but it hinges on finding a suitable 'exclusion restriction', which may be difficult to find[18]. Helpman et al. (2008) successfully use an index for 'religion' as their exclusion restriction. However, such a variable is differenced out of models with fixed effects which is why the Heckman model in this study does not have such an exclusion restriction. As the reader will see, it is possible to do so but not ideal, reflecting a disadvantage of the Heckman sample selection correction over the PPML estimator in this context.

I estimate three models, a traditional OLS model on the truncated data set, as well as the Heckman and PPML models to evaluate the consistency of my findings regarding trade creation and trade diversion. In what follows, I state the specifications used in each case.

[17] The Tobit analyzed in this study is the Eaton and Tamura (1994) Tobit model, also referred to in Martin (2008).

[18] The exclusion restriction requires a variable to be correlated with inclusion in the sample (dependent variable in the selection equation), but uncorrelated with the error term in the outcome equation.

2.4.1.1 Pooled OLS with Country Fixed Effects

The benchmark specification in this work is a pooled cross-sectional Ordinary Least Squares (OLS) model with country fixed effects taking the following form:

$$\ln(I_{kjt}) = \beta_0 + \beta_1 t + \beta_2 \ln\big(GDP_{Pyt} \cdot GDP_{jt}\big) + \beta_3 \ln\left[\left(\frac{GDP}{pop}\right)_{Pyt}\left(\frac{GDP}{pop}\right)_{jt}\right] + \beta_4 \ln\big[RER_{Pyt} \cdot RER_{jt}\big] + \gamma_3\big(PM_{jt}\big) + \gamma_4\big(PNM_{jt}\big) + C_1 + \cdots + C_N + \varepsilon_{kjt} \quad (11)$$

$$for\ j = 1, \dots, N$$

Recall the data are at the one digit level of commodity disaggregation so the subscript *k* indexes commodities[19]. Considering only Paraguayan imports are analyzed, the subscript *j* indexes exporting countries and *Py* reflects Paraguayan values. Parameters β_2 and β_3 reflect traditional gravity variables used to predict 'natural' flows of trade from the rest of the world to Paraguay. The product of the real exchange rates is also included in this specification considering the time dimension is an essential component of this data set (Bayoumi & Eichengreen, 1997; Winters & Soloaga, 2001). It is defined as each country's real exchange rate versus the US dollar, so an increase in its value reflects a real depreciation of the local purchasing power. The last two parameters in (8), γ_3 and γ_4, reflect the re-parameterization of the difference-in-difference estimator used to capture measures of trade creation and trade diversion[20]. As mentioned previously, I use 1995 as the start year for Mercosur considering it was the year regional trade was originally planned to be liberated, as well as when convergence to the common external tariff began.

The specifications in this study include a linear time trend intended to capture aggregate factors such as global inflation and economic growth. This is a departure from the usual practice in the literature which is to include year fixed effects. Year fixed effects are omitted because of an almost perfect linear combination with the trade diversion parameter, γ_4. It is

---

[19] There are ten commodity codes in the data set are: S1-0 Food and live animals; S1-1 Beverages and tobacco; S1-2 Crude materials, inedible, except fuels; S1-3 Mineral fuels, lubricants and related materials; S1-4 Animal and vegetable oils and fats; S1-5 Chemicals; S1-6 Manufactured goods classified chiefly by material; S1-7 Machinery and transport equipment; S1-8 Miscellaneous manufactured articles; S1-9 Commodities & transactions not classified according to kind.

[20] Recall M_j included in the difference-in-difference specification is time invariant and dropped with fixed effects.

not possible to include year dummy variables because after 1995, they form a linear combination with the sum of dummy variables PM_{jt} and PNM_{jt}.

2.4.1.2 Heckman Sample Selection Model

I also implement a Heckman sample selection model with accounts for the zero-valued trade observations discarded using traditional OLS methods. This is a two-step procedure where in the first stage or selection equation the existence of a trading relationship is modeled using a probit estimator. Hence, the positive-valued trade observations take a value of one and the remaining 'missing' observations zero. The predicted values of this model are used to construct the inverse Mills ratio, which is augmented to the outcome equation in the second stage, the gravity model in this application. The gravity equation is estimated on the uncensored observations only. As in the benchmark, this model is estimated using country fixed effects.

This model takes the following form:

Selection Equation:

$$T_{kjt} = \beta_0 + \beta_1 t + \beta_2 \log(GDP_{Pyt} \cdot GDP_{jt}) + \beta_3 \log\left[\left(\frac{GDP}{pop}\right)_{Pyt}\left(\frac{GDP}{pop}\right)_{jt}\right] +$$

$$\beta_4 \ln[RER_{Pyt} \cdot RER_{jt}] + \gamma_3(PM_{jt}) + \gamma_4(PNM_{jt}) + C_j + \cdots + C_N \ + u_{kjt}$$

$$\text{for } \ j = 1, \dots, N$$

$$T_j = \begin{cases} 1 & Trade\ Volume > 0 \\ 0 & Trade\ Volume = 0 \end{cases} \tag{12}$$

Outcome Equation:

$$\ln(I_{kjt}) = \beta_0 + \beta_1 t + \beta_2 \log(GDP_{Pyt} \cdot GDP_{jt}) + \beta_3 \log\left[\left(\frac{GDP}{pop}\right)_{Pyt}\left(\frac{GDP}{pop}\right)_{jt}\right] +$$

$$\beta_4 \ln[RER_{Pyt} \cdot RER_{jt}] + \ \gamma_3(PM_{jt}) + \gamma_4(PNM_{jt}) + \lambda(\hat{T}_{kjt}) + C_j + \cdots +$$

$$C_N + \theta_{kjt} \qquad \text{for } j = 1, \dots, N$$

This approach has several advantages. First, it adds flexibility to the comparable Tobit model, which restricts the censoring mechanism to be part of the outcome equation. The two-step procedure models the processes separately. Using an OLS estimator in the second stage rather than maximum-likelihood methods permits the adoption of weaker distributional assumptions of the joint distribution of the error terms of both models (Cameron and Travedi, 2005). A priori, any potential correlation between the error terms of the first and second stages can be verified and corrected by this procedure. A lack of correlation between the error terms Cov($\varepsilon_{kjt}, u_{kjt}$)=0, is indicative of no sample selection bias. Note ε_{kjt} refers to the error term of the model prior to the implementation of the correction procedure.

This specification usually requires the existence of a suitable 'exclusion restriction'. The exclusion restriction requires a variable to be correlated with inclusion in the sample (dependent variable in the selection equation), but uncorrelated with the error term in the outcome equation. Helpman et al. (2008) constructed a 'religion' index increasing in similarity between trading partners. However, this is a time invariant index and not suitable for a specification with country fixed effects. As a result, this model does not include an exclusion restriction. This is possible, but it is likely to introduce multicollinearity into the outcome equation. However, in such exercises multicollinearity is attenuated by the extent of variation in the explanatory variables of the selection equation, that is, the better the probit model distinguishes between trading and non-trading partners (p. 551 Cameron & Travedi, 2005; p. 564 Wooldridge, 2002). Although the lack of a suitable exclusion restriction is a weakness of this model, I propose multicollinearity does not pose a threat to the identification of this model considering the large number of countries and years in the data set.

2.4.1.3 Poisson Pseudo-Maximum Likelihood Estimator

As discussed above, estimating the linearized form of the gravity model (by taking the log of both sides) will lead to inconsistent estimates when the variance of the error term in the untransformed gravity model depends on the explanatory variables. Silva & Tenreyro (2006) first derive a general form of a multiplicative exponential form of the gravity model and then propose to use the *Poisson Pseudo Maximum Likelihood* (PPML) estimator. Although PPML is usually used for count data, the authors argue that it is numerically equivalent to the Non-linear Least Squares (NLS) estimator that could be used to estimate the multiplicative exponential model[21]. Finally, since import flows (the dependent variable) are in levels, zero-trade flows do not pose a challenge in this estimation procedure.

Silva & Tenreyro (2006) study this estimator on cross-sectional data. However, this estimator has been implemented in panel data environments[22], which is what I estimate. For this application, I implement the above estimator using the following specification:

$$I_{kjt} = \exp\left\{\beta_0 + \beta_1 t + \beta_2 \ln\left(GDP_{Pyt} \cdot GDP_{jt}\right) + \beta_3 \ln\left[\left(\frac{GDP}{pop}\right)_{Pyt} \left(\frac{GDP}{pop}\right)_{jt}\right] + \beta_4 \ln\left[RER_{Pyt} \cdot RER_{jt}\right] + \gamma_3 PM_{jt} + \gamma_4 PNM_{jt} + C_1 + \ldots + C_N\right\} + \varepsilon_{kjt} \tag{13}$$

$$for\ j = 1, \ldots, N,$$

which is an application of the general from,

$$I_{kjt} = \exp\left(\boldsymbol{x}'_{\boldsymbol{jt}}\hat{\beta}\right) + \varepsilon_{kjt}\,. \tag{14}$$

Then from equation (14), minimized squared errors lead to the following set of K *m* first order conditions for panel data:

$$\sum_{j=1}^{N}\sum_{t=1}^{T}\left[I_{kjt} - \exp\left(\boldsymbol{x}'_{\boldsymbol{jt}}\hat{\beta}\right)\right]\exp\left(\boldsymbol{x}'_{\boldsymbol{jt}}\hat{\beta}\right)\boldsymbol{x}'_{\boldsymbol{jmt}}\hat{\beta} = 0 \tag{15}$$

m= 1,…,M. M is the number of explanatory variables.

---

[21] This refers directly to equations (6) and (7) in Silva & Tenreyro (2006), where (6) is the general form of the multiplicative exponential model,

$$y_i = \exp(x_i\beta) + \varepsilon_i$$

which can also be expressed as,

$$y_i = \exp(x_i\beta)\,\eta_i.$$

This equation is transformed by taking the logarithm of both sides, leading to equation (7) in Silva & Tenreyro (2006)

$$\ln y_i = x_i\beta + \ln \eta_i.$$

[22] It is noteworthy, that while Silva & Tenreyro (2006) recommend estimating the gravity model in levels, Magee (2008) took the log of the dependent variable using the fixed effect PPML estimator.

2.4.1.4 Event fixed effects

To improve the accuracy of the estimates, I run an additional set of regressions including dummy variables for major political and macroeconomic shocks which are unlikely to be adequately accounted for by the linear time trend variable. The selected events, which have been mentioned in the section 2.2 are listed in Table 7.

Table 7 Event fixed effects included in the model specifications.

Year	Event
1973	1st Oil Crises
1979	2nd Oil Crises
1982	Falklands War
1984	Operations of Itaipú dam begin in 1984.
1989	Argentine hyperinflation. Democratic government in Paraguay.
1990	Brazilian hyperinflation.
1991	Mercosur agreement is signed.
1994	Brazilian inflation and Mexican financial crises.
1995	Start of Mercosur policies and Paraguayan financial crisis.
1997	Asian financial crises.
1999	Brazilian recession and end of *Plan Real*.
2002	Argentine financial crises and end of *Plan Convertibilidad*.

2.4.2 Sensitivity to sample size

Estimating the PPML and Heckman sample selection models in-full, implies adding 47,571 zero-valued observations to the data set which represent 78% of all observations. Considering the disproportionate number of zero trade observations in the full zero-imputed data set, I gradually remove countries with limited number of trade occurrences with Paraguay during the sample period, with the objective of reducing the censored share of observations to less than 20%[23]. This exercise is used to analyze the sensitivity of the gravity variables, the trade creation, and the trade diversion parameters across the three specifications as the number of countries included in the data set is reduced.

Table 8 describes the six sample sizes that were used to estimate the three specifications in this work. Each country may have a total of 410 (10 merchandise categories x 41 years) frequencies of exports to Paraguay. I gradually remove countries with few export frequencies in order to reduce the share of censored observations and to analyze the

[23] Cameron & Travedi (2005) list seven studies with the share of zeros in the data set range from 20 – 90% (p. 666).

estimates when only major trading partners are included. The first row in Table 8 describes the full data set representing 100% of Paraguayan trade with all 152 countries in the sample from 1970-2010. The second row describes the sample size when 66 countries remain in the sample. Notice the sample size is reduced to 26,860 export flows while the share of censored observations decreases to 54%. Still, these 66 countries represent 99.6% of trade relative to the full data set. Similarly, when the sample size is reduced to include 44 countries, these represent 97.4% of trade relative to the full data set.

Table 8 Sample Sizes uses to estimate the Poisson, Heckman, and OLS models.

N	Frequencies	Coun-tries	% of Total Trade	Positive-valued observations	Censored share
60,780	Full	152	100%	13,209	78%
26,860	<41	66	99.6%	12,403	54%
18,040	<124	44	97.4%	10,850	40%
16,400	<138	40	80.8%	10,327	37%
11,070	<200	27	77.6%	8,194	26%
7,790	<250	19	74.9%	6,395	18%

There is a noticeable decrease in the share of total trade in the sample when the sample size is reduced to 16,400 observations. Notice the percentage of total trade decreases from 97.35% to 80.0%, which responds to the removal of 'China' from the data set[24]. When *N* equals 7,790 the share of censored observations is reduced to 18%, though the 19 countries in this sample continue to represent about 75% of all imports from all countries during the given time period.

2.4.3 Interval Analysis

Finally, I introduce an interval analysis of Mercosur imports. The objective of this exercise is to identify whether there is a noticeable 'a priori' influence of Mercosur as often described in the literature. This exercise may also illustrate whether regional trade has been disproportionately high, or whether recent increases in regional trade may be genuinely attributable to the creation of Mercosur. I use the full data set to estimate the three models with adjustments to their specification for the interval analysis.

[24] The treatment of China in this data set may be challenging, considering Paraguay holds diplomatic relations with Taiwan rather than China, while the UN Comtrade database recognizes the People's Republic of China rather than Taiwan.

The standard models of the previous section compare levels of import flows from member and non-member countries before and after 1995, implying the base years for this comparison are 1970-1994. To capture 'a priori' effects of the agreement the base years are reduced to 1970-1979 and three year dummy variables for membership to the agreement are included from 1980 and on[25]. Hence, while the previous set of models capture average changes in the levels of trade at a single point in time, the interval analysis captures changes in the level of trade at ten distinct three-year intervals.[26] The 10 year period from 1970-79 is considered a sufficiently long and representative period to compare more recent years against. A shorter (longer) period would make finding a significant difference relative to that period easier (more difficult) because import flows in the early 1970s may have been disproportionately low compared to current levels (because import flows into the late 1980s may be less distinguishable from current levels).

A general specification holding the dynamic time path of trade creation is proposed in (16):

$$
\begin{aligned}
Imports_{kjt} = {} & \beta_0 + \beta_1 t + \beta_2 \ln\left(GDP_{Pyt} \cdot GDP_{jt}\right) + \beta_3 \ln\left[\left(\frac{GDP}{pop}\right)_{Pyt}\left(\frac{GDP}{pop}\right)_{jt}\right] \\
& + \beta_4 \ln\left[RER_{Pyt} \cdot RER_{jt}\right] + \delta_1\left(PM_{j8082}\right) + \delta_2\left(PM_{j8385}\right) \\
& + \delta_3\left(PM_{j8688}\right) + \delta_4\left(PM_{j8991}\right) + \delta_5\left(PM_{j9294}\right) + \delta_6\left(PM_{j9597}\right) \\
& + \delta_7\left(PM_{j9800}\right) + \delta_8\left(PM_{j0103}\right) + \delta_9\left(PM_{j0406}\right) + \delta_{10}\left(PM_{j0710}\right) \\
& + \gamma_4\left(PNM_{jt}\right) + C_1 + \cdots + C_N + \varepsilon_{kjt} \\
& \qquad for\ j = 1, \ldots, N
\end{aligned}
\tag{16}
$$

The measure of each of the δ coefficients reflects the percent change in imports from Mercosur countries over the years it is indexed in comparison to imports levels during the base years. Hence, if there are significant 'a priori' increases in regional trade before the creation of Mercosur, one would expect δ_{8991} (1989-1991) or δ_{9294} (1992-1994) to be positive and statistically significant. The interpretation of γ_4 remains unchanged, it measures average percent change in levels of imports from non-member countries starting in 1995.

---

[25] Three year intervals collapse three years of data into a single estimate.

[26] The dummy variable for the last interval is actually a four year interval due to the number of years in the exercise (2007-2010).

2.5 Results

2.5.1 Models with year fixed effects and complete sample.

The first set of results reported in Table 9 compares the output of the three main models (11), (12), and (13), using the full data set. This data set includes a total of 152 countries holding 62,780 observations. The estimates across the three models are consistent to the extent all coefficients have the same signs, but the statistical significance of these varies across the models. In the OLS model, all coefficients are statistically significant except the coefficient on the *Non-member* dummy variable. As the gravity model predicts, imports increase in the product of country GDPs by a relatively large magnitude, 2.6% for every 1% increase in the product of country GDPs.

Table 9 Estimation Results with country fixed effects and full data set.

	Pooled OLS		Heckman two-step				Poisson	
			Selection		Outcome			
	β	p-val.	β	p-val.	B	p-val.	B	p-val.
Time (β_1)	-0.10**	0.00	-0.02	0.50	-0.120**	0.00	-0.05	0.23
GDP (β_2)	2.60**	0.00	0.37**	0.00	3.35**	0.00	1.42	0.30
GDPperCap.(β_3)	-1.43**	0.01	-0.10	0.16	-1.85**	0.00	0.99	0.53
RER (β_4)	-0.16**	0.04	-0.01	0.44	-0.17**	0.00	-0.31*	0.07
PM, *Member* (γ_3)	1.49**	0.00	0.11	0.39	1.34**	0.00	0.82**	0.00
PNM, *Non-member* (γ_4)	0.15	0.12	0.28**	0.00	0.42**	0.00	-0.16	0.19
Constant	-86.9**	0.00	-17.53**	0.00	-115.4**	0.00		
Inverse Mills (λ)					1.70**	0.00		
Fixed Effects	Yes		Yes		Yes		Yes	
Countries	152		152		152		152	
Observations	13,209		62,780		13,209		62,780	

Note: *, **, *** imply significance at the 10%, 5%, and 1% levels, respectively.

The negative signs on *GDP per capita* have been discussed in the first chapter of this thesis. They are consistent with the Helpman-Krugman theory of trade considering the negative signs on *GDP per capita* imply Paraguay imports less from countries with high capital endowment ratios. While foreign GDPs per capita are driving the negative sign on this variable, Paraguay's own real exchange rate is driving the negative sign on the *RER* variable. Since all *RERs* in the sample are calculated in relation to the U.S. dollar, a real depreciation of Paraguay's currency makes imports more expensive, obtaining the negative, statistically significant negative sign on *RER*. The last two estimates in the pooled OLS model are the 'difference' estimates measuring average percent change with Mercosur and non-Mercosur countries since the policy intervention year, 1995. Hence, the 1.488 estimate on *Member* reflects an average expansion in imports of 322% $[(e^{1.44} - 1) * 100]$.

The models incorporating the zero trade flows, the Heckman and Poisson models, do not contradict these results, but somewhat attenuate them. While the magnitude of the GDP variables is slightly greater in the Heckman model, the coefficient capturing 'trade creation' is slightly smaller than the OLS model. Specifically, by the Heckman model the measure of trade expansion is 283% and also obtains a statistically significant expansion of imports from *Non-members* of 51%. Notice this result indicates imports from non-Mercosur countries has increased on average, indicating there is no evidence of trade diversion in Paraguay due to Mercosur. The coefficient on the Inverse Mills (λ) is positive and statistically significant at the 95% confidence level, which is evidence of sample selection bias. Finally, the results obtained in the Poisson model depart from the previous models in that the only statistically significant coefficient is the measure of trade creation, though much lower than previous models (126%). This reduction in the effect of regional trade agreements is consistent with the estimation of results of Silva & Tenreyo (2006), who also found a substantial decrease on the trade agreement dummy variable of the Poisson specification in comparison to their benchmark OLS models. The *RER* coefficient which is negative and statistically significant at the 95% confidence level in both the OLS and Heckman models is statistically significant at the 90% confidence level in the Poisson model.

2.5.2 Models with country and 'event' fixed effect using the full data set.

In this section I discuss the results incorporating the dummy variable for the 'event fixed effects' explained in the methods sections. These results are reported in Table 10.

Four of the years included in the specification are statistically significant using the three estimators: 1979, 1989, 1994, and 2002. The first is the year the second oil crises occurred. In 1989 the military dictatorship ended in Paraguay, and even though this was a political event, it is likely it had economic ramifications. Several different economic events took place in 1994. First, Brazilian inflation levels reached a temporary high of 2,075%. At the same time, Brazil implemented a currency stabilization policy called the *Plan Real* resulting in a currency appreciation and stabilization which lasted until 1999. Additionally, the Mexican currency devaluation crisis, popularly referred to as the 'Tequila Crises', also took place in

1994. It is difficult to disentangle whether the 1994 dummy variable is capturing Brazil's inflation or currency appreciation, since the latter took effect in July 1994. Whichever it is, both events are likely to have affected Paraguayan imports from Brazil in the same direction. However, it is very unlikely it is capturing the 'Tequila Crises' since this occurred in December 1994[27]. Finally, 2002 is the year the *Plan Convertibilidad* ended. Notice the negative sign on these coefficients imply a contraction of Paraguayan import flows that year.

Although several of the selected 'event' fixed effects obtain statistical significant results, the estimates on the remaining variables seem to be insensitive to the inclusion of these dummy variables since there is little variation in their magnitudes. The coefficients of *GDP* and *GDP per capita* are quite insensitive to the inclusion of event fixed events in the OLS and the Heckman models. For example, the magnitude of the estimate of *GDP* decreases from 2.605 to 2.420 while that for *GDP per capita* decreases from -1.43 to -1.32 in the OLS model. The magnitudes of *GDP* and *GDP per capita* are slightly greater in both, with and without event fixed effects, in the Heckman specification. However, within the Heckman models *GDP* and *GDP per capita* have a slightly lower magnitude after including events fixed effects. *GDP* decreases from 3.345 to 3.150 and the magnitude of *GDP per capita* falls from -1.853 to -1.709.

Finally, the Poisson model obtains the lowest set of coefficients for these variables, though neither *GDP* nor *GDP per capita* are statistically significant in either specification, with and without event fixed effects. It is possible this is due large share of zero-valued observations imputed in the data set, which is a reason we analyze these estimates on smaller sample sizes in the next section.

The *RER* variable endures more noticeable changes after including event fixed effects in some of the models. In the OLS model it slightly decreases in magnitude (from -0.163 to -0.117), and loses statistical significance (from 0.05 to 0.10). Its magnitude also decreases slightly in the Heckman model, from -0.176 to -0.128, though there is no change in statistical significance (0.01). However, in the Poisson model the *RER* variable is no longer statistically significant in the model with event fixed effects. This result is not completely surprising considering a lot of the information held in the year dummy variables is related to

[27] Although it may seem unlikely that a currency crisis in Mexico may an impact on Paraguayan import flows, the far reaching effects of this crisis in Brazil and especially Argentina are well-documented. An excellent narrative of these events is provided in "The Return of Depression Economics" by Paul Krugman (1999).

price changes and exchange rate volatility, giving the real exchange rate variable less to control for.

Table 10 Estimation Results with country fixed effects and full data set.

	Pooled OLS		Heckman two-step				Poisson	
			Selection		Outcome			
	β	p-val.	β	p-val.	β	p-val.	β	p-val.
Time (β_1)	-0.093**	0.00	-0.009**	0.01	-0.117**	0.00	-0.052	0.24
GDP (β_2)	2.420**	0.00	0.427**	0.00	3.150**	0.00	1.281	0.37
GDP per Cap. (β_3)	-1.318**	0.02	-0.104	0.13	-1.709**	0.00	1.175	0.47
RER (β_4)	-0.117*	0.09	-0.006	0.57	-0.128**	0.00	-0.250	0.16
PM, *Member* (γ_3)	1.419**	0.00	0.261**	0.05	1.407**	0.00	0.947**	0.00
PNM, *Non-member* (γ_4)	0.121	0.27	0.411**	0.00	0.467**	0.00	-0.027	0.83
1973	-0.021	0.81	0.103*	0.08	0.078	0.64	0.469**	0.02
1979	0.206*	0.05	0.084	0.12	0.260*	0.09	0.100**	0.03
1982	0.041	0.69	-0.029	0.60	0.003	0.99	-0.075	0.39
1984	-0.497**	0.00	-0.234**	0.00	-0.680**	0.00	-0.105	0.45
1989	-0.361**	0.00	-0.040	0.47	-0.399**	0.01	-0.268**	0.01
1991	0.248**	0.02	0.036	0.51	0.258*	0.08	0.153	0.14
1994	0.542**	0.00	0.185**	0.00	0.662**	0.00	0.399**	0.00
1995	0.000*	0.00	0.000**	0.00	0.000	0.25	0.000	0.68
1997	0.346*	0.00	-0.152**	0.00	0.198	0.15	0.007	0.95
1999	0.072	0.48	-0.267**	0.00	-0.153	0.29	-0.402**	0.00
2002	-0.247**	0.02	-0.149**	0.00	-0.362**	0.01	-0.221**	0.02
Constant	-79.99**	0.00	-20.05**	0.00	-20.05**	0.00		
Inverse Mills (λ)					1.548**	0.00		
Fixed Effects	Yes		Yes		Yes		Yes	
Countries	152		152		152		152	
Observations	13,209		60,780		13,209		60,780	
RESET test p-values	0.058		0.000		0.000		0.519	

Note: *, **, *** imply significance at the 10%, 5%, and 1% levels, respectively.

As a check of the functional form of the reported specifications I ran a heteroskedasticity-robust RESET test (Ramsey, 1969). This test assumes there is no functional form misspecification in the null hypothesis and is realized including the quadratic predicted value, $(\hat{I}_{kjt})^2$ as an additional explanatory variable and analyzing its statistical significance. The p-values for this test are reported in the line of Table 10. Notice this parameter is statistically significant at the 90% confidence level for both the OLS or Heckman specifications, providing evidence of functional form misspecification, that is, the linear specification is inappropriate. However, the null hypothesis of no functional form misspecification cannot be rejected in the case of the Poisson estimator.

Regarding our main parameters of interest capturing trade creation and trade diversion, I find there is no qualitative difference across the estimators or the inclusion of event fixed effects. Table 11 below summarizes the estimates for trade creation and trade diversion. The OLS estimates for trade creation are the most insensitive to the inclusion of event fixed effects. The OLS estimate is 1.488 in Table 9 and 1.419 in Table 10. Consequently, their exponentiated values reported in Table 11 are 343% and 313%, respectively. The corresponding estimates of the Heckman model are slightly attenuated for selection bias but the signs and statistical significance are equivalent. Finally, the estimates for trade creation in the Poisson model are also statistically significant but much lower in magnitude than even the Heckman specification. By the Poisson model, Paraguayan imports from Mercosur countries have only increased 126% on average since 1995. The measure increases to 158% after accounting for event fixed effects. Finally, only the trade diversion estimates of the Heckman specification estimates, captured in PNM, *Non-member* (γ_4), are statistically significant. Since these measures are positive, there is no evidence of trade diversion.

Table 11 Measures of Trade Creation & Trade Diversion by Estimator and Inclusion of Event Fixed Effects.

	Pooled OLS	**Heckman two-step**	**Poisson**
PM, *Member* (γ_3)			
No Event FE	343%***	283%***	126%***
Event FE	313%***	308%***	158%***
PNM, *Non-member* (γ_4)			
No Event FE	17%	51%***	-15%
Event FE	13%	60%***	-3%

Notes: FE abbreviates effects. All values calculated from estimates by $[(e^{\gamma_i} - 1) * 100]$ and '***' imply significance of the estimate at the 1% level.

The reduction of the trade creation in the Poisson model to about half the measure of the same estimate in the OLS model is consistent with the work of Silva & Tenreyro (2006). Though their results were based on cross-sectional data, they found their dummy variable for preferential trade agreement in the Poisson model to be about one third of the magnitude of the estimate in the OLS model. A commentary on this may be that the large error terms associated to Mercosur countries are disproportionately weighting these observations, leading to large coefficient estimates of the membership dummy variable. On the contrary, the Poisson estimator weights all observations equally and consequently leads to a smaller measure of trade creation. Regardless, Silva & Tenreyro (2006) attribute this difference in magnitudes to biases associated to the inaccuracies of estimating gravity equations using log-linear forms.

2.5.3 Sensitivity to sample size

An additional exercise involves analyzing the estimates after reducing the number of countries in the sample until only Paraguay's major exporting partners are included in the data set. This exercise was performed by gradually removing countries with very few export flows to Paraguay. Recall, if each country can potential have 410 export flows to Paraguay (41 years x 10 commodity classifications) countries with a single export observation over the given time period imply they have 409 zero export flows in the data set. The specifications were estimated using several sample sizes, the smallest of which included 19 of the 152 exporting partners. These 19 countries reflect those that have at least 250 export observations to Paraguay representing 75% of all Paraguayan imports over the given period. This list of 19 countries is reported in Table A.2 of the Appendix.

There are two objectives in reducing the sample of countries in the data set. First, it will test the sensitivity of the trade creation and trade diversion estimates to differing number of countries used for the specifications. Second, it will also illustrate how the significance of the remaining gravity variables changes as the ratio of zero Paraguayan import flows is reduced in the data set. For example, the share of zero import flows in the smallest data set is 18%, much lower than the original 78%.

Table 12 reports the estimation results using the smallest data set previously discussed. Specifically, the sample holds 19 countries, 7,790 observations, 18% of which are zero Paraguayan import flows. Although some of the magnitudes differ substantially from the estimates based on the original sample size, the new coefficients are consistent in signs. The inverse Mills ratio continues to be statistically significant and therefore continues to support evidence of sample selection bias. Additionally, the number of statistically significant coefficients of the Poisson model almost doubles when the sample is reduced to the 19 major trading partners.

Table 12 Estimation Results with country fixed effects using the reduced data set.

	Pooled OLS		Heckman two-step				Poisson	
			Selection		Outcome			
	β	p-val.	β	p-val.	β	p-val.	β	p-val.
Time (β_1)	-0.16**	0.00	-0.06**	0.00	-0.08**	0.01	-0.05*	0.07
GDP (β_2)	5.11**	0.00	1.32**	0.00	3.77**	0.00	2.50**	0.00
GDP per Cap. (β_3)	-4.23**	0.00	-0.73*	0.04	-3.78**	0.00	-1.56*	0.08
RER (β_4)	-0.23	0.29	0.01	0.94	-0.28**	0.05	-0.37**	0.01
PM, *Member* (γ_3)	0.93**	0.00	0.50**	0.00	0.58*	0.06	0.55**	0.00
PNM, *Non-member* (γ_4)	0.09	0.47	0.32**	0.00	-0.30	0.22	-0.14*	0.10
1973	-0.08	0.44	-0.00	0.99	-0.11	0.67	0.15*	0.07
1979	0.34**	0.01	-0.18	0.11	0.56**	0.04	0.17**	0.00
1982	0.13	0.41	-0.14	0.22	0.31	0.24	0.05	0.56
1984	-0.32**	0.02	-0.13	0.24	-0.15	0.55	-0.02	0.91
1989	-0.32*	0.07	0.16	0.22	-0.45*	0.09	-0.19**	0.00
1991	0.20	0.16	-0.02	0.84	0.22	0.39	0.12	0.30
1994	0.52**	0.01	0.32**	0.02	0.20	0.48	0.35**	0.00
1995	0.00**	0.03	0.00	0.90	0.00	0.10	0.00	0.19
1997	0.24	0.11	-0.07	0.60	0.34	0.18	0.19**	0.04
1999	-0.18	0.28	-0.18	0.15	0.01	0.98	-0.30**	0.00
2002	-0.5**	0.01	0.02	0.85	-0.50*	0.05	-0.29**	0.01
Constant	-162.0**	0.00	-48.82**	0.00	-101.9**	0.00		
Inverse Mills (λ)					-3.151**	0.01		
Fixed Effects	Yes		Yes		Yes		Yes	
Countries	19		19		19		19	
Observations	6,395		7,790		6,395		7,790	
RESET test p-values+	0.263		0.502		0.172		0.245	

Note: *, **, *** imply significance at the 10%, 5%, and 1% levels, respectively. + The RESET test p-value indicates none of the models are incorrectly specified when in estimating the models on the reduced data of 19 major trading partners.

The magnitude of the gravity coefficients in the OLS model, *GDP* and *GDP per capita*, are much greater in the reduced data set, increasing from 2.4 to 5.1 and -1.3 to -4.2, respectively. This is expected considering most of the countries in the reduced sample are industrialized country with relatively higher *GDPs* and *GDPs per capita* than the countries in the full sample. That is, the higher scale of these variables is influencing the magnitudes of these coefficients.

The pattern of greater magnitudes in the *GDP* and *GDP per capita* coefficients continues in the outcome equation of the Heckman model. Finally, while the *GDP* and *GDP per capita* coefficients were not statistically significant in the Poisson using the full data set, both are statistically significant using the reduced data set. Once again the Poisson estimates are lower than those of the corresponding OLS and Heckman estimates. Interestingly, the coefficient on *RER* is not statistically significant in the OLS model, but it remains negative and statistically significant in both the Heckman and the Poisson models. The relatively low magnitudes of the *RER* coefficients in both the Heckman and Poisson models, -0.28 and -0.37, respectively, suggest import volumes are relatively insensitive to currency depreciations.

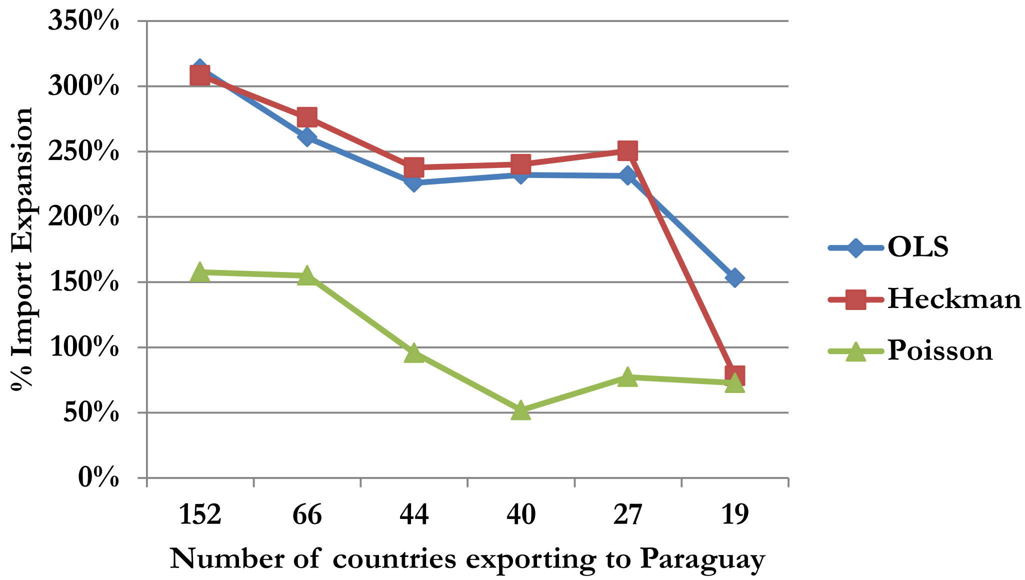
Finally, the parameters of interest measuring trade creation and trade diversion also remain consistent across the three models of the reduced sample. As before, the order of magnitudes are maintained with the estimate for Member (γ_3) being the greatest in the OLS model, followed by the Heckman and the Poisson estimates at 0.928, 0.576, and 0.547, respectively. These values are considerably lower than those estimated using the full sample (1.419, 1.407, 0.947), which illustrates that regional imports are disproportionate even when compared to only major international exporters to Paraguay. These estimates imply imports from Mercosur countries have increased 153%, 78%, and 73% since 1995 for each of the specifications, respectively.

Regarding the measure for trade diversion, *Non-member* (γ_4), I find no qualitative difference between the results based on the full and reduced data sets using the OLS and Heckman specifications. This is because the trade diversion estimates are not statistically significant in these specifications. However, I obtain different results in the Poisson specification. While the trade diversion estimate is not statistically significant in the full data set, it is negative and statistically significant at the 90% confidence level in the reduced data set. The trade diversion estimate in the Poisson specification is -0.143 and its interpretation is that the average level of imports from non-member countries has decreased by 13.36% since 1995. This figure represents slight evidence of trade diversion, but not enough to claim Mercosur has been mostly trade diverting for Paraguay. Recall the measure of trade creation in the Poisson specification, 73%, outweighs the measure of trade diversion.

To further illustrate the variability of the different estimators for trade creation across different sample sizes, I summarize the measures of trade creation in Figure 4 below. This figure reports the measures of trade creation using multiple sample sizes beginning with 152 countries and gradually decreasing these to 66, 44, 40, 27, and finally 19 countries[28]. The corresponding shares of censored observation are 78%, 54%, 40%, 37%, 26%, and 18%, respectively.

The OLS estimates for trade creation across the various sample sizes decrease from 313% to 153%. Those of the Heckman specification are very similar, starting at 308% but decreasing to 78% in the data set with the fewest number of countries. Finally, the Poisson model estimates begin with 158% and end with 73%. All these estimates are statistically significant at the 95% confidence level with the exception of the estimate of the Heckman model with 19 major trading partners (it is significant at the 90% confidence level).

Figure 4 Measure of Trade Creation in Paraguay attributable to Mercosur by number of exporting countries, 1970-2010.



For all sample sizes, the Poisson estimate is consistently below those of the OLS and Heckman methodologies demonstrating to be the most insensitive of the estimators. As mentioned previously, the lower magnitudes of the Poisson estimates are consistent with the

[28] Regression results using 153 and 19 countries are reported in Table 10 and 12, respectively. The regression results using 66, 44, 40, and 27 countries are reported in the appendix in tables A.3-A.7.

findings of Silva & Tenreyro (2006) and these differences are attributable to biases associated with the inaccuracies of estimating gravity equations using log-linear models.

2.5.4 Interval Analysis

The results of the interval analysis discussed in the methods section are reported in Table 13. These results are comparable to those presented in Table 10 since both sets of results were estimated on the same number of observations. The coefficients on *GDP* and *GDP per capita,* as expected, are similar in magnitude to those reported in Table 10. All the *RER* coefficients reported in Table 13 are negative and statistically significant with the Poisson estimator having the greatest magnitude (-0.373). As reported in Table 10, the coefficient on *Non-member* (γ_4), the measure for trade diversion, is positive and statistically significant in the Heckman specification only. Its estimate in the Heckman model (0.481) implies an average increase in imports from non-member countries in the order of 62%. As a result, none of these models present any evidence of trade diversion and I have not included a more detailed analysis of this parameter.

In terms of the event fixed effects, the statistically significant years across the three models are: 1989, 1994, and 2002 which are the year democracy began in Paraguay, the year of high Brazilian inflation, as well as the end of the *Plan Convertibilidad* in Argentina, respectively. These results are somewhat consistent with those reported in Table 10 with the exception that the 2nd oil crisis of 1979 was statistically significant across all three estimators in the Table 10. The 2nd oil crisis is significant at the 90% level of confidence in the OLS specification only in Table 13.

Table 13 Estimation Results of Interval Analysis, full data set and country fixed effects.

	Pooled OLS		Heckman two-step				Poisson	
			Selection		Outcome			
	β	p-val.	β	P-val.	β	P-val.	β	p-val.
Time (β_1)	-0.097**	0.00	-0.009**	0.01	-0.119**	0.00	-0.066	0.13
GDP (β_2)	2.429**	0.00	0.423**	0.00	3.135**	0.00	1.283	0.41
GDPperCap. (β_3)	-1.297**	0.02	-0.105	0.13	-1.688**	0.00	1.323	0.45
RER (β_4)	-0.123*	0.08	-0.005	0.65	-0.131**	0.00	-0.373**	0.03
PM, 1980-82	0.138	0.26	0.184	0.46	0.075	0.82	-0.373**	0.03
PM, 1983-85	0.440**	0.00	-0.234	0.25	0.279	0.40	0.172	0.46
PM, 1986-88	-0.209	0.58	-0.317	0.12	-0.456	0.17	-0.150	0.64
PM, 1989-91	0.234	0.26	-0.085	0.71	0.070	0.83	0.220	0.58
PM, 1992-94	0.483	0.20	0.246	0.38	0.333	0.30	0.489	0.28
PM, 1995-97	1.236**	0.00	0.386	0.17	1.192**	0.00	1.035*	0.06
PM, 1998-00	1.242**	0.00	0.680*	0.05	1.210**	0.00	1.152*	0.06
PM, 2001-03	1.546**	0.00	0.130	0.61	1.426**	0.00	1.528**	0.02
PM, 2004-06	1.748**	0.00	-0.085	0.73	1.520**	0.00	1.529**	0.02
PM, 2007-10	2.142**	0.00	0.132	0.60	1.915**	0.00	1.408**	0.05
PNM, (γ_4)	0.176	0.11	0.406**	0.00	0.481**	0.00	0.194	0.23
1973	-0.028	0.75	0.102*	0.08	0.065	0.69	0.366**	0.03
1979	0.201*	0.06	0.085	0.12	0.249	0.10	-0.027	0.76
1982	0.030	0.77	-0.033	0.55	-0.003	0.98	-0.005	0.97
1984	-0.531**	0.00	-0.228**	0.00	-0.705**	0.00	-0.170	0.19
1989	-0.353**	0.00	-0.041	0.46	-0.393**	0.01	-0.224*	0.10
1991	0.261**	0.01	0.034	0.54	0.265*	0.07	0.195	0.12
1994	0.544**	0.00	0.179**	0.00	0.650**	0.00	0.357**	0.01
1995	0.000**	0.00	0.000**	0.00	0.000	0.19	0.000	0.64
1997	0.358**	0.00	-0.152**	0.00	0.219	0.11	0.054	0.69
1999	0.095	0.33	-0.272	0.00	-0.119	0.41	-0.344**	0.00
2002	-0.235**	0.02	-0.149**	0.00	-0.348**	0.01	-0.201**	0.02
Constant	-80.65**	0.00	-19.89**	0.00	-108.3**	0.00		
Inverse Mills (λ)					1.475**	0.00		
Fixed Effects	Yes		Yes		Yes		Yes	
Countries	152		152		152		152	
Observations	13,209		60,780		13,209		60,780	

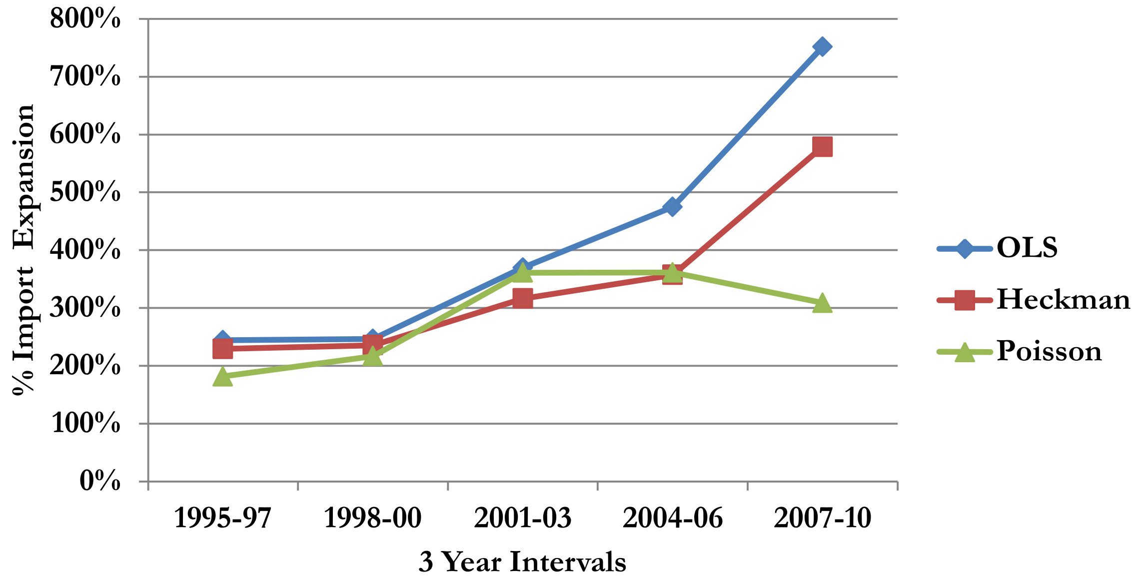
Note: *, **, *** imply significance at the 10%, 5%, and 1% levels, respectively.

The main results in this section are related to the coefficients on the interval dummy variables representing three-year averages from 1980-1982 to 2007-2010, this last period being of four years for practical purposes. These coefficients measure how much imports from Mercosur countries have increased with respect to the average level of imports between 1970 and 1979, the base years. A statistically significant result in any period dummy before 1995 is indicative of 'a priori' effects, an increase in trade related to firms presumably preparing themselves for business in soon to be more open markets. An inspection of the variables '*PM, 1980-82*', '*PM, 1983-85*', '*PM, 1986-88*', '*PM, 1989-91*' and '*PM, 1992-94*' reveals that none but *PM, 1983-85* are statistically significant in the OLS model. None of these coefficients are statistically significant in the Heckman model, and only *PM, 1980-82* is statistically significant with a negative sign in the Poisson model. These results not only suggest that there was no 'a priori' effect of Mercosur, but support the argument that the increases in regional trade observable after 1995 are genuinely attributable to the trade agreement.

The period dummy variables after 1995 illustrate how regional trade has evolved since Mercosur was created. An apparent trend in the period dummy variables since 1995 is that they are increasing in all models. These estimates capture the measure by how much imports from Mercosur countries have increased with respect to the group of base years 1970-1979. In short, these are measures of relative trade creation. The results are consistent in signs and statistical significance though the magnitudes of the period dummy variables vary by estimator.

Figure 5 below presents a graphical depiction of the three-year average levels of import expansion from Mercosur countries beginning in 1995. The magnitudes of these average values are very similar from 1995-97 to 2001-03. The OLS estimates of trade creation are the highest in all years, with import expansions of 244% for the 1995-1997 period up to 751% during the 2007-2010. Similarly, the Heckman model attenuates these estimates by correcting for sample selection bias. The estimates of the Heckman model for the mentioned periods increase from 229% in the first three year interval to 579% in the final three year interval. Finally, the Poisson estimates are somewhat less variable and begin with a measure of 182% in the first three year interval increasing to a measure of 361% for the 2001-2003 period before decreasing slightly to 309% for the final four-year interval. The result of the final period dummy (2007-10) contradicts the increasing tendencies observed in the OLS and Heckman estimates. Nevertheless, these estimates reflect overwhelming evidence of trade creation and higher levels of regional trade in more recent years.

Figure 5 Measure of Trade Creation in Paraguay attributable to Mercosur by three year intervals from 1995-2010.



2.6 Conclusion

I used the gravity model of international trade to analyze how the creation of Mercosur has affected import flows into Paraguay and compared the results of traditional and recently used estimators in gravity model applications. The main challenge in such studies is how 'zero trade flows' are addressed. After controlling for both country fixed effects as well as political and macroeconomic shocks, I compared a benchmark pooled OLS model to an application of the Heckman model and a Poisson model. While the Heckman model corrected an underestimation of 'gravity' variables in the pooled OLS estimates attributable to sample selection bias, there was no noticeable difference in the impact of membership to Mercosur between these two models. The Poisson estimates for trade creation were consistently lower than those of the other specifications, a finding consistent with the seminal work of Silva & Tenreyro (2006). Additionally, in testing the sensitivity of the estimates to varying sample sizes or trading partners, I found that of the three estimates of trade creation the Poisson estimators had the least variation. I found slight evidence of trade diversion using the Poisson estimator using Paraguay's 19 major trading partners only, but this trade diversion was clearly outweighed by trade creation in the same specification.

Additionally, a criticism of claims of trade creation in regional trade agreements is that it is difficult to disentangle the influences of historical and cultural from genuine policy effects. Including fixed effects in the models addresses these criticisms. Moreover, the dummy variables in of the interval analysis just discussed presented no evidence of trade creation before 1995, 'a priori' effects, which supports the trade creating effects of Mercosur. Much to the contrary, the main surges of regional imports occurred after Mercosur's creation, with even greater increases in imports in more recent times. These findings are strong evidence the trade creation observed in Paraguay is attributable to the trade agreement. Additionally, I found little evidence of trade diversion, only when the Poisson specification is estimated on Paraguay's 19 major trading partners, but outweighed by the measure of trade creation. The overwhelming evidence of trade creation found is this study demonstrates Mercosur has been beneficial to the Paraguayan economy.

Bibliography

Anderson, James A. 1979. "A Theoretical Foundation for the Gravity Equation." *American Economic Review*, 69 (1).

Anderson, J., and E. van Wincoop. 2003. "Gravity with Gravitas: A Solution to the Border Puzzle." *Americkan Economic Review*, 93.

Baldwin, R. and D. Taglioni. 2006. "Gravity for Dummies and Dummies for Gravity." NBER working paper No. 23516, Cambridge, MA: National Bureau of Economic Research.

Bayoumi, T., & Eichengreen, B. "Is regionalism simply a diversion? Evidence from the evolution of the EC and EFTA." Chapter 6, *Regionalism versus Multilateral Trade Arrangements*, Takatoshi Ito and Anne O. Krueger, editors. Chicago: University of Chicago Press, 1997.

Bergstrand, J. 1985. "The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence." *The Review of Economics and Statistics*, 20.

Berlinski Julio, Kume, Vaillant, Miranda, Ons, and Romero. 2005. "Aranceles a las Importaciones en el Mercosur: El Camino al Arancel Externo Común" Departamento de Economía, Universidad de la República, Uruguay.

Bhagwati, Jadish. 1996. "Preferential Trading Areas and Multilateralism- Strangers, Friends, or Foes?". Chapter 2, *Trading Blocs: Alternative Approaches to Analyzing Preferential Trade Agreements*, Jadiwh Bhagwati, Pravin Krishna, and Arvind Panagariya, editors. Cambridge, MA" The MIT Press, 1999.

Cameron, A. Colin and P. Travedi. 2005. *Microeconometrics: Methods and Applications*. New York, NY: Cambridge University Press.

Carrere, Celine. 2006. "Revisiting the Effects of Regional Trade Agreements on Trade Flows with Proper Specification of the Gravity Model." *European Economic Review*, 50.

Clausing, K. 2001. "Trade Creation and Trade Diversion in the Canada – United States Free Trade Agreement." *Canadian Journal of Economics*, 34 (3).

Deardorff, Alan. 1997. "Determinant of Bilateral Trade: Does Gravity Work in a Classical World?" In *The Regionalization of the World Economy,* ed. by Jeffrey Frankel. Chicago: University of Chicago Press.

Eaton, J., and A. Tamura. 1994. "Bilateralism and Regionalism in Japanese and US Trade and Direct Foreign Investment Patterns." *Journal of the Japanese and International Economics*, 8.

Egger, 2000. "A Note on the Proper Econometric Specification of the Gravity Equation." *Economic Letters*, 66 (1).

Feenstra R. C. 2002. "Border Effects and the Gravity Equation: Consistent Methods for Estimation." *Scottish Journal of Political Economy*, 49 (5).

Frankel, Jeffrey A. 1999. *No Single Currency Regime is Right for all Countries*. Princeton University, International Finance Section, Essays in International Finance, August, 215.

Frankel, Jeffrey A., E. Stein, and S.J. Wei. 1997. *Regional Trading Blocs in the World Economic System*. Washington, DC: Institute for International Economics.

Fukao, Kyoji, T. Okubo, and R.M. Stern. 2003. "An Econometric Analysis of Trade Diversion under NAFTA." *North American Journal of Economics and Finance*, 14 (1).

Heckman J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica*, 47 (1).

Helpman, E. 1987 "Imperfect Competition and International Trade: Evidence from Fourteen Industrial Countries," *Journal of the Japanese and International Economics*, 1 (1).

Helpman, E., Melitz, M., & Rubinstein Y. 2008. "Estimating Trade Flows: Trading Partners and Trading Volumes." *The Quarterly Journal of Economics*, 123 (2).

Helpman, E. & Krugman P. 1985. *Market Structure and Foreign Trade*, Cambridge: MIT Press,

Krueger, Anne O. 1999. "Trade Creation and Trade Diversion under NAFTA." *National Bureau of Economic Research*, Working Paper 7429 (December).

Krugman, Paul. 1991. "The Move to Free Trade Zones." *Economic Review* 76 (6).

Krugman, Paul. 1993. "Regionalism versus Multilateralism: Analytical Notes". Chapter 19, *Trading Blocs: Alternative Approaches to Analyzing Preferential Trade Agreements*, Jadish Bhagwati, Pravin Krishna, and Arvind Panagariya, editors. Cambridge, MA. The MIT Press, 1999.

Linders, Gert-Jan M. and Henri L.F. de Groot (2006). "Estimation of the Gravity Equation in the Presence of Zero Flows." Tinbergen Institute Discussion Paper, 2006-072/3.

Magee, C. 2008. "New Measures of Trade Creation and Trade Diversion." *Journal of International Economics*, 75 (2).

Martin W., Pham C. 2008. "Estimating the Gravity Model when Zero Trade Flows Are Important." mimeo, the World Bank.

Matyas, L. 1998. "The Gravity Model: Some Econometric Considerations." *The World Economy*, 21.

Matyas, L. 1997. "Proper Econometric Specification of the Gravity Model." *The World Economy*, 20.

Ramsey, J.B. 1969. "Tests for Specification Errors in Classical Linear Least Squares Regression Analysis." *Journal of the Royal Statistical Society B*, 31.

Santos Silva, J. M. C. and Tenreyro, S. 2006. "The Log of Gravity." *The Review of Economics and Statistics*, 88(4).

Soloaga, Isidro and L. Alan Winters. 2001. "Regionalism in the Nineties: What Effect on Trade?" *North American Journal of Economics and Finance* 12 (1).

Summers, Lawrence. 1991. "Regionalism and the World Trading System." Symposium Sponsored by the Federal Reserve Bank of Kansas City, *Policy Implications of Trade and Currency Zones*.

Tinbergen,J. 1962. *Shaping the World Economy: Suggestions for an international Economic Policy*. The Twentieth Century Fund, New York.

Viner, Jacob. 1950. *The Customs Union Issue*. New York: Carnegie Endowment for International Peace.

Wonnacot, Paul and Mark Lutz. 1989. "Is There a Case for Free Trade Areas?" In: Schott, Jeffrey J., ed., *Free Trade Areas and U.S. Trade Policy,* pp. 5-84. Washington, D.C.: Institute for International Economics.

Wooldridge, J. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

World Bank. 2011. Data retrieved from World Development Indicators Online (WDI) database.

Yeats, Alexander J. 1998. "Does Mercosur's Trade Performance Raise Concerns about the Effects of Regional Trade Arrangements?" *World Bank Economic Review*, 12 (1).

Appendix

Table A.1 Exporting Countries to Paraguay 1970-2010, Chapter 1

Afghanistan	Gabon	Nigeria
Albania	Gambia	Norway
Algeria	Georgia	Occ. Palestinian Terr.
Angola	Germany	Oman
Argentina	Ghana	Pakistan
Australia	Greece	Panama
Austria	Grenada	Peru
Bahamas	Guatemala	Philippines
Bahrain	Guinea	Poland
Bangladesh	Guyana	Portugal
Barbados	Haiti	Rep. of Korea
Belarus	Honduras	Rep. Moldova
Belgium	Hungary	Romania
Belize	Iceland	Russian Federation
Bermuda	India	Rwanda
Bolivia	Indonesia	Saudi Arabia
Bosnia Herzegovina	Iran	Senegal
Botswana	Iraq	Serbia
Brazil	Ireland	Sierra Leone
Brunei Darussalam	Israel	Singapore
Bulgaria	Italy	Slovakia
Burundi	Jamaica	Slovenia
Cambodia	Japan	South Africa
Cameroon	Jordan	Spain
Canada	Kenya	Sri Lanka
Central African Rep.	Kuwait	Sudan
Chile	Kyrgyzstan	Suriname
China	Lao People's Dem. Rep.	Swaziland
China, Hong Kong	Latvia	Sweden
China, Macao SAR	Lebanon	Switzerland
Colombia	Lesotho	Syria
Congo	Liberia	Tajikistan
Costa Rica	Libya	TFYR of Macedonia
Côte d'Ivoire	Lithuania	Thailand
Croatia	Luxembourg	Togo
Cuba	Madagascar	Trinidad and Tobago
Cyprus	Malaysia	Tunisia
Czech Rep.	Malta	Turkey
Dem. R. of Congo	Mauritania	Turkmenistan

Denmark	Mauritius	Uganda
Dominica	Mexico	Ukraine
Dominican Rep.	Mongolia	United Arab Emirates
Ecuador	Morocco	United Kingdom
Egypt	Mozambique	U. Rep. of Tanzania
El Salvador	Myanmar	Uruguay
Equatorial Guinea	Namibia	USA
Estonia	Nepal	Uzbekistan
Ethiopia	Netherlands	Venezuela
Finland	New Caledonia	Viet Nam
France	New Zealand	Yemen
French Polynesia	Nicaragua	Zimbabwe
FS Micronesia	Niger	

Table A.2 Reduced Data Set of Exporting Countries to Paraguay 1970-2010.

USA	Japan
Argentina	Mexico
Belgium	Netherlands
Brazil	Rep. of Korea
Canada	Spain
Chile	Switzerland
Denmark	United Kingdom
France	Uruguay
Germany	USA
Italy	

Table A.3 Estimation Results with country fixed effects on 66 countries.

	Pooled OLS		Heckman two-step				Poisson	
			Selection		Outcome			
	β	p-val.	β	p-val.	β	p-val.	β	p-val.
Time (β_1)	-0.099	0.000	-0.026	0.000	-0.117	0.000	-0.053	0.257
GDP (β_2)	2.770	0.000	0.607	0.000	3.185	0.000	1.286	0.387
GDP per Cap. (β_3)	-1.757	0.002	-0.144	0.117	-1.936	0.000	1.182	0.483
RER (β_4)	-0.084	0.240	0.021	0.074	-0.076	0.050	-0.243	0.175
PM, *Member* (γ_3)	**1.283**	**0.000**	**0.378**	**0.005**	**1.324**	**0.000**	**0.936**	**0.002**
PNM, *Non-member* (γ_4)	**0.097**	**0.363**	**0.474**	**0.000**	**0.299**	**0.041**	**-0.045**	**0.717**
1973	0.013	0.886	0.098	0.117	0.060	0.711	0.484	0.016
1979	0.233	0.020	0.005	0.938	0.227	0.133	0.101	0.024
1982	0.062	0.540	-0.060	0.318	0.031	0.838	-0.073	0.398
1984	-0.511	0.000	-0.253	0.000	-0.612	0.000	-0.102	0.462
1989	-0.393	0.000	-0.016	0.791	-0.405	0.006	-0.270	0.012
1991	0.204	0.043	0.106	0.076	0.239	0.098	0.152	0.146
1994	0.508	0.000	0.291	0.000	0.612	0.000	0.397	0.001
1995	0.000	0.000	0.000	0.000	0.000	0.026	0.000	0.629
1997	0.333	0.001	-0.083	0.160	0.285	0.033	0.017	0.883
1999	0.028	0.785	-0.173	0.003	-0.048	0.724	-0.398	0.000
2002	-0.336	0.002	-0.083	0.146	-0.368	0.005	-0.237	0.010
Constant	-90.135	0.000	-27.9	0.000	-103.4	0.000		
Inverse Mills (λ)					0.798	0.032		
Fixed Effects	Yes		Yes		Yes		Yes	
Countries	66		66		66		66	
Observations	12,403		26,860		12,403		26,860	

Table A.4 Estimation Results with country fixed effects on 44 countries.

	Pooled OLS		Heckman two-step				Poisson	
			Selection		Outcome			
	β	p-val.	β	p-val.	β	p-val.	β	p-val.
Time (β_1)	-0.103	0.000	-0.059	0.000	-0.122	0.000	-0.100	0.006
GDP (β_2)	3.069	0.000	1.357	0.000	3.499	0.000	2.864	0.007
GDP per Cap. (β_3)	-2.175	0.001	-0.837	0.000	-2.439	0.000	-0.385	0.769
RER (β_4)	-0.054	0.516	0.056	0.000	-0.038	0.396	-0.231	0.216
PM, *Member* (γ_3)	**1.182**	**0.000**	**0.463**	**0.001**	**1.217**	**0.000**	**0.671**	**0.000**
PNM, *Non-member* (γ_4)	**0.091**	**0.406**	**0.408**	**0.000**	**0.201**	**0.199**	**-0.005**	**0.961**
1973	-0.029	0.746	0.084	0.239	-0.003	0.987	0.615	0.002
1979	0.325	0.000	-0.053	0.442	0.309	0.054	0.120	0.013
1982	0.102	0.341	-0.065	0.347	0.082	0.604	-0.057	0.539
1984	-0.503	0.000	-0.266	0.000	-0.569	0.001	-0.070	0.636
1989	-0.357	0.000	-0.013	0.855	-0.365	0.018	-0.349	0.000
1991	0.187	0.072	0.117	0.097	0.211	0.164	0.095	0.338
1994	0.512	0.000	0.302	0.000	0.578	0.000	0.305	0.005
1995	0.000	0.000	0.000	0.053	0.000	0.012	0.000	0.609
1997	0.288	0.009	-0.110	0.128	0.253	0.086	-0.002	0.983
1999	-0.075	0.509	-0.158	0.024	-0.117	0.427	-0.430	0.000
2002	-0.403	0.001	-0.117	0.094	-0.430	0.003	-0.277	0.003
Constant	-98.077	0.000	-49.6	0.000	-111.51	0.000		
Inverse Mills (λ)					0.516	0.288		
Fixed Effects	Yes		Yes		Yes		Yes	
Countries	44		44		44		44	
Observations	10,850		18,040		10,850		18,040	

Table A.5 Estimation Results with country fixed effects on 40 countries.

	Pooled OLS		**Heckman two-step**				**Poisson**	
			Selection		**Outcome**			
	β	**p-val.**	**β**	**p-val.**	**β**	**p-val.**	**β**	**p-val.**
Time (β_1)	-0.105	0.000	-0.048	0.000	-0.126	0.000	-0.062	0.018
GDP (β_2)	3.057	0.000	1.261	0.000	3.612	0.000	3.173	0.000
GDP per Cap. (β_3)	-2.117	0.002	-0.886	0.000	-2.521	0.000	-2.411	0.031
RER (β_4)	-0.033	0.661	-0.002	0.917	-0.037	0.395	-0.462	0.004
PM, *Member* (γ_3)	**1.201**	**0.000**	**0.389**	**0.006**	**1.224**	**0.000**	**0.418**	**0.003**
PNM, *Non-member* (γ_4)	**0.099**	**0.369**	**0.254**	**0.000**	**0.197**	**0.171**	**-0.147**	**0.098**
1973	-0.018	0.832	0.054	0.460	0.005	0.977	0.163	0.059
1979	0.327	0.000	-0.039	0.591	0.311	0.053	0.174	0.000
1982	0.086	0.434	-0.079	0.273	0.057	0.722	0.040	0.686
1984	-0.512	0.000	-0.245	0.001	-0.596	0.001	-0.043	0.759
1989	-0.358	0.000	0.033	0.661	-0.349	0.025	-0.156	0.029
1991	0.184	0.087	0.111	0.134	0.217	0.162	0.117	0.328
1994	0.602	0.000	0.241	0.002	0.674	0.000	0.312	0.001
1995	0.000	0.000	0.000	0.803	0.000	0.005	0.000	0.112
1997	0.288	0.013	-0.007	0.925	0.274	0.065	0.163	0.064
1999	-0.080	0.515	-0.083	0.267	-0.112	0.452	-0.324	0.002
2002	-0.422	0.001	-0.098	0.187	-0.452	0.003	-0.297	0.007
Constant	-98.431	0.000	-43.8	0.000	-115.6	0.000		
Inverse Mills (λ)					0.735	0.219		
Fixed Effects	Yes		Yes		Yes		Yes	
Countries	40		40		40		40	
Observations	10,327		16,400		10,327		16,400	

Table A.6 Estimation Results with country fixed effects on 27 countries.

	Pooled OLS		Heckman two-step				Poisson	
			Selection		Outcome			
	β	p-val.	β	p-val.	β	p-val.	β	p-val.
Time (β_1)	-0.109	0.000	-0.036	0.000	-0.123	0.000	-0.049	0.063
GDP (β_2)	3.074	0.000	0.858	0.000	3.419	0.000	2.341	0.001
GDP per Cap. (β_3)	-2.036	0.022	-0.498	0.004	-2.251	0.000	-1.316	0.113
RER (β_4)	-0.001	0.984	-0.024	0.246	-0.015	0.772	-0.337	0.012
PM, *Member* (γ_3)	**1.198**	**0.000**	**0.444**	**0.003**	**1.254**	**0.000**	**0.572**	**0.000**
PNM, *Non-member* (γ_4)	**0.091**	**0.452**	**0.252**	**0.001**	**0.175**	**0.308**	**-0.149**	**0.077**
1973	-0.006	0.949	-0.014	0.879	-0.005	0.978	0.149	0.059
1979	0.332	0.004	-0.106	0.236	0.295	0.118	0.157	0.000
1982	0.143	0.245	-0.087	0.333	0.113	0.539	0.028	0.755
1984	-0.425	0.000	-0.162	0.065	-0.474	0.013	-0.053	0.710
1989	-0.363	0.002	0.139	0.147	-0.323	0.079	-0.193	0.001
1991	0.276	0.047	0.023	0.810	0.281	0.119	0.162	0.166
1994	0.602	0.001	0.285	0.006	0.676	0.001	0.389	0.000
1995	0.000	0.001	0.000	0.709	0.000	0.007	0.000	0.115
1997	0.405	0.004	-0.042	0.678	0.382	0.032	0.207	0.027
1999	-0.024	0.862	-0.103	0.280	-0.057	0.747	-0.288	0.004
2002	-0.501	0.001	-0.050	0.600	-0.509	0.004	-0.317	0.005
Constant	-101.3	0.000	-30.5	0.000	-110.75	0.000		
Inverse Mills (λ)					0.720	0.394		
Fixed Effects	Yes		Yes		Yes		Yes	
Countries	27		27		27		27	
Observations	8,194		11,070		8,194		11,070	