TWO ESSAYS ON THE ESTIMATION OF AGRICULTURAL TRADE FLOWS: MODEL SELECTION AND ENDOGENEITY OF TRADE AGREEMENTS

by

David R. Thibodeau

Submitted in partial fulfilment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
November 2015

TABLE OF CONTENTS

LIST OF TABLES

ABSTRACT

This thesis investigates three econometric issues relevant to the estimation of gravity models of agriculture trade, namely the issues of zero trade flows, heteroskedasticity, and the endogeneity of preferential trade agreements (PTAs). Recent studies have employed the Poisson pseudo-maximum likelihood (PPML) estimator to address zero trade and heteroskedasticity, while a second line of work has argued for better modelling of zero trade through limited dependent variable models, with little consideration for the heteroskedasticity problem. The first essay of this thesis proposes the use of a Box-Cox double hurdle (BCDH) model as a way to address both econometric issues. The second essay in this thesis challenges Sun and Reed's (2010) conclusion that NAFTA has no impact on members' agricultural trade. The two-step estimator developed by Treza (1998) provides evidence of substantial endogeneity bias in their estimated effects for nearly all preferential trade agreements, including NAFTA.

LIST OF ABBREVIATIONS USED

ASEAN	Association of Southeast Asian Nations
BCDH	Box-Cox double hurdle
COMESA	Common Market for Eastern and Southern Africa
EU15	the first 15 countries of the European Union
EU25	the additional 10 countries that entered into the European union in 2004
GDP	Gross domestic product
MRT	Multilateral resistance term
NAFTA	North American Free Trade Agreement
NLS	Nonlinear least squares
PPML	Poisson pseudo-maximum likelihood
PTA	Preferential trade agreement
WLS	Weighted least squares

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Stephen Clark, who persistently encouraged the completion of this work. I have had the fortune of working with Professor Clark on a number of projects over the years. He has had by far the greatest influence on my development as an economist, for which I am extremely grateful.

I would also like to acknowledge the entire faculty at the Department of Business & Social Sciences for their support over the years. Through my career I have seen first-hand how their dedication to student learning has translated into highly-skilled professionals, who work tirelessly towards the betterment of the agriculture sector.

Foremost, I would like to thank my wife Jessica, and my children Justin, Madeleine and Jonah, whose love and support made this work possible.

CHAPTER 1: INTRODUCTION

There is a growing body of empirical work for which bilateral trade is modelled as a function of the economic sizes and distance between two regions. Given the similarity to Newton's law of gravity, these type of models are commonly referred to as gravity equations. Stated more formally, various economic theories suggest that the following relationship should hold for the value of bilateral trade (T_{ij}) between countries i and j:

$$T_{ij} = b_0 Y_i^{b_1} Y_j^{b_2} \boldsymbol{R}_{\boldsymbol{ij}}^{\boldsymbol{\gamma}} \tag{1}$$

where Y_i is gross domestic product (GDP) in country i, Y_j is GDP in country j, $\boldsymbol{R}_{\boldsymbol{ij}}$ is a set of trade resistance factors predicted to influence trade by the underlying theory, b_0 is an intercept term, and b_1, b_2 and $\boldsymbol{\gamma}$ are the elasticities of trade with respect to the various factors (e.g. Eaton and Kortum, 2002; Anderson and van Wincoop, 2003; Melitz and Ottaviano, 2008). This constant elasticity model has been the foundation of numerous econometric studies on international trade, with the log-linear form being a popular empirical specification (Santos Silva and Tenreyro, 2006; Xiong and Chen, 2014; Liu, 2009).

The popularity of the gravity equation can be linked to the model's apparent success in previous studies, which is likely linked to its theoretical underpinnings. Numerous empirical studies provide support for this model[1], plus several recent economic trade theories (e.g. Helpmen et al, 2008; Eaton and Kortum, 2002; Anderson and van Wincoop, 2003; Melitz and Ottaviano, 2008). As an econometric approach to trade analysis, one of the main benefits of the gravity equation is that it provides a retrospective examination of trade flows using observed data. In contrast, computable equilibrium models of trade have

[1] Santos Silva and Tenreyro (2006) provide numerous examples of applied work.

been criticised for having parameters with questionable econometric foundations (Hertel et al, 2007). Overall, the gravity equation is regarded as a useful analytical tool for determining how factors such as trade organizations, preferential trade agreements, and tariff and non-tariff barriers influence bilateral trade (e.g. Liu, 2009; Hoekman and Nicita 2011; Sun and Reed, 2010).

The main theoretical models supporting the gravity equation also provide predictions on the spatial pattern of zero trade. Not surprisingly, these models suggest that the likelihood of export zeroes increases with trade costs (e.g. Helpmen et al, 2008; Eaton and Kortum, 2002; Melitz and Ottaviano, 2008). Therefore, empirical studies should find that distance and importer remoteness is negatively correlated with the probability of trade. Extensions of the monopolistic competition model of trade, in which trade costs are passed on to consumers, suggest that zero trade can arise due to a choke-price or a fixed market entry cost (beachhead cost) in the importing country (Melitz and Ottaviano, 2008; Helpmen et al, 2008). These models suggest a single profitability condition or hurdle arising from trade costs, and that firms should export to any country for which this condition is satisfied.

Addressing zero trade flows has been a focus of more recent studies using the gravity equation. The empirical model has traditionally been estimated using the log-linear function implied by a number of theoretical models (Sanso et al 1993; Bergstrand, 1989; Feenstra et al, 2001). However, when applying the gravity equation to datasets with many regions and differentiated products, researchers will likely encounter a significant number of observations with zero trade. Since the log of zero is undefined, the log-linear form necessitates removal of zero observations from the sample, a practice that could introduce selectivity bias (Heckman, 1979). Several recent trade studies have addressed this issue by using limited dependent variable models, with a log-linear version of Heckman's sample selection model being a popular choice (e.g. Olper and Raimondi, 2008; Portugal-Perez and Wilson, 2012; Cipollina and Salvatici 2010; Haq et al, 2011)

A second line of work has advocated against using the log transformation of the gravity model (e.g. Silverstovs and Schumacher, 2009; Liu, 2009; Hoekman and Nicita, 2011; Felbermayr and Kohler 2010; Sun and Reed 2010). While this approach avoids the log of zero issue, the main argument against the log-linear model has been one of consistency. Previous studies have identified that the log-linear model will not provide consistent elasticity estimates under heteroskedasticity. The conditional expectation of T_{ij} is a function of the regression error variance after transforming the model from the log scale back to the original scale. Because of the correlation between the variance and the regressors under heteroskedasticity, the parameters of the log-linear model are not consistent elasticity estimates for the conditional expectation of T_{ij} (Mullahy, 1998; Manning, 1998).

Efforts to address zero trade and heteroskedasticity in the gravity equation have led to a notable divergence in econometric methods. Santos Silva and Tenreyro (2006) suggest using the Poisson pseudo-maximum likelihood (PPML) estimator of the untransformed model as a way to address both issues. PPML avoids the retransformation problem, and has been shown to be robust to heavy clustering at zero for the dependent variable (Santos Silva and Tenreyro, 2011). However, it has been argued that PPML does not properly handle zero trade, because all zero observations are treated equally. Some zeroes might be at the margin of trade, while other zeros could reflect cases far from the margin, with a very low probability of trade (Jayasinghe et al, 2010). This concern has led to the use of standard limited dependent variable models to better explain the censoring rule for zeros, and to address the issue of selectivity bias in the gravity equation. However, nearly all trade studies using limited dependent variable models have paid no attention to the problems arising from heteroscedasticity.

The first essay in this thesis proposes the use of a Box-Cox double hurdle (BCDH) model as a way to address both the zero trade and heteroskedasticity issues. In addition to providing a better approximation to the conditional mean of trade, data-dependent transformations of the variables can produce a model which better reflects the assumptions made in the statistical analysis, such as homoskedasticity and normally distributed errors (Box and Cox, 1964). An empirical exercise using data from Santos Silva and Tenreyro (2006) demonstrates that the BCDH with dependent errors can outperform all other nested models. Furthermore, a heteroskedastic version of the P-test proposed by Davidson and MacKinnon (1981) provides evidence that favours the BCDH over PPML. Comparing the results of both estimators reveals notable differences in the effects of key gravity equation variables, such as GDP, distance and preferential trade agreements (PTAs).

The differences in the estimated PTA effects in the first essay is particularly noteworthy, given that the gravity model is a commonly used tool for analysing the ex-post trade enhancement effects of PTAs. Trade liberalization through bilateral and regional PTAs has greatly reduced international trade barriers, and recent analysis suggests that there are substantial gains to be had through trade liberalization (Hoekman and Nicita, 2011). However, only recently have econometric techniques been used to address the possible endogeneity of PTAs. Baier and Bergstrand (2007) argue that omitted variables that inhibit trade might be positively correlated with the formation of a PTA, leading to a negative correlation between the error term of the gravity equation and the PTA variable. As a result, the effects of PTAs on bilateral trade might be substantially underestimated.

The second essay in this thesis explores the endogeneity issue by re-evaluating Sun and Reed's (2010) conclusion that NAFTA has no effect on members' agricultural trade. The two-step estimator developed by Treza (1998), which includes an endogeneity correction, provides evidence of bias in Sun and Reed's (2010) estimates on nearly all preferential trade agreements (PTAs). The estimates after including a correction term for endogeneity show that NATFA increased members' trade by 58% to 105%, from 1996 to 2007. In addition, the estimated effects for a shared PTA in 2007 increase from 53% to 235% for ASEAN, 90% to 208% for COMESA, 60% to 140% for the EU15 countries, and 79% to 230% for the EU25 countries[2]. These findings have important implications outside of the data analysed in this study. Nearly all current agricultural trade studies using the gravity model have not explored this endogeneity problem. The evidence produced in this study suggests that the estimated effects of PTAs found in recent papers could be substantially biased.

The outline of this thesis is as follows. Chapter 2 starts with a discussion on some of the main econometric issues faced when estimating the gravity equation, and the current methods being used in applied work. A review of the literature reveals a notable

[2] The agreements studied include the North American Free Trade Agreement (NATFA), the Association of Southeast Asian Nations (ASEAN), the Common Market for Eastern and Southern Africa (COMESA), the first 15 countries of the European Union (EU15), and the additional 10 countries that entered into the European union in 2004 (EU25).

split between those advocating the PPML estimator, and those recommending the use of limited dependent variable models, such as Heckman's sample selection. Based on this literature review, this thesis proposes the use of the BCDH model in Section 2.4, along with a non-nested model selection test in Section 2.5 for comparing the BCDH and PPML. Empirical results supporting the BCDH model are presented in Section 2.6. Chapter 3 starts with a discussion of the issues around PTAs and agricultural trade. Section 3.3 highlights some of the most recent empirical evidence on the effects of PTAs using gravity models, while Section 3.4 discusses the potential for endogeneity bias in these estimated effects. Sections 3.5 to 3.8 provide a re-examination Sun and Reed's (2010) results, with details on Treza's (1998) estimator given in section 3.7. Chapter 4 highlights the major findings of the two empirical studies.

CHAPTER 2: A BOX-COX DOUBLE HURDLE GRAVITY MODEL

2.1 Introduction

Despite a long history of empirical studies using the gravity equation, there is still ongoing discussions on how to best estimate the model. Recent studies have focused primarily on issues related to sample selection and retransformation bias (e.g. Santos Silva and Tenreyro, 2006; Egger et al, 2011; Helpman et al, 2008; Xiong and Chen, 2014)). The evidence produced suggests that the traditional practice of estimating the log-linear model over only positive observations is not appropriate. There is likely unobserved and correlated factors that influence both the probability of trade and the trade level. If not addressed, the results of the log-linear model would be subject to selectivity bias (Heckman, 1979). A second issue of concern has been the retransformation bias of the log-linear model with heteroskedastic errors. This issue appears to be first noted in the trade literature by Santos Silva and Tenreyro (2006). In the case of the log-linear gravity model, the conditional expectation of T_{ij} is a function of the regression error variance after transforming the model from the log scale back to the original level scale. If there is correlation between the error variance and the regressors in the log-linear model, the parameters are not consistent elasticity estimates for the conditional expectation of the dependent variable (Mullahy, 1998; Manning, 1998).

While selectivity bias and heteroskedasticity have been at the forefront of recent econometric discussions on the gravity equation, empirical studies have generally focused on only one of these two issues at a time. Santos Silva and Tenreyro (2006) recommend Poisson pseudo-maximum likelihood (PPML) to address the issue of heteroskedasticity. PPML can also be estimated over zero observations, although no specific modelling of a censoring rule is made. Due to interest in modelling the probability of trade, and concerns over how PPML handles zero observations, other studies have used limited dependent variable models (e.g. Olper and Raimondi, 2008; Portugal-Perez and Wilson, 2012; Cipollina and Salvatici 2010; Haq et al, 2011). Xiong and Chen's (2014) trade study appears to be the only one that incorporates the concerns of non-constant variance into a limited dependent variable model. This lack of concern over heteroskedasticity is surprising, since it is well known among household survey analysts that limited dependent variable models produce inconsistent estimates under heteroskedasticity and non-normality (Jones and Yen, 2000). It appears that these relevant econometric issues have been largely overlooked in recent trade studies.

This study attempts to address the above concerns by evaluating the Box-Cox double hurdle (BCDH) model against other popular methods of estimating the gravity equation, such as PPML and Heckman's sample selection model. Data dependent transformations such as the Box-Cox may address heteroskedasticity and non-normality, and can provide a better approximation to the conditional mean (Box and Cox, 1964). Given that the data studied by Santos Silva and Tenreyro (2006) was used to promote PPML as an estimator of the gravity equation, this study evaluates the BCDH model using the same data. The BCDH model specified in this study nests the log-linear sample selection and Tobit models, providing simple parametric tests against specifications used in previous studies. Since BCDH and PPML are non-nested models, a test proposed by Davidson and MacKinnon (1981) is used to evaluate the performance of these two specifications.

The model selection tests conducted in this study strongly support the BCDH over all nested alternatives, including the log-linear sample selection model. The non-nested model testing results also favour the BCDH over PPML. To better understand the importance of these findings, elasticity estimates and binary effects are compared to see if BCDH and PPML produce significantly different results. The two models generate notable differences in the effects of key gravity equation variables, such as GDP, distance and preferential trade agreements. After accounting for multilateral resistance through importer and exporter fixed effects, PPML provides evidence that trade agreements enhance trade between members and non-members. In contrast, the BCDH provides evidence of trade diversion, through a reduced probability of trade and lower conditional trade level.

The next two sections of this study discuss some of the econometric issues faced when estimating the gravity equation, and the current methods being used in applied work. A review of the literature reveals a notable split between those advocating the PPML estimator, and those recommending the use of limited dependent variable models, such as Heckman's sample selection. Based on these findings, this study proposes the use of the BCDH model in Section 2.4, and proposes a non-nested model selection test in Section 2.5 for comparing the BCDH and PPML. Empirical results supporting the BCDH model are presented in Section 2.6, followed by concluding remarks in Section 2.7.

2.2 Model Specification Issues for the Gravity Equation

Numerous specification issues for the gravity equation have been discussed in recent years. The study by Anderson and van Wincoop (2003) resulted in a major shift in how the gravity equation is specified. The authors shows that the traditional gravity equation (1) is misspecified, because the model omits a measure of the exporter's openness to trade, and the importer's aggregate price level. These factors, which Anderson and van Wincoop (2003) refer to as multilateral trade resistance, can be accounted for by using importer and exporter fixed effects. These additions to (1) yield the following equation

$$T_{ij} = b_0 Y_i^{b_1} Y_j^{b_2} \boldsymbol{R}_{\boldsymbol{ij}}^{\boldsymbol{\gamma}} e^{a_i d_i + a_j d_j} \tag{2}$$

where a_i and a_j are parameters on the dummy variables d_i and d_j, which identify the exporter and importer, respectively. Each importer and exporter fixed effect is estimated across the full sample, so that equation (2) can be estimated with cross-sectional data. Unlike a standard fixed effects model for panel data, a time dimension is not required to identify a_i and a_j. Starting with a sample of *N* regions, adding a new region would create two new fixed effects to estimate, as well as 2*N* additional observations for estimating all parameters of the model. Asymptotically, the fixed effects are not estimated over a finite number of observations, so there is no small sample bias (i.e. incidental parameters problem) for the maximum likelihood estimator of equation (2).[3]

Retransformation bias is another concern that appears to have been discussed in the health economics literature well before the trade literature. Manning (1998) discusses several issues related to retransformation of the log-linear model, including consistent estimation of the partial derivatives when the error term exhibits non-constant variance. The case of normally distributed heteroskedastic errors for log-linear models such as (1) or (2) is particularly illustrative of the retransformation problem. Grouping the gravity

---

[3] Greene (2004) gives a detailed examination of the incidental parameters problem for several limited dependent variable models, including the Tobit and Probit models.

equation regressors within a vector $\boldsymbol{x_{ij}}$, the regression model with log normally distributed heteroskedastic errors u_i can be defined as:

$$ln(T_{ij}) = ln(\boldsymbol{x_{ij}})\boldsymbol{\beta} + ln(u_{ij})$$

$$ln(u_{ij}) \backsim N\left(0, \sigma^2(\boldsymbol{x_{ij}})\right)$$

Transforming back to $E[T_{ij}|\boldsymbol{x_{ij}}]$ yields:

$$E[T_{ij}|\boldsymbol{x_{ij}}] = e^{ln(\boldsymbol{x_{ij}})\boldsymbol{\beta} + 0.5\sigma_{ij}^2}$$

$$\frac{\partial E[T_{ij}|\boldsymbol{x_{ij}}]}{\partial x_{ijk}} = E[T_{ij}|\boldsymbol{x_{ij}}]\left[\frac{\beta_k}{x_{ijk}} + 0.5\frac{\partial \sigma_{ij}^2}{\partial x_{ijk}}\right] \quad (3)$$

Equation (3) illustrates that under heteroskedasticity the estimates $\boldsymbol{\beta}$ of the log-linear model are no longer consistent estimates of the elasticities for the untransformed dependent variable. On the other hand, when the variance of $ln(T_{ij})$ is constant, β_k is a consistent estimate of the elasticity of $E[T_{ij}|x_{ij}]$ with respect to x_{ijk}. Santos Silva and Tenreyro (2006) show that if the variance of $ln(T_{ij})$ is constant, then the variance of T_{ij} must be proportional to $E[T_{ij}|x_{ij}]^2$. They argue that while the variance is likely increasing in the mean, it is unlikely that the variance is exactly proportional to the mean squared, and that some amount of bias will generally occur when estimating the log-linear model. It is important to note that this issue extends to other areas of applied work using the log-linear model.

Even though nonlinear least squares (NLS) avoids the retransformation problem, Santos Silva and Tenreyro (2006) also argue against this estimator. NLS gives more weight to observations with larger values for the conditional expectation, which are likely the observations with the largest variance. This issue can be seen by considering the first-order conditions of the NLS estimator:

$$\sum_n^N exp\left(ln\left(\boldsymbol{x_{ijn}}\right)\boldsymbol{\beta}\right)ln\left(\boldsymbol{x_{ijn}}\right)\left(T_{ijn} - exp\left(ln\left(\boldsymbol{x_{ijn}}\right)\boldsymbol{\beta}\right)\right) = \mathbf{0} \tag{4}$$

More weight is given to observations where $exp\left(ln\left(\boldsymbol{x_{ijn}}\right)\boldsymbol{\beta}\right)$ is large, due to the curvature of the exponential function. Since more weight is likely given to noisy observations, the NLS estimator is likely very inefficient.

Concerns over possible selectivity bias have also been raised in recent studies. It has been argued that positive trade follows the gravity model specification, while the zeroes are generated by a separate process. Therefore, the gravity models defined by (1) or (2) are not the correct specification $E\left[T_{ij}|\boldsymbol{x_{ij}}\right]$. If this assumption is true, the expected value of trade is a double-index model, which can be represented as the product of a model defining the expected value of trade over positive trade flows, and a model defining the probability of trade, with appropriate intercept shifts where applicable (Santos Silva et al, 2015). For example, the conditional expected value of trade for the log-linear sample selection model is

$$E\left[T_{ij}|\boldsymbol{x_{ij}}\right] = exp\left(ln\left(\boldsymbol{x_{ij}}\right)\boldsymbol{\beta} + \sigma^2/2\right)\Phi\left(\boldsymbol{z_{ij}}\boldsymbol{\alpha} + \rho\sigma\right) \tag{5}$$

where and $\boldsymbol{z_{ij}}$ and $\boldsymbol{\alpha}$ are regressors and parameter estimates from a probit model, σ is the variance of the error term in the gravity equation, ρ is an estimate of the population correlation coefficient between the residuals of the selection equation and the gravity equation, and $\Phi(\cdot)$ is the standard normal cumulative distribution function.

2.3 Addressing Specification Issues – PPML vs. Sample Selection

Efforts to address zero trade and heteroskedasticity in the log-linear gravity equation have led to a notable divergence in econometric methods. Santos Silva and Tenreyro (2006) recommend an estimator that gives equal weight to all observations, which amounts to the following first order conditions:

$$\sum_{n}^{N} ln(\boldsymbol{x_{ijn}})\left(T_{ijn} - exp(ln(\boldsymbol{x_{ijn}})\boldsymbol{\beta})\right) = \boldsymbol{0} \tag{6}$$

for a dataset with *N* observations. The pseudo maximum likelihood estimator with conditions (6) is equivalent to a constant elasticity PPML model, for which $E[T_{ij}|\boldsymbol{x_{ij}}] = exp(ln(\boldsymbol{x_{ij}})\boldsymbol{\beta}) \propto VAR[T_{ij}|\boldsymbol{x_{ij}}]$, where $\propto$ indicates the conditional mean is proportional to the variance. While PPML assumes the variance of the dependent variable is proportional to the conditional mean, the estimator has been shown to be robust to deviations from this assumption. PPML will produce consistent estimates as long as the model is correctly specified, regardless of the distribution of the dependent variable (Gourieroux et al, 1984). As a result, heavy clustering at zero for the dependent variable should not affect the consistency of the PPML estimates, as was shown by Santos Silva and Tenreyro's (2011) simulation study.

Since PPML was initially proposed for estimating the gravity equation, it has been applied in a number of trade studies (e.g. Silverstovs and Schumacher, 2009; Liu, 2009; Hoekman and Nicita, 2011; Felbermayr and Kohler 2010; Sun and Reed 2010). It is fairly common for studies using PPML to compare the results to OLS, either on the truncated sample, or the OLS estimates obtained when adding a small constant to the dependent variable. These results show that OLS estimates on preferential trade agreements, distance, common language and colonial ties tend to be greater in magnitude than PPML estimates (Felbermayr and Kohler, 2010; Sun and Reed 2010; Silverstovs and Schumacher, 2009). Liu (2009) also compared PPML to the Tobit model, and found that the Tobit produced unreasonably high effects for WTO membership. The Tobit estimates from Liu (2009) suggest that the mean value of exports would increase 300% if either the importer or exporter was a WTO member, or increase 700% if both trade partners were WTO members. In contrast, the PPML estimates suggest increases of only 23% and 60%, respectively.

A second line of work has generally ignored the warnings of heteroskedasticity, focusing instead on modelling the censoring rule for positive trade using traditional limited dependent variable models (Olper and Raimondi, 2008; Portugal-Perez and Wilson, 2012; Cipollina and Salvatici 2010; Haq et al, 2011). In addition to directly addressing the zero observations, these models allow for analysis on the probability and level of trade. Jayasinghe et al (2010) also argue that limited dependent variable models might be superior to PPML, because PPML does not explicitly model the zero observations. As the authors point out, some zeroes may be at the margin of trade, while other zeros could reflect cases far from the margin where there is a very low probability of observing trade.

While several models have been developed that address zero observations and the selectivity issue, such as the Tobit, Double Hurdle and the Two-part model, the log-linear version of Heckman's (1979) sample selection model has been a popular choice in trade research (e.g. Olper and Raimondi, 2008; Portugal-Perez and Wilson, 2012; Cipollina and Salvatici 2010; Haq et al, 2011). The log-linear sample selection is a first hurdle dominance model with the following censoring rule:

$$ln(T_{ij}) = \begin{cases} y^*_{2ij} & if y^*_{1ij} > 0 \\ unobserved & if y^*_{1ij} \leq 0 \end{cases}$$

where y_{1ij}^* is a latent variable for the selection equation, which determines whether or not trade occurs between countries, and y_{2ij}^* is the latent variable for the expected level of trade. This framework appears consistent with the gravity equations derived from monopolistic competition. For example, the model derived by Helpman et al (2008) suggests a single profitability condition, for which exports should occur when the firm's profits in the export market exceed the fixed market entry cost (i.e. $y_{1i}^* > 0$). However, if one considers limits on the primary factors of production, a firm may choose to service only the most profitable markets. The desired export level to some of the markets would be zero, even if the profitability condition is satisfied for these markets.

In addition to arguing the superiority of the sample selection model over PPML, some recent empirical trade studies have estimated both models for comparison. This has produced some unexpected results. Portugal-Perez and Wilson (2012) find that PPML produces a negative effect for preferential trade agreements, and Jayasinghe et al (2010) estimate an elasticity with respect to distance of only -0.08 using PPML, much lower than expected given the results of previous studies. However, since neither of these studies used a model selection test, the choice of sample selection over PPML still remains an open question. Santos Silva et al (2015) appear to be the first to use a non-nested model selection test to compare the sample selection and PPML models when estimating bilateral trade flows. Despite the additional parametrization of the sample selection model, their results strongly support the PPML model.

It is important to note that the log-linear sample selection model is also subject to the heteroskedasticity problem. In fact, Greene (2003) shows that heteroskedastic errors is a property of the two-step estimator of the sample selection model, which has been used in several trade studies (e.g. Olper and Raimondi, 2008; Portugal-Perez and Wilson, 2012; Cipollina and Salvatici 2010). Xiong and Chen (2014) appears to be the only trade study so far that addresses the issue of heteroskedasticity within a limited dependent variable setting. The authors propose a two-step estimator, in the same spirit of the Heckit procedure, but use a Method of Moments (MM) estimator of the exponential functional form in the second step. The MM estimates are consistent under heteroskedasticity given a correct specification for $E\left[T_{ij}|\boldsymbol{x_{ij}}\right]$.

For their MM estimator, Xiong and Chen (2014) use the following orthogonality conditions over positive trade observations:

$$E_p\left[e_{ij}ln\left(\boldsymbol{x_{ij}}\right)\right] = \boldsymbol{0} \quad (7)$$

$$E_p\left[e_{ij}\lambda_{ij}\right] = 0 \quad (8)$$

where $e_{ii} = T_{ij} - exp\left(ln\left(\boldsymbol{x_{ij}}\right)\boldsymbol{\beta}\right) - \omega\lambda_{ij}$, $\lambda_{ij} = \phi\left(\boldsymbol{z_{ij}\alpha}\right)/\Phi\left(\boldsymbol{z_{ij}\alpha}\right)$ is the inverse Mill's ratio, and $\phi(\cdot)$ is the probability density function. When $\omega = 0$, conditions (7) and (8) reduce to a PPML model estimated over positive observations. Using the same dataset as Santos Silva and Tenreyro (2006), Xiong and Chen (2014) estimate a gravity model with importer and exporter fixed effects. Their insignificant inverse Mill's ratio coefficient suggests that selectivity bias is not an issue for this dataset.

Finally, it is worth noting the results of several simulation studies that have attempted to address the question of how to best estimate the gravity equation. Santos Silva and Tenreyro (2006) and (2011) simulated data for a dependent variable using a constant elasticity model. Their results show that PPML works well under various forms of heteroskedasticity. The authors also considered a variation of the Tobit model, which they find generates estimates with substantial bias. However, the Tobit model is misspecified given how the dependent variable was generated, so it is unclear whether the substantial bias is the result of heteroskedastic errors or misspecification of the conditional mean.

A simulation study by Martin and Pham (2008) appears to be the definitive paper for researchers who want to estimate the Heckman sample selection model but not address the PPML model (e.g. Haq et al, 2011; Cipollina and Salvatici 2010, Dutt and Traca, 2010). The results from Martin and Pham (2008) suggest that PPML performs poorly when the fraction of zero trade observations is high. However, this finding may be more to do with incorrect model specification than with the treatment of zeros in the PPML estimator. For generating a dependent variable with zeros, the authors deviate from the PPML specification by subtracting an ad-hoc constant term from the conditional mean[4].

It is worth noting that Martin and Pham's earlier 2007 working paper shows that the Heckman sample selection model performs poorly, and they expressed concern over the popular use of this estimator given its "dismal performance". The authors softened this conclusion in their 2008 paper, in which they conclude that the maximum likelihood estimator, and not the two-step estimator, of the sample selection model performs reasonably well given a valid exclusion restriction. The difference in performance between the two estimators of the sample selection model appears to have been overlooked in the trade literature, as both approaches are commonly used.[5]

---

[4] The simulated dependent variable is left censored at zero.
[5] Olper and Raimondi (2008), Portugal-Perez and Wilson (2012), Cipollina and Salvatici (2010) and Helpman et al (2008) use the two-step procedure, while Haq et al (2011), Jayasinghe et al (2010), Santos Silva et al (2015) and Disdier and Marette (2010) use the maximum likelihood estimator.

Xiong and Chen (2014) conducted several Monte Carlo experiments to demonstrate the consistency of their Heckit MM estimator. Their first two simulations are based on a sample selection data generating process, while the third they claim follows Santos Silva and Tenreyro's (2006) design. However, they add to the third data generating process a censoring point that is independent of the explanatory variables, which was not used in the original study. Santos Silva and Tenreyro (2011) are critical of studies that use an uncorrelated censoring point, and they demonstrate under the conditions of their simulation that PPML works well when the censoring is related to the explanatory variables. Overall, the simulation results of Xiong and Chen (2014) support the MM estimator of the exponential form over the standard Heckit, threshold Tobit and PPML estimators.

Several of the studies mentioned earlier compared coefficients across various models without considering the differences in functional form. Consider, for example, a case where data are generated using a log-linear sample selection model, but one estimates a constant elasticity PPML model with these data. In general, the PPML estimates are not expected to be the same as the parameters in the sample selection model, because the functional forms for $E\left[T_{ij}|\boldsymbol{x_{ij}}\right]$ are different:

PPML: $E\left[T_{ij}|\boldsymbol{x_{ij}}\right] = exp\left(ln\left(\boldsymbol{x_{ij}}\right)\boldsymbol{\beta}\right)$

Log-linear Sample Selection[6]: $E\left[T_{ij}|\boldsymbol{x_{ij}}\right] = exp\left(ln\left(\boldsymbol{x_{ij}}\right)\boldsymbol{\gamma} + \sigma^2/2\right)\Phi\left(\boldsymbol{z_{ij}\alpha} + \rho\sigma\right)$

The marginal effects are clearly different in the above example if x_{ijk} is in both the exponential function and the distribution function $\Phi(\cdot)$. In this particular example,

[6] From Yen and Rosinski (2008).

comparing the estimates $\boldsymbol{\beta}$ to $\boldsymbol{\gamma}$ says nothing about whether PPML produces significantly different results for the estimated effects of explanatory variables.

Overall, recent empirical work and simulation studies provide no clear guidance or consensus on how the gravity equation should be estimated. There are numerous empirical examples where PPML and the sample selection model generate unusual results. The likely causes include incorrect model specification, heteroskedasticity and non-normality. The conclusions drawn from past simulation studies appear to be sensitive to the data generating process, and how the estimates of different models are being interpreted. With the above issues in mind, the next section proposes the use of a BCDH model for estimating the gravity equation.

2.4 The Box-Cox Double Hurdle

Instead of a fixed functional form, a data dependent transformation like the Box-Cox may benefit trade studies that use limited dependent variable models. In addition to providing a better approximation to the conditional mean of trade, data-dependent transformations of the variables can produce a model which better reflects the assumptions made in the statistical analysis, such as homoskedasticity and normally distributed errors (Box and Cox, 1964). Even though Manning (1998) shows that retransformation bias is also an issue for the Box-Cox, transforming both the dependent and explanatory variables should help reduce this bias by stabilizing the variance of the error term. These properties are particularly relevant to trade studies using limited dependent variable models, since in general the consistency of estimates from these models relies on both homoskedasticity and normality.

The idea of using a flexible data transformation on the dependent and explanatory variables has received little consideration for the gravity equation. There appears to be only two notable studies in the gravity trade literature that have considered this approach. Sanso et al (1993) considered the Box-Cox transformation with separate transformation coefficients on the dependent and explanatory variables. Celik and Guldmann (2007) also considered the Box-Cox transformation on both the dependent and explanatory variables, but restricted the model to only one transformation coefficient for all regressors. The log-linear specification commonly used to estimate the gravity model is rejected in both studies, although Sanso et al (1993) conclude that the log-linear specification is a reasonable approximation for their data. Sanso et al (1993) also present encouraging results on heteroskedasticity. They find for their data that the Box-Cox transformation is effective at reducing and even eliminating the heteroskedasticity present in the log-linear model.

This study extends the work using the Box-Cox transformation towards the issue of zero trade through a double hurdle model. Similar to the sample selection model, the double-hurdle model recognizes that zero outcomes can arise by a selection or participation decision. In both models, factors determining participation can be outside of the trade level decision. However, the double hurdle is more general than the sample selection model, in that it allows for the possibility of zeros in the trade level equation. As noted below, the logarithmic transformation on the dependent variable imposes non-zero trade outcomes, so the double-hurdle is not suitable for a log-linear model. However, relaxing this assumption through a Box-Cox transformation opens up the possibility of zero outcomes in the trade level equation.

The BCDH model with dependence between the participation and level equations was recently considered by Jones and Yen (2000) for the analysis of household survey data. Their model is generalized is this study by applying Box-Cox transformations on both the dependent and explanatory variables. The BCDH specification studied in this paper begins with the usual censoring rule for the observed dependent variable y_i, and latent variables y_{1i}^* and y_{2i}^*:

$$y_i = \begin{cases} y_{2i}^* & if\, y_{2i}^* > 0 \; and \; y_{1i}^* > 0 \\ 0 & otherwise \end{cases}$$

In contrast to the sample selection model, zero can be a desired level for the observed dependent variable when the first censoring rule is satisfied (i.e. when $y_{1i}^* > 0$). The assumed relationship between the latent variables and the Box-Cox transformed explanatory variables $\boldsymbol{x_{1i}}$ and $\boldsymbol{x_{2i}}$ is:

$$y_{1i}^* = \boldsymbol{x_{1i}}^{\boldsymbol{T}} \boldsymbol{B_1} + u_{1i} \quad \text{(Participation equation)}$$

$$y_{2i}^* = \boldsymbol{x_{2i}}^{\boldsymbol{T}} \boldsymbol{B_2} + u_{2i} \quad \text{(Trade flow equation)}$$

where $\boldsymbol{B_1}$ and $\boldsymbol{B_2}$ are the estimated parameter vectors, u_{1i} and u_{2i} are bivariate normal error terms, and

$$[u_{1i}, u_{2i}]' \sim BVN(0, \Sigma)$$

$$\Sigma = \begin{bmatrix} 1 & \rho\sigma \\ \rho\sigma & \sigma^2 \end{bmatrix}$$

$$x_{jik}{}^{T} = \frac{x_{jik}{}^{\lambda_k} - 1}{\lambda_k}, \quad \forall\, j \in \{1, 2\}$$

Instead of imposing the same transformation on all variables, the model allows for unique transformation coefficients λ_k for each variable. However, to reduce the number of estimated parameters, the participation and trade flow equations is this study use the same set of transformation coefficients.

Using the Box-Cox transformation on the observed trade flow y_i yields

$$y_i{}^T = \frac{y_i{}^{\lambda_y} - 1}{\lambda_y}$$

for which the latent variable y_{2i}^* equals $y_i{}^T$ over positive trade observations. Combining all of the previous assumptions produces the following likelihood function (L) for the observed trade flows y_i:

$$L = \prod_{i=1}^{N} \left[1 - \Psi\left(\boldsymbol{x_{1i}}^T \boldsymbol{B_1}, \frac{\boldsymbol{x_{2i}}^T \boldsymbol{B_2} + 1/\lambda_y}{\sigma}, \rho\right)\right]^{1-I_i} \times$$

$$\left[\Phi\left(\frac{\boldsymbol{x_{1i}}^T \boldsymbol{B_1} + (\rho/\sigma)(y_i^T - \boldsymbol{x_{2i}}^T \boldsymbol{B_2})}{\sqrt{1-\rho^2}}\right) \frac{y_i^{\lambda_y - 1}}{\sigma} \phi\left(\frac{y_i^T - \boldsymbol{x_{2i}}^T \boldsymbol{B_2}}{\sigma}\right)\right]^{I_i} \qquad (9)$$

where $\Psi(\cdot)$ is the bivariate standard normal cumulative distribution function, $\Phi(\cdot)$ and $\phi(\cdot)$ are the standard normal distribution and density functions, and $I_i = 1$ when $y_i > 0$. Details on the derivation of the likelihood function can be found in Jones and Yen (2000). For the remainder of this paper the acronym BCDH refers specifically to the model defined by equation (9). By imposing the logarithmic transformation on the dependent variable (i.e. $\lambda_y = 0$), equation (9) becomes the likelihood function of the log-transformed sample selection model.

Although the BCDH model nests a number of popular limited dependent variable models with different censoring rules, one should be cautious when interpreting the results of nested model specification tests. The transformation coefficient on the dependent variable is not independent of heteroskedasticity in the error term. In fact, the coefficient is biased in the direction that stabilizes the error variance. This leads to uncertainty on whether a significant transformation coefficient is capturing a better fit for the conditional mean of trade, an error stabilizing effect, or both (Blaylock and Smallwood, 1985). Under heteroskedasticity, a test for $\lambda_y = 0$ cannot be interpreted as a test of the double hurdle censoring rule against the single hurdle dominance rule implied by the sample selection model. However, since the focus of this study is on the overall fit of the model being studied, and not the censoring rules at work, this issue is not of concern.[7]

2.5 Non-nested Model Selection

This study is primarily focused on testing the BCDH specification against the PPML estimator of the gravity equation. Unfortunately, testing these competing specifications presents two challenges. First, a non-nested model selection test must be considered when comparing these two estimators, and second, PPML is quasi-maximum likelihood. Non-nested tests based on a properly specified conditional density function, such as the one proposed by Vuong (1989), are not applicable for PPML. One approach that appears suitable is the P-test proposed by Davidson and MacKinnon (1981). This test nests competing specifications of the conditional expected value within a compound model. The recent paper by Santos Silva et al (2015) considered a heteroskedastic version of the P-test, which was used to evaluate the log-linear sample selection model against PPML.

[7] Blaylock and Smallwood (1985) suggest specifying a functional form for the error variance if functional form testing on the conditional mean is of interest. This approach has been applied to various limited dependent variable models in past research, and would be of interest if substantial heteroskedasticity is suspected for the BCDH model.

The P-test compares two competing specifications for the conditional expectation of a dependent variable y_i:

$$H_o: E[y_i|\boldsymbol{X_i}] = f(\boldsymbol{X_i}, \boldsymbol{B})$$

$$H_a: E[y_i|\boldsymbol{X_i}] = g(\boldsymbol{X_i}, \boldsymbol{A})$$

where $\boldsymbol{X_i}$ is the set of regressors, and $\boldsymbol{A}$ and $\boldsymbol{B}$ are sets of parameter estimates for each specification. The compound model that nests the alternative and null hypotheses is:

$$y_i = (1-a)f(\boldsymbol{X_i}, \boldsymbol{B}) + ag(\boldsymbol{X_i}, \boldsymbol{A}) + u_i \tag{10}$$

where u_i is regression error term. Because the estimates a, $\boldsymbol{A}$ and $\boldsymbol{B}$ are not identifiable in regression model (10), Davidson and MacKinnon (1981) recommend replacing $\boldsymbol{A}$ with the estimates $\widehat{\boldsymbol{A}}$, which are obtained from the model specified under the alternative (i.e. H_a). To simplify the test for non-linear models, the authors recommend linearizing the model at the estimates of $\boldsymbol{B}$ (i.e. $\widehat{\boldsymbol{B}}$) and $a = 0$. This yields the following regression equation:

$$y_i - f\left(\boldsymbol{X_i}, \widehat{\boldsymbol{B}}\right) = \nabla f\left(\boldsymbol{X_i}, \widehat{\boldsymbol{B}}\right)\boldsymbol{b} + a\left(g\left(\boldsymbol{X_i}, \widehat{\boldsymbol{A}}\right) - f\left(\boldsymbol{X_i}, \widehat{\boldsymbol{B}}\right)\right) + u_i \tag{11}$$

where $\boldsymbol{b}$ and a are parameters to be estimated.

For regression model (11) the usual t-statistic can be used to test the null hypothesis $a = 0$ against the alternative of $a = 1$. While a two-tailed test is valid, Santos Silva et al (2015) recommend a one-tailed test in the direction of the alternative hypothesis. Their recommendation is based on the work of Fisher and McAleer (1979), who argue that one should only view the deviation away from the null towards the alternative, as evidence in favour of the model specified under the alternative[8]. Since either of the two competing models can be specified under the null, the test should be conducted for both cases. This leads to the possibility that both models are rejected, or that one fails to reject the null in both cases.

Santos Silva et al (2015) extend the P-test in two directions. First, the authors consider a weighted least squares version of the test, for which the error variance is assumed proportional to the predicted values of the model specified under the null. In other words, they use the skedastic function $f(\boldsymbol{X_i}, \widehat{\boldsymbol{B}})$. The authors argue that this assumption is suitable when the conditional expectation of the dependent variable is heavily clustered near zero. Another approach that might be of value is the two-step FGLS estimator of the skedastic function parameters (Greene, 2003). One could extend the heteroskedastic P-test in this direction by considering a two-step estimator of k, for the skedastic function $f(\boldsymbol{X_i}, \widehat{\boldsymbol{B}})^k$. In addition to the test proposed by Santos Silva et al (2015), this study considers the above two-step estimator starting at an initial value of $k = 1$.

Santos Silva et al (2015) also consider the P-test for various single and double index models, including the PPML estimator and the log transformed sample selection model. The authors conduct a simulation study to investigate the power of this test, and find that it performs well with their heteroskedasticity correction for sample sizes

[8] An interesting example of where a two-tailed test can lead one astray can be found in Greene (2003). Greene compares two non-nested consumption functions. With model 1 under the null he gets α=1.014 and a t-value of 62.86, which gives very strong support for model 2. Greene then places model 1 under the alternative to verify his initial finding, which yields an estimate of α=-10.68 and a t-value of -7.19. Based on the two-tailed test, he rejects both models. However, a *t*-value of -7.19 should not have been interpreted as a deviation away from the null towards the alternative. The one-tailed test gives unambiguous support for model 2, which is not surprising given that Greene's initial estimate of α was very close to 1.

typically encountered in applied work. They recommend conditioning on the parameters of the selection equation (e.g. y_{1i}^{*}) to improve the power of the test when the model under the null can be obtained as a limiting case of the alternative model. To avoid this issue, this study conditions on the Box-Cox transformation parameters when conducting the P-test.[9]

2.6 Revisiting the Log of Gravity: BCDH vs PPML

To better understand the performance of the BCDH against PPML and the sample selection model, this paper evaluates the gravity equations studied by Santos Silva and Tenreyro (2006). Their two models are extensions of the traditional gravity equation (1) and the gravity equation (2) with importer and exporter fixed effects. Santos Silva et al (2015) find evidence to suggest that these gravity models should be estimated using PPML instead of the log-linear sample selection model, when using the dataset studied by Santos Silva and Tenreyro (2006). Thus, it seems fitting to evaluate the BCDH using the same data.

The dataset consists of a single cross-section of aggregate bilateral trade flows among 136 countries, for which zero trade accounts for approximately 48% of the 18,360 observations. Given the rich dataset at hand, the same set of explanatory variables is considered for the participation and trade flow equations. The sample information and nonlinearity of the model should allow for identification of the BCDH parameters without the need of an ad-hoc exclusion restriction in the trade flow equation. The traditional gravity model contains a number of common explanatory variables used in previous

---

[9] When all transformation coefficients equal zero, the BCDH reduces to the log-linear sample selection model, for which $E[y_i|\boldsymbol{x_i}]$ is nearly identical to the conditional expectation of the PPML model.

studies, including importer and exporter GDP, distance, a measure of importer and exporter remoteness, and various dummy variables meant to capture the effects of preferential trade agreements, common language and colonial ties. Full details on the models and data can be found in Santos Silva and Tenreyro (2006).

There are two points to note on specification before discussing the results. First, the Box-Cox transformation on all continuous variables was initially considered. However, the optimization algorithm had difficulties converging near the optimum with transformations on importer and exporter remoteness.[10] Since the estimated transformation coefficients appeared to be converging to zero, the log transformation was imposed on these variables. The final specification includes Box-Cox transformations on bilateral trade, distance, importer and exporter GDP, and importer and exporter per-capita GDP. Secondly, the trade agreement dummy variable is nearly a perfect predictor of positive trade flows, with trade occurring over 93% of observations with an agreement. More importantly, this variable is likely endogenous within the selection equation, and it is likely that in nearly all cases the causality is in the direction of positive trade leading to an agreement. Consequently, the trade agreement dummy variable was dropped from the selection equation, since it is likely not a valid predictor of positive trade.

2.6.1 The Traditional Gravity Equation

Table 2.1 gives the likelihood ratio tests for various restrictions on the BCDH model. For the traditional gravity equation, independence of the error terms is rejected at the 1% significance level. It has been argued that survey data generally lack enough information to identify the correlation between error terms for the double hurdle model

[10] See Appendix A for details on the optimization algorithm.

(Smith, 2003). The results here suggest that this issue may not be present for trade datasets typically used in applied work. Imposing the logarithmic transformation on all continuous regressors is also rejected at the 1% level. The restriction $\lambda_y = 0$ yields a Box-Cox version of the log transformed sample selection model. This restriction is appealing since it simplifies the calculation of the conditional mean and the elasticities. However, the likelihood ratio test also rejects imposing the logarithm of trade. Given the previous two findings, it is not surprising that for the traditional gravity equation the log-linear sample selection model is strongly rejected in favour of the BCDH.

Table 2.1 Nested specification tests for the traditional gravity equation

Restricted Model	Restrictions	# of Restrictions	Likelihood Ratio Test [1]
Independent BCDH	$\rho = 0$	1	297.449 *
Dependent BCDH, log(x)	$\lambda_k = 0, \forall\, k$	5	700.038 *
Log sample selection, $x_{jik}{}^T$	$\lambda_y = 0$	1	1068.569 *
Log-linear sample selection	$\lambda_y = 0$ and $\lambda_k = 0, \forall\, k$	6	1613.936 *

[1] The test statistic is two times the difference of the log-likelihoods. * indicates significance at the 1% level or higher.

Table 2.2 presents the results of the P-test under three different models of heteroskedasticity in equation (9). The null is rejected at the 5% significance level for both models when assuming the variance is constant or proportional to the mean. However, for the values of k estimated using the two-step procedure, the P-test rejects the PPML model in favour of the BCDH. Overall, the test results suggest that for the traditional gravity equation the BCDH is at least as good as, if not better, than the proposed PPML model. It is important to note that these results are specific to both the dataset and specification of the gravity equation studied.

Table 2.2 Non-nested model selection tests for the traditional gravity equation

	P-test [1]		
Model under the Null	$\sigma_i^2 = \sigma^2$	$\sigma_i^2 = f(X_i, \hat{B})$	$\sigma_i^2 = f(X_i, \hat{B})^k$
PPML	1.881 *	3.472 *	8.630 * (*k*=1.85)
BCDH	12.177 *	13.117 *	-1.400 (*k*=1.92)

[1] * indicates significance at the 5% level of higher. σ_i^2 is the variance of the error term for observation *i*. The estimates for *k* are based on two-step weighted least squares.

Table 2.3 Traditional gravity equation estimates

	BCDH [1]		PPML [1,2]
Variable	**Selection**	**Trade Level**	**Trade Level**
Exporter's real GDP	0.179 (0.021) *	1.400 (0.176) *	0.732 (0.027)*
Importer's real GDP	0.034 (0.005) *	0.277 (0.039) *	0.741 (0.027)*
Exporter's per capita real GDP	0.001 (0.001)	0.008 (0.005)	0.157 (0.053)*
Importer's per capita real GDP	0.001 (0.001)	0.007 (0.006)	0.135 (0.045)*
Distance	-0.229 (0.072) *	-1.954 (0.638) *	-0.784 (0.055)*
Log exporter's remoteness	0.380 (0.049) *	2.747 (0.307) *	0.660 (0.134)*
Log importer's remoteness	0.138 (0.050) *	-0.028 (0.275)	0.561 (0.119)*
Contiguity dummy	-0.378 (0.092) *	1.790 (0.507) *	0.193 (0.104)
Common language dummy	0.361 (0.034) *	2.79 0(0.252) *	0.746 (0.135)*
Colonial tie dummy	0.176 (0.036) *	1.513 (0.228) *	0.025 (0.150)
Landlocked exporter dummy	-0.099 (0.035) *	-0.828 (0.213) *	-0.863 (0.157)*
Landlocked importer dummy	-0.217 (0.036) *	-2.541 (0.231) *	-0.696 (0.141)*
Trade agreement dummy	..	2.383 (0.337) *	0.181 (0.089)*
Openness dummy	0.171 (0.028) *	-0.212 (0.169)	-0.107 (0.131)
λ_{trade}	0.132 (0.005) *		
$\lambda_{exporter\ real\ GDP}$	0.043 (0.005) *		
$\lambda_{importer\ real\ GDP}$	0.102 (0.006) *		
$\lambda_{exporter\ per\ capita\ real\ GDP}$	0.582 (0.064) *		
$\lambda_{importer\ per\ capita\ real\ GDP}$	0.544 (0.096) *		
$\lambda_{distance}$	0.110 (0.037) *		

[1] Standard errors appear in parentheses. * indicates significance at the 5% level or higher. Standard errors for PPML were calculated using White's robust covariance matrix (White, 1980)
[2] For PPML, all stochastic variables are in logarithms.

Estimates for the traditional gravity model are given in Table 2.3 for the BCDH and PPML. The BCDH specification, along with conflicting signs in the trade level and selection equations, complicates inference on the effects of explanatory variables. It is better to explore the empirical results through the estimated elasticities and binary effects. Table 2.4 compares the elasticities with respect to the continuous variables for both estimators. Elasticity estimates for the BCDH are decomposed into the elasticities of probability, the elasticities of the expected value conditional on trade occurring (i.e. $E[y_i|y_i > 0]$) and the elasticities of the unconditional level $E[y_i]$. The ability to decompose the elasticities in this manor is one rationale for using limited dependent variable models. The unconditional elasticity estimates are comparable to the PPML estimates, since both are elasticities for the $E[y_i]$. Table 2.4 also includes standard errors of the elasticities, which are approximated using the delta method (Greene, 2003).

Because the gravity equation with fixed effects is a more appropriate specification, the focus here will be on the parameters that are only identified in the traditional model. According to Feenstra et al (2001), the elasticity estimates for importer and exporter GDP provide insights on which theoretical model best characterizes aggregate trade. If the exporter GDP elasticity is greater than the importer GDP elasticity, then the results are consistent with the theoretical predictions of monopolistic competition. Fixed costs and economies of scale under monopolistic competition make it more profitable for businesses to situate within the largest markets due to transportation costs. This creates the so-called home market effect (Hanson and Xiang, 2002). Feenstra et al (2001) show that a reverse home market effect exists under prefect competition with national product differentiation (i.e. the Armington model). The Armington model would be supported if the importer GDP elasticity is greater than the exporter GDP elasticity.

Table 2.4 Elasticities for the traditional gravity equation

	BCDH [1]			
Variable	Probability	Conditional Level	Unconditional Level	PPML
exporter's real GDP	0.761 (0.216) *	0.568 (0.108) *	1.329 (0.129) *	0.732
importer's real GDP	0.679 (0.193) *	0.539 (0.097) *	1.218 (0.117) *	0.741
exporter's per capita real GDP	0.238 (0.071) *	0.157 (0.038) *	0.395 (0.046) *	0.157
importer's per capita real GDP	0.217 (0.067) *	0.059 (0.036)	0.276 (0.041) *	0.135
Distance	-0.858 (0.260) *	-0.757 (0.133) *	-1.615 (0.177) *	-0.784
exporter's remoteness	0.530 (0.206) *	0.332 (0.111) *	0.863 (0.156) *	0.660
importer's remoteness	0.193 (0.117)	-0.161 (0.067) *	0.032 (0.099)	0.561

[1] Standard errors appear in parentheses. * indicates significance at the 5% level or higher.

The GDP elasticities presented in table 2.4 suggest a home market effect when looking at the BCDH estimates, while the PPML results suggest that the elasticities are equal. These findings are investigated more formally using a Wald test[11]. For PPML, the restriction of equal elasticities for importer and exporter GDP is not rejected at a 5% significance level. In contrast, the unconditional importer and exporter GDP elasticities estimated using the BCDH are significantly different at the 5% level, and since the exporter GDP elasticity is larger, the test supports a home market effect for aggregate trade. The BCDH estimates are also significantly larger than PPML, and go against the finding that the trade-to-GDP ratio decreases as total GDP rises (Santos Silva and Tenreyro, 2006). The BCDH results are not entirely against this observation, since the

[11] Technically, a nonlinear version of the Wald test for the BCDH estimates, which is based on a linear approximation of the elasticity function (see Greene, 2003).

conditional GDP elasticities are less than one. Therefore, conditional on having bilateral trade, smaller countries are generally more open to trade.

Baldwin and Harrigan (2011) extended the Feenstra et al (2001) analysis by considering theoretical predictions for the probability of trade. Their analysis shows notable differences in the expected effect of importer GDP across different models. The Eaton and Kortum (2002) comparative advantage model predicts that larger importer size decreases the probability of bilateral trade. In contrast, the heterogeneous firm trade (HTF) model with monopolistic competition developed by Melitz (2003) predicts a positive effect, when the model is extended to asymmetric countries. The results from this study suggest that importer GDP has a positive effect on the probability trade, which supports the extended Melitz (2003) model.[12]

The evidence presented on the effects of GDP appear to support the BCDH over PPML. The positive effect of importer GDP on the probability of trade supports the findings of Baldwin and Harrigan (2011), and suggests that aggregate trade is best characterized as an equilibrium under monopolistic competition. Therefore, one would expect for the unconditional level that the exporter GDP elasticity is greater than the importer GDP elasticity. This is supported by the BCDH results, while the PPML results suggest that the GDP elasticities are equal.

The expected effects of per-capita GDP on aggregate trade appear to be somewhat ambiguous. Under monopolistic competition, Bergstrand (1989) claims that exporter per capita GDP is a proxy for the exporter's capital-labour endowment ratio, which should

[12] Baldwin and Harrigan (2011) also investigate the spatial pattern of export prices. They find that the extended Melitz (2003) model does not correctly predict per unit export values. They propose a product quality HFT model which, according to their empirical work, gives the correct predication for export unit values.

have a positive effect on the trade of capital intensive goods, and a negative effect on the trade of labour intensive goods. In addition, Bergstrand's (1989) model shows that a positive coefficient on per-capita income is an indication that trade is largely in luxuries, while a negative coefficient suggests that trade is largely in necessities. The exporter and importer per capita GDP elasticities estimated using BCDH are both positive, and slightly larger than the PPML estimates. The BCDH results show that the effect of importer per-capita GDP is largely attributable to the impact on the probability of trade.

Both models estimated in this study suggest that exporter remoteness has a positive effect on trade, with the BCDH producing a larger elasticity estimate. The positive effect of exporter remoteness seems counterintuitive. Santos Silva and Tenreyro (2006) argue that more remote countries will have higher than expected bilateral trade flows, because their remoteness reduces the number of available trading partners. The BCDH partially supports this idea through a positive conditional trade elasticity. However, the lack of trading partners would imply a negative effect on the probability of observing trade, which is not supported by the estimates.

While PPML suggests a significant and positive relationship between the expected level of trade and importer remoteness, the BCDH model finds no statistically significant effect. However, the BCDH results suggest that the conditional trade level is negatively affected by importer remoteness. This may reflect additional trade costs associated with shipping goods to remote locations. Finally, it is worth noting that a positive effect on the probability of trade is suggested by the Eaton and Kortum (2002) model. Since the probability of trade does not appear to be affected by importer remoteness, the BCDH results provide further evidence against the Eaton and Kortum (2002) model.

Table 2.5 gives the effects of binary variables on the expected value of trade by simulating a zero to one change for each variable, holding all other explanatory variables at their mean. To better compare the BCDH and PPML estimates, the effect is normalize by the conditional or unconditional trade level to give the relative change in the expected value. The binary effects on the unconditional mean are directly comparable to the estimated binary effects using PPML. The probability effects are the change in the probability of trade for a zero to one change in the dummy variable. Table 2.5 also gives approximate standard errors for the BCDH binary effects, calculated using the delta method (Greene, 2003).

Table 2.5 Effects of binary variables for the traditional gravity equation

	BCDH [1]			
Variable	**Probability**	**Conditional Level [2]**	**Unconditional Level [2]**	**PPML[2]**
Contiguity	-0.090 (0.047)	1.454 (0.330) *	0.371 (0.267)	0.213
Common language	0.110 (0.044) *	0.451 (0.117) *	1.333 (0.216) *	1.109
Colonial tie	0.052 (0.023) *	0.247 (0.075) *	0.582 (0.119) *	0.025
Landlocked exporter	-0.027 (0.015)	-0.110 (0.045) *	-0.227 (0.053) *	-0.578
Landlocked importer	-0.057 (0.028) *	-0.369 (0.043) *	-0.540 (0.041) *	-0.502
Trade agreement	0.000 (0.000)	0.882 (0.170) *	0.882 (0.170) *	0.199
Openness	0.048 (0.021) *	-0.220 (0.043) *	-0.007 (0.062)	-0.101

[1] Standard errors appear in parentheses. * indicates significance at the 5% level or higher.
[2] These estimates are the relative change in the expected value from a 0 to 1 change the binary variable, holding all other explanatory variables are there mean level.

The landlocked dummies are the only variables not identified in the fixed effects model. The PPML estimates for the traditional gravity equation suggest that the effects of being landlocked are fairly symmetric between importers and exporters (-50% vs. -58%, respectively). In contrast, the estimates using the BCDH suggest that being a landlocked importer reduces trade by 54% at the mean, while being a landlocked exporter only reduces trade by 23%. It is unclear what theses asymmetries reflect. Recent empirical studies provide little evidence for or against asymmetric effects, since it appears that most studies use only a single landlocked variable (e.g. Liu, 2009; Portugal-Perez and Wilson, 2012). Hoekman and Nicita (2011) allowed for separate effects, but find using PPML that the negative impact of being a landlocked exporter is substantially greater than the negative impact of being a landlocked importer, which is opposite to the findings from the BCDH model estimated in this study.

2.6.2 The Gravity Equation with Fixed Effects

The second gravity equation estimated in this study has importer and exporter fixed effects added to the traditional model to control for unobserved multilateral resistance terms (MRTs). Because there is only a single cross-section of data, these fixed effects subsume the GDP, remoteness and landlocked variables. In addition, there are two specification issues that need mentioning before discussing the results. First, Santo Silva and Tenreyro (2006) dropped the openness dummy variable in their analysis, even though the matrix of dummy variables including fixed effects has full column rank. Given the interest in identifying possible trade diversion, the openness variable is kept in the model for this study.

The second specification issue relates to the use of fixed effects in probit models. Fixed effects in these models are not identifiable when the corresponding dependent variable takes on the same value across all observations (Greene, 2004). Importer and exporter fixed effects should be dropped from the BCDH selection equation when (1) a country exports to all other countries, (2) a country imports from all other countries, or (3) a country has no exports or no imports. For the dataset used in this study, nine exporter and two importer fixed effects had to be dropped. These exclusions might be rather benign, since the countries impacted are similar in size, representing 9 of the 21 largest nations in the sample[13].

Table 2.6 gives the likelihood ratio tests for various restrictions on the BCDH gravity equation with fixed effects. The results parallel those presented for the traditional gravity equation, with all restricted models strongly rejected. Using the same dataset, Xiong and Chen (2014) find using their sample selection model that selectivity bias is not an issue. This conclusion is also supported by the maximum likelihood estimates of the log-linear sample selection model in this study, since the covariance of the error terms is not significantly different from zero at the 5% level. However, this finding appears to be the result of model misspecification. The log-linear sample selection model is rejected in favour of the BCDH, and independence of the error terms in the BCDH model is also strongly rejected.

Table 2.6 Nested specification tests for the fixed effects gravity equation

Restricted Model	Restrictions	# of Restrictions	Likelihood Ratio Test [1]
Independent BCDH	$\rho = 0$	1	147.391 *
Dependent BCDH, log(x)	$\lambda_k = 0, \forall\, k$	1	29.253 *
Log sample selection, ${x_{jik}}^T$	$\lambda_y = 0$	1	1196.558 *
Log-linear sample selection	$\lambda_y = 0$ and $\lambda_k = 0, \forall\, k$	2	1225.564 *

[1] The test statistic is two times the difference of the log-likelihoods. * indicates significance at the 1% level or higher.

Table 2.7 presents the results of the P-test for the gravity equation with fixed effects. These non-nested model selection tests yield results similar to the traditional

[13] The one required exclusion, which is needed to avoid the linear dependence between the fixed effects and the intercept term, is also the third largest economy in the dataset.

gravity equation. Both PPML and BCDH are rejected under the homoskedastic and variance equal to conditional mean assumptions. For the two-step weighted least squares approach, the null of PPML is rejected at the 5% significance level, while the null of BCDH is not rejected. For this test the estimate of a is also very close to zero when BCDH is under the null. With an estimate of a less than 1e-7 in magnitude, there is essentially no departure from the BCDH model towards PPML. Overall, the test results again show that the BCDH is at least as good as, if not better, than the proposed PPML estimator.

Table 2.7 Non-nested model selection tests for the gravity equation with fixed effects

	P-test [1]		
Model under the Null	$\sigma_i^2 = \sigma^2$	$\sigma_i^2 = f(X_i, \hat{B})$	$\sigma_i^2 = f(X_i, \hat{B})^k$
PPML	4.033 *	9.569 *	12.411 * (*k*=1.77)
BCDH	19.576 *	6.098 *	-1.040 (*k*=1.86)

[1] * indicates significance at the 5% level of higher. σ_i^2 is the variance of the error terms for observation *i*. The estimates for *k* are based on two-step weighted least squares.

Table 2.8 Estimates for the gravity equation with fixed effects

	BCDH [1]		PPML [1,2]
Variable	**Selection**	**Trade Level**	**Trade Level**
Distance	-0.216 (0.055) *	-0.830 (0.211) *	-0.784 (0.042) *
Contiguity dummy	-0.549 (0.118) *	1.247 (0.254) *	0.344 (0.094) *
Common language dummy	0.323 (0.050) *	0.944 (0.123) *	0.378 (0.095) *
Colonial tie dummy	0.326 (0.053) *	1.310 (0.130) *	0.095 (0.135)
Trade agreement dummy	..	1.137 (0.202) *	0.390 (0.077) *
Openness dummy	-0.193 (0.074) *	-0.404 (0.147) *	0.246 (0.117) *
λ_{trade}	0.082 (0.002) *		
$\lambda_{distance}$	0.153 (0.030) *		

[1] Standard errors appear in parentheses. * indicates significance at the 5% level or higher. Standard errors for PPML were calculated using White's robust covariance matrix (White, 1980)
[2] For PPML, distance is in logarithms

Estimates for the gravity model with fixed effects are given in Table 2.8 for both models. All PPML estimates except colonial ties are significant at the 5% level or higher. As noted earlier, it is better to explore the empirical results of the BCDH through estimated elasticities and binary effects. These are given in Table 2.9, along with comparable estimates using PPML. The table also includes the standard errors of the BCDH elasticities and binary effects, which are approximated using the delta method (Greene, 2003).

The BCDH estimates suggest that distance plays a significant role in reducing both the probability and conditional level of trade. Overall, the BCDH produces an elasticity of trade with respect to distance of -1.94, while PPML produces an elasticity half the size. PPML appears to routinely generate distance elasticities less than one (e.g. Hoekman and Nicita, 2011; Liu, 2009; Siliverstovs and Schumacher, 2009). The BCDH estimate is even more striking when considering the meta-analysis conducted Disdier and Head (2008). The authors conducted an extensive review of previous studies and find that 90% of the estimated distance elasticities are between -0.28 and -1.55. The BCDH elasticity estimate is also very close to the estimate produced by Xiong and Chen (2014), although the results here suggest a lesser impact on the likelihood of observing trade compared to their results.

The estimated shared border effect is similar for the two estimators, but the BCDH does produce an unusual negative effect on the probability of observing trade. The BCDH estimates suggest that for two trading nations a shared border increases the average trade flow by 118%. However, a shared border appears to decrease the probability of trade by 22%. Zero trade between neighbouring countries seems unusual. However, the sample has 360 observations with a shared border, of which 37% have zero trade. The log-linear sample selection estimates in this study also show that a shared border negatively affects the probability of trade, while Xiong and Chen's (2014) sample selection estimates, produced using the same data, suggest no significant effect. The main difference in specification is the addition of the openness variable in this study.

Table 2.9 Elasticities and binary effects for the gravity equation with fixed effects

	BCDH [1]			
Variable	**Probability**	**Conditional Level**	**Unconditional Level**	**PPML**
Elasticity				
Distance	-0.506 (0.025) *	-1.436 (0.040) *	-1.942 (0.047) *	-0.784
Binary Effect [2]				
Contiguity	-0.216 (0.046) *	1.181 (0.261) *	0.448 (0.220) *	0.411
Common language	0.116 (0.017) *	0.480 (0.088) *	0.761 (0.111) *	0.460
Colonial tie	0.117 (0.018) *	0.773 (0.110) *	1.109 (0.139) *	0.100
Trade agreement	0.000 (0.000)	0.758 (0.170) *	0.758 (0.170) *	0.476
Openness	-0.072 (0.027) *	-0.141 (0.064) *	-0.233 (0.061) *	0.279

[1] Standard errors appear in parentheses. * indicates significance at the 5% level or higher.
[2] Binary effects on the conditional level, unconditional level, and for PPML represent the relative change in the expected value, from a 0 to 1 change the binary variable, holding all other explanatory variables are there mean level.

The dummy variables for common language and colonial ties attempt to capture the effects of historical connections. PPML estimates suggest that only common language has an impact on trade, with a shared language increasing the mean level by 46%. The BCDH estimates suggest that both factors have a positive effect on trade. According to these estimates, common language increasing the mean trade level by 76%. The effect of colonial ties estimated using BCDH nearly doubles after accounting for MRTs. While

PPML suggests no relationship between the trade level and colonial ties, the BCDH estimates indicate that average trade tends increase by over 100% from having a shared colonial heritage.

The final two binary variables in Table 2.9 deal with the effects of preferential trade agreements (PTAs). The trade agreement dummy equals one if the importer and exporter are part of the same PTA. The openness dummy equals one if either the exporter or importer are part of any PTA. The combination of these two variables gives the net effect of PTAs, while the openness variable on its own captures trade diversion. Both models produce significant estimates for the openness variable after accounting for MRTs. However, the PPML model suggests trade enhancement of 28% at the mean level of trade, while the BCDH suggests trade diversion of 23%.

Once MRTs and trade diversion are considered, the BCDH estimates produce a net increase of 52% in mean trade due to PTAs. In contrast, trade enhancement pushes the estimated net trade effect using PPML to 76%. The contradictory evidence on the effect of PTAs on member/non-member trade is concerning. While the gravity model appears to offer a straightforward approach to estimating potential trade diversion, the findings illustrate that estimates can be sensitive to model specification. The BCDH results also suggest that PTAs are increasing the likelihood of zero trade between members and non-members, while decreasing member/non-member trade flows by 14% on average. It should be noted that neither estimator addresses the likely endogeneity issue for PTAs, so caution is advised when using these estimates.

2.7 Conclusion

Although one empirical example does not provide conclusive evidence to suggest that the BCDH should be used in all gravity trade studies, the results of this particular application are promising. Based on the test results and a comparison of the estimates, future gravity trade analysis should strongly consider the BCDH model. Because the model attempts to correct for heteroskedasticity, non-normality and misspecification of the conditional mean, any gravity study that is interested in modelling the censoring rule should give the BCDH serious consideration. This model might also be well suited for disaggregated trade data with many zeroes, given that the BCDH model nests several different censoring rules, and has a flexible functional form. Therefore, the BCDH model might be particularly useful for the analysis of agricultural trade flows.

CHAPTER 3: ENDOGENOUS PREFERENTIAL TRADE AGREEMENTS IN AGRICULTURAL TRADE MODELS

3.1 Introduction

Despite the long history of using the gravity equation to estimate the effects of preferential trade agreements (PTAs), only recently has concerns about the endogeneity of PTAs been raised (Baier and Bergstrand, 2007; Egger et al, 2011). It has been argued that endogeneity of PTAs is likely caused by omitted variables. Factors that inhibit trade, such as regulatory costs, are typically omitted from the gravity equation. These factors might be positively correlated with the formation of a PTA, leading to a negative correlation between the error term of the gravity equation and the PTA variable. If true, the coefficients on the shared PTA variables could be substantially underestimated (Baier and Bergstrand, 2007).

A likely example of this downward bias can be found in Sun and Reed (2010). Their results show that the North American Free Trade Agreement (NAFTA) has no statistically significant effect on members' agricultural trade. Sun and Reed (2010) do not consider endogeneity bias, and conclude instead that other factors, such as GDP growth, were the primary causes of increased agricultural trade between NAFTA countries. The conclusion that NAFTA does not impact intra-block agricultural trade is also supported by Koo et al (2006). However, the results of these studies for NAFTA conflict with the finds of Jayasinghe and Sarker (2008), and Lambert and McKoy (2009). Given the conflicting evidence and potential for endogeneity bias, it is of interest to further investigate this issue.

This study re-evaluates Sun and Reed (2010) on two fronts. First, their Poisson Pseudo-Maximum Likelihood (PPML) estimator is compared against two competing models, namely the Box-Cox double hurdle (BCDH) and a two-part PPML estimator. Secondly, a control function for a binary endogenous variable is used to investigate possible endogeneity bias. The results suggest that the trade data studied by Sun and Reed (2010) is reasonably modelled using PPML. In addition, the two-step estimator developed by Treza (1998) provides evidence of endogeneity bias in Sun and Reed's (2010) estimates. There is a significant increase in most of the estimated shared PTA effects after including a correction term for endogeneity, especially for NAFTA. The estimates show that NATFA increased member trade by 58% to 105%, from 1996 to 2007. Furthermore, the estimated effects of a shared PTA in 2007 increase from 53% to 235% for ASEAN, 90% to 208% for COMESA, 60% to 140% for the EU15 countries, and 79% to 230% for the EU25 countries[14].

The remainder of this study is organized as follows. Section 3.2 discusses the issues around PTAs and agricultural trade. Section 3.3 highlights some of the most recent empirical evidence on the effects of PTAs using gravity models, while Section 3.4 discusses the potential for endogeneity bias in these estimates. Sections 3.5 to 3.8 provide a re-examination Sun and Reed's (2010) results, with details on Treza's (1998) estimator given in section 3.7. Concluding remarks are given in section 3.9.

[14] This variable captures the impact of EU membership on the countries that were not in the EU15.

3.2 Preferential Trade Agreements and Agricultural Trade

The growth in PTAs through the 2000s has been remarkable. The WTO received notification of 380 negotiated PTAs in 2007, of which 205 were in force (Sun and Reed, 2010). By 2015, these numbers increased to 612 notifications and 406 agreements in force (WTO, 2015). With the number of active agreements doubling from 2007 to 2015, PTAs are likely having a profound effect on world trade. However, it is not fully known how these agreements impact agriculture and agri-food commodities. The effects of these agreements are likely not as pronounced for agricultural trade, since the general trend has been to exclude agriculture from PTA provisions that reduce tariff and non-tariff barriers (Lambert and McKoy, 2009). However, several major agreements such as the European Union (EU), the Caribbean Community and Common Market (CARICOM), the Southern Common Market (MERCOSUR), and NAFTA have extended preferential coverage to agricultural products (Koo et al, 2006).

This expansion of bilateral and regional PTAs has occurred over a backdrop of failed multilateral negotiations. Despite accounting for a small share of total world trade, agriculture has likely played a significant role in deterring multilateral agreements. For example, a number of countries were unwilling to proceed with the Doha Round negotiations until a suitable agreement on agriculture was in place (Grant and Lambert, 2008). Subsequently, agriculture has likely contributed to the proliferation of regional PTAs. With fewer parties, regional PTAs are easier to negotiate, and agricultural protectionism easier to maintain. As noted by Jayasinghe and Sarker (2008), there are divided opinions over the role of these PTAs in moving towards more open multilateral trade. While some view regional PTAs as a step towards a multilateral agreement, others believe these agreements will likely impede multilateralism.

The surge in PTAs will likely have significant negative welfare implications if these agreements obstruct multilateral negotiations. Part of the welfare concern is directed at the general use of non-tariff barriers for all countries, against a backdrop of tariff reduction given to select nations. The use of non-tariff measures has also increased over time, with a greater number of commodities covered, and more countries utilizing them. A positive correlation between non-tariff measures and per-capita income has also been observed, along with a rise in non-tariff measures as a major policy component for restricting global agriculture trade (Hoekman and Nicita (2011).

Agricultural goods continue to face greater tariff and non-tariff barriers compared to manufactured goods, with the difference in restrictiveness being typically greater for higher income countries (Jayasinghe and Sarker, 2008; Hoekman and Nicita, 2011). The side-effect of this dichotomy in trade policy is more restrictive market access for low income countries, because agricultural products comprise a greater share of their total exports. For example, in Africa, and South and Central America, agricultural products account for approximately 29% and 12% of total exports, respectively. Agriculture also accounts for over 50% of national GDP in some African countries (Lambert and McKoy, 2009). The dependence on agriculture in these countries highlights the need to understand the trade creation and diversion effects of PTAs, and their subsequent welfare effects.

There has also been debate among economists over the net trade creation effects of PTAs. The potential for trade diversion is of concern, as these agreements could improve market access for higher cost producers, diverting trade away from more efficient producing regions (Sun and Reed, 2010). The negative economic effects of regional PTAs can be illustrated using an example from NAFTA. Under NAFTA, US textile duties were eliminated for Mexico. Increases in Mexican textile imports coincided with a reduction in U.S. imports from China and India, despite these countries being a lower cost supplier than Mexico (Koo et al, 2006). Subsequently, NAFTA related trade diversion had a negative impact on textile producers in non-member nations. However, there is also a potential for trade enhancement, especially when looking at product categories like agri-food. If the majority of intra-block agricultural products are not substitutes for extra-block products, the PTA may not directly affect agricultural trade with non-members. Furthermore, greater intra-block trade could increase consumer income, resulting in greater demand for products from non-member countries (Koo et al, 2006).

Even the intra-block effects of PTAs on agricultural trade are typically seen as being modest due to the continued protectionism under these agreements. However, as argued by Grant and Boys (2012) for the effect of WTO membership on agriculture trade, the effects of PTAs go beyond observed reductions in tariffs. PTAs can have positive effects on trade through greater certainty for international transactions. More transparent polices, better coordination, and legal recourses for addressing trade distorting practices can encourage greater bilateral trade, despite continued tariff and non-tariff barriers. The effect of PTAs on non-agricultural commodities might also impact relative agricultural production costs. This could lead to trade opportunities for agriculture, even if the PTAs do not include tariff reduction for agricultural products.

3.3 The Estimated Effects of PTAs on Agricultural Trade

The gravity model has remained popular in the analysis of PTAs due its parsimonious specification, and consistent results across numerous applications (Grant and Boys, 2012). The gravity model also appears well suited for studying PTAs compared to other approaches. Computable equilibrium models have been widely used for assessing PTAs, but have been criticised for having parameters with questionable econometric foundations (Hertel et al, 2007), and for being prospective instead of based on actual observation (Krueger, 1999). Analysis using descriptive statistics and measures has also been used. However, it is unlikely that this approach can disentangle the trade creation and diversion effects of PTAs from other major determinants of trade (Jayasinghe and Sarker, 2008). On the other hand, the gravity model has the benefit of being retrospective, with observed data being used to help identify PTA effects from other factors related to bilateral trade.

Not surprisingly, there are several notable studies that used a gravity model to assess the trade impact of PTAs on agriculture commodities. Koo et al (2006) assess various PTAs, with specific focus on the Association of Southeast Asian Nations (ASEAN) agreement, the Andean Community (CAN) currency union, the EU, and NAFTA. Their PTA specification differs from more recent studies, in that they include an additional regressor to capture the average effect of the 131 PTAs accounted for in the study. This was done out of concern that a specific trade diverting dummy (e.g. a NAFTA diversion dummy) might be correlated with an omitted PTA. Overall, their results suggest that PTAs generally increase agricultural trade through trade creation and trade enhancement with non-members. Surprisingly, Koo et al (2006) find that NAFTA has no significant effect on agricultural trade.

The above mentioned finding for NAFTA contrasts with the more recent work by Jayasinghe and Sarker (2008). One of the main differences between the two studies is the use of disaggregated time-series data on agricultural trade in Jayasinghe and Sarker (2008). The authors looked specifically at NAFTA's effect on major agri-food products, using pooled cross-sectional data from 1985 to 2000. Their results show that NAFTA increased intra-block trade greatly for red meat, grains and oilseeds, vegetables, fruits and sugar, while reducing the degree of openness to trade with non-members for all commodities except oilseeds.

Unlike the previously mentioned studies, Lambert and McKoy (2009) allow for asymmetric diversion through separate importer and exporter diversion dummies. Their results show that WTO membership causes diversion for agriculture goods, while CARICOM, the Central American Common Market (CACM), CAN, and the Common Market for Eastern and Southern Africa (COMESA) appear to cause diversion in food trade. Their results also suggest that NAFTA and EU memberships have increased intra and extra-block trade. Lambert and McKoy's (2009) pooled estimates using data from 1995 to 2004 suggest that NAFTA increased intra-block agriculture trade by 145% over this time period.

The studies mentioned previously used traditional log-linear models that are subject to the retransformation issue and selectivity bias discussed in the previous chapter. The research by Sun and Reed (2010) addresses the effects of PTAs while also accounting for zero trade and heteroskedasticity using the PPML estimator. The most reliable estimates produced from this study are likely from the gravity model with fixed country pair effects and time dummies. Sun and Reed (2010) account for diversion in a similar manor as Lambert and McKoy (2009). Their dummy variable specification for PTAs allows for enhancement for member/non-member trade, export diversion, import diversion, or even pure trade diversion (i.e. diversion completely offsetting PTA member trade enhancement). Their study also allows for time-varying PTA effects, to capture the phase-in time of these agreements.

The results of Sun and Reed (2010) indicate that ASEAN, the EU, and the Southern African Development Community (SADC) agreements generate large increases in agricultural trade among members. Their findings also show that the EU generated significant export and import diversion, while SADC increased agricultural trade with non-members over the time period studied. Surprisingly, their results show that NAFTA only generated export diversion for agriculture products. They conclude that other factors, such as GDP growth, were the primary causes of greater trade among NAFTA countries. Overall, Sun and Reed (2010) find limited evidence for multilateral trade promotion, most likely due to continued barriers between PTA members and non-members.

Grant and Boys (2012) also used the PPML estimator in their analysis of WTO membership. Their findings go against the traditional view that the WTO has limited impact on trade in agricultural commodities. Their estimates show that membership increases intra-block agricultural trade by 161% on average, with low to middle income countries benefiting substantially from WTO membership. In addition to their PPML model, Grant and Boys (2012) estimate a Heckman sample selection model, the results of which suggest that shared WTO membership increases the probability of trade by approximately 50%.

The study by Grant and Lambert (2008) is particularly noteworthy, as it appears to be the only agricultural trade study that follows Baier and Bergstrand's (2007) recommendation of using a log-linear gravity model with fixed country pair and time-varying exporter and importer effects. Baier and Bergstrand (2007) recommend this model to address the endogeneity of variables such as PTAs. Overall, Grant and Lambert (2008) find that the PTAs studied had a greater positive effect on trade in agriculture products than in non-agriculture products. According to their results, NAFTA, EU and ASEAN memberships have increased agriculture trade by 137%, 400% and 63%, respectively. The authors also highlight the time-varying nature of PTAs, by showing that over one-third of the effect of NAFTA occurred after a 12 year phase-in period.

3.4 Addressing the Endogeneity of Trade Agreements

The existence of a PTA depends upon the benefits generated from the policy, which in turn are likely related to pre-existing trade relations. PTAs are generally formed between large economies located within the same region. These factors are also linked to the level of trade in the gravity equation (Egger et al, 2011). Countries may also select into PTAs for reasons that are unobserved by the researcher, and possibly correlated with the level of trade, leading to endogeneity of PTAs within the empirical gravity equation. Baier and Bergstrand (2007) argue that these omitted variables will likely result in a downward bias in the estimated trade enhancing effects of PTAs. As discussed in this section, a number of recommendations have been given to address the endogeneity problem, some of which create additional econometric issues.

Addressing endogeneity in the gravity equation starts with accounting for Anderson and van Wincoop's (2003) multilateral resistance terms through pair or nation dummies. These terms are correlated with trade cost factors such as PTAs. If omitted, multilateral resistance will be absorbed into the error term, which in turn means that the error term will be correlated with the trade cost variables. Since multilateral resistance is difficult to measure, the simplest approach is to include importer and exporter fixed effects (i.e. nation dummies). These variables should by time-varying when using panel data, to account for changes in price levels over time. However, as noted by Sun and Reed (2010), time varying nation dummies may preclude identification of parameters of interest, such as the trade diversion effects of PTAs.

Nation dummies alone cannot address the possibility of unobserved factors that simultaneously determine the trade level and PTAs. Baldwin and Taglioni (2006) recommend the use of fixed country pair and time-varying nation dummies to address endogeneity bias for policy variables such as currency unions and PTAs. While this approach should greatly reduce the bias, there are two downsides to the approach highlighted by the authors. First, the impact of time-invariant regressors such as distance cannot be identified. Secondly, the effects of policy variables will only be identified through time variation, so one requires significant time variation in the policy to identify its effect on trade.

It is likely that future agricultural trade studies using the gravity equation will fall short of using Baldwin and Taglioni's (2006) suggestion of country pair dummies and time varying nation dummies. Studies will likely use only one of these two fixed effects specifications, to allow analysis on other determinates of trade. However, the simultaneity of PTAs cannot be overlooked. A more parsimonious approach to addressing endogeneity bias is needed. Several control function (i.e. two-step residual substitution) and instrumental variable (IV) approaches were considered for the gravity equation by Baier and Bergstrand (2004). Their results are not encouraging, with the PTA estimates highly unstable across the various methods. The authors conclude that neither IVs nor control functions adequately addresses the endogeneity problem.

Baier and Bergstrand's (2004) IV results likely reflect the lack of suitable instruments, an issue that also likely affected the similar study conducted by Magee (2003). In contrast, it is unclear whether Baier and Bergstrand's (2004) control function results reflect an issue with the approach. Their reported estimates appear to be based on a traditional gravity equation without controlling for multilateral resistance. Therefore, the lacklustre findings might have resulted from the omitted multilateral resistance terms, which are easily accounted for using importer/exporter dummy variables. The more recent article by Egger et al (2011) shows promise for the control function approach, with the estimated effects of PTAs increasing by 188% after controlling for endogeneity. To date, the control function approach has not been considered for gravity models of agricultural trade.

Only one agricultural trade study mentioned earlier addresses endogeneity of PTAs using fixed country pair and time-varying importer and exporter effects. The other studies mentioned have either overlooked the issue (e.g. Jayasinghe and Sarker, 2008; Lambert and McKoy, 2009; Grant and Boys, 2012), argued that endogeneity is not an issue for agricultural goods (e.g. Koo et al, 2006), or claimed incorrectly that the model specification used in the study addressed the endogeneity problem (e.g. Sun and Reed, 2010). The assumption made by Koo et al (2006) that endogeneity of PTAs is likely not an issue for agricultural goods probably has merit, given that trade agreements are likely formed based on trade in non-agricultural goods. However, the results of Lambert and McKoy (2009) suggest otherwise. Their observation regarding the estimated effects of WTO membership being linked to limited trade potential for non-member countries points to an endogeneity problem. There is likely a set of unobserved factors which simultaneously make these countries less open to both agricultural trade and WTO membership.

3.5 The Econometric Models

This study explores the issue of endogeneity bias for PTA variables in agricultural trade models by re-evaluating the empirical results in Sun and Reed (2010). As noted earlier, this study did not find any statistically significant trade creation effects for NAFTA. The authors claim that a three-way fixed effects model (i.e. country pair and time dummies) adequately addresses the endogeneity problem. While the three-way effects model for panel data is an improvement over other approaches, such as a traditional gravity or fixed importer/exporter effects model, the article by Baldwin and Taglioni (2006) shows that this specification will not fully address the endogeneity problem. Given the statistically insignificant estimate on the NAFTA membership variable in Sun and Reed (2010), it is worth investigating if endogeneity bias is an issue for their results.

Sun and Reed (2010) estimate using PPML the following gravity model with country pair dummies a_{ij} and time dummies time dummies a_t:

$$X_{ijt} = exp\big(a_{ij} + a_t + b_1 lnGDP_{it} + b_2 lnGDP_{jt} + b_3 lnPOP_{it} + b_4 lnPOP_{jt} + \sum_m \sum_t \gamma_{mt}\, D_t PTA^m_{ijt} + \sum_m \sum_t \eta_{mt}\, D_t PTA^m_{jt} + \sum_m \sum_t \omega_{mt}\, D_t PTA^m_{it} + e_{ijt}\big) \quad (1)$$

where X_{ijt} is the dollar value of exports from country *i* to country *j* in time period *t*, GDP is gross domestic product, POP is population, PTA^m_{ij} is a shared PTA dummy for agreement m, PTA^m_j is a dummy variable that equals one when only the importer is part of agreement m, PTA^m_i is a dummy variable that equals one when only the exporter is part of agreement m, and D_t is a dummy variable identifying the year. The set of coefficients on the PTA variables capture a number of possible scenarios for the impact of PTAs, and allows for differences in the diversion (or enhancement) effects across imports and exports.

The general sample consists of bilateral trade flows for 81 countries, for the years 1993, 1996, 1999, 2002, 2005 and 2007. Several observations were dropped due to missing data, leaving 38,037 usable observations. However, country pair fixed effects cannot be estimated using PPML when the dependant variable takes on a value of zero for all years, given that the solution is $a_{ij} = -\infty$.[15] Sun and Reed (2010) dropped these observations from the sample, leaving 31,955 observations for their three-way fixed effects model. The PTAs included in the study are ASEAN, COMESA, the European Union 15 (EU15), the European Union 25 (EU25), NAFTA, and SADC. Further details on the model and data sources can be found in Sun and Reed (2010).

[15] Issues of convergence for PPML, especially when using STATA, have been reported in serval articles (e.g. Liu, 2009; Santos Silva and Tenreyro, 2011; Portugal-Perez and Wilson, 2012). It is worth noting that the convergence issues encountered in this study were eliminated after conditioning out the country pair fixed effects. Computational time was also greatly reduced by using the sparse matrix data type in Gauss.

This empirical study provides an opportunity to apply the model selection work from Chapter 2, to determine if Sun and Reed's (2010) results are related to model misspecification. Using additional variables from Sun and Reed (2010), a BCDH model can be estimated with the trade flow equation

$$X_{ijt}^{T} = a_{ij} + a_t + b_1 GDP_{it}^{T} + b_2 GDP_{jt}^{T} + b_3 POP_{it}^{T} + b_4 GDP_{jt}^{T} + \sum_m \sum_t \gamma_{mt} D_t PTA_{ijt}^{m} + \sum_m \sum_t \eta_{mt} D_t PTA_{jt}^{m} + \sum_m \sum_t \omega_{mt} D_t PTA_{it}^{m} + e_{ijt} \quad (2)$$

and a selection equation based on the latent variable model:

$$y_{1ijt}^{*} = a_0 + a_i + a_j + a_t + z_1 GDP_{it}^{T} + z_2 GDP_{jt}^{T} + z_3 POP_{it}^{T} + z_4 GDP_{jt}^{T} + z_5 lnDist_{ij} + z_6 Border_{ij} + z_7 Colony_{ij} + \sum_m \sum_t \gamma_{mt} D_t PTA_{ijt}^{m} + \sum_m \sum_t \eta_{mt} D_t PTA_{jt}^{m} + u_{ijt} \quad (3)$$

where a_i and a_j are nation dummies (i.e. importer/exporter fixed effects), $Dist_{ij}$ is the distance between countries i and j, $Border_{ij}$ is dummy variable indicating a shared border, and $Colony_{ij}$ is a dummy variable indicating a shared colonial heritage. The superscript T denotes the Box-Cox transformation, for which the transformation is allowed to differ by variable, but not by equation. The shared PTA variables are dropped from the selection equation, because PTAs are likely formed from existing trade relations, instead of agreements leading to positive trade.[16]

The use of nation dummies instead of country pair dummies in the selection equation is due to concerns over the incidental parameter problem for the probit model, which can lead to large biases in the parameter estimates for panels with less than 20 years (Greene, 2004). Because the trade level equation in the double-hurdle follows a Tobit model, it is likely safe to extrapolate from the work of Greene (2004) on the

[16] The data used by Sun and Reed (2010) also has a dummy variable indicating common language. This variable was also not included in the selection equation because it is a near prefect predictor of positive trade.

incidental parameter problem in the Tobit model. Greene (2004) shows that a fixed effects Tobit using panel data with as few as five years has merit, with the bias in the error variance dropping off substantially as the panels increase past five years. It is important to note that the bias in the Tobit model is not just a problem for inference, as the marginal effects are also a function of the variance.

The sample size in Sun and Reed (2010) is likely at the lower limit for using country pair fixed effects in a BCDH model. Because of this, an alternative model using nation dummies is considered. Time-varying nation dummies are preferred, since the multilateral resistance terms should change across time (Baldwin and Taglioni, 2006). However, time-varying effects would subsume the trade diversion variables. As a compromise, the following alternative trade flow equation is considered for the BCDH:

$$X_{ijt}^{T} = a_0 + a_i^{93-99} + a_i^{02-07} + a_j^{93-99} + a_j^{02-07} + \sum_t D_t^* a_{jt} + a_t + b_1 GDP_{it}^{T} + b_2 GDP_{jt}^{T} + b_3 POP_{it}^{T} + b_4 GDP_{jt}^{T} + z_5 Dist_{ij}^{T} + +z_6 Border_{ij} + z_7 Colony_{ij} + z_8 ComLang_{ij} + \sum_m \sum_t \gamma_{mt} D_t PTA_{ijt}^{m} + \sum_m \sum_t \eta_{mt} D_t^* PTA_{jt}^{m} + \sum_m \sum_t \omega_{mt} D_t^* PTA_{it}^{m} + e_{ijt} \tag{4}$$

where $ComLang_{ij}$ is a dummy variable indicating a common language, and D_t^* is a dummy variable indexing the years 2005 and 2007.[17] Nation dummies are specified for the periods 1993 to 1999, and 2002 to 2007, which allows for identification of the trade diversion effects for the two most recent years in the dataset.

As noted in the previous chapter, the use of limited dependent variable models has been motivated in part by concerns over the treatment of zero trade in the PPML model. Some zeroes represent cases far away from possible trade, while other zeroes are near the margin of positive trade between two nations. PPML treats both cases equally. One

[17] The selection equation is (3), but with a Box-Cox transformation on distance.

simple solution to this issue is to dispense with estimating the PPML model over zero trade. Consider the two-part model based on the decomposition

$$E[y] = Prob[y > 0]E[y|y > 0] \tag{5}$$

where the two parts on the right hand side of (5) are independent (Mullahy, 1998). With independence, the researcher is free to choose any binary choice model, such as the probit or logit model, and use any model or estimator for the conditional expectation. Egger et al (2011) were the first to propose this two-part model for the gravity equation, yet no example of this approach can be found in the agricultural trade literature. This study considers a two-part model using a probit model based on (3) with the log transformations imposed, and the trade model (1) estimated over positive observations using PPML.

3.6 Sun and Reed Revisited – Part 1: Model Selection

Before addressing the potential endogeneity bias in the Sun and Reed (2010) results, its worth testing the various models discussed earlier, to try and determine which one best fits the data. Table 3.1 presents the results of the P-test for non-nested model selection, including the Santos Silva et al (2015) version with $\sigma_i^2 = f(X_i, \hat{B})$, and the WLS version proposed in the previous chapter. All models with country pairs preform reasonably well based on the R^2 (i.e. the square of the product-moment correlation coefficient for the actual and predicted trade flows). The WLS P-test is unable to identify a preferred model when comparing both the country pair BCDH and the two-part model to PPML. Both the WLS P-test and the relatively low R^2 value show that the nation dummy BCDH model does not perform nearly as well as PPML. Overall, the PPML model used by Sun and Reed (2010) appears to be the best model considered for this particular dataset.

Table 3.1 Non-nested model selection tests for the Sun and Reed trade data

Model under the Null[2]	R^2	P-test[1] $\sigma_i^2 = \sigma^2$	$\sigma_i^2 = f(X_i, \hat{B})$	$\sigma_i^2 = f(X_i, \hat{B})^k$
PPML vs C. Pair BCDH				
Country Pair BCDH	0.946	5.869 *	10.567 *	-0.375 (*k*=1.552)
PPML	0.969	-3.310	1.559	-3.624 (*k*=1.643)
PPML vs N.D. BCDH				
Nation Dummy BCDH	0.557	64.664 *	48.184 *	10.576 (*k*=1.243) *
PPML	0.969	1.711 *	4.638 *	1.482 (*k*=1.546)
PPML vs 2P-PPML				
Two-part PPML	0.964	-12.303	-8.437	-6.577 (*k*=1.704)
PPML	0.969	-7.072	1.447	0.087 (*k*=1.642)

[1] * indicates significance at the 5% level of higher. σ_i^2 is the variance of the error term for observation *i*. The estimates for *k* are based on two-step weighted least squares.
[2] The PPML models include country pair fixed effects

The model selection results also suggest that the WLS P-test is an improvement over the Santos Silva et al (2015) version, which assumes the variance is proportional to the mean. The R^2 values for the country pair fixed effects BCDH and PPML models are very close, suggesting that both models provide a comparably strong fit to the observed trade flows. However, only the WLS P-test is in agreement with this conclusion drawn from the R^2 values. The test results for the two-part PPML model also suggest a similar issue. The R^2 values for the standard PPML and two-part PPML models are nearly identical, yet the PPML model is almost rejected at the 5% level when the variance is assumed proportional to the mean. Based on the R^2 values, the PPML model also significantly outperforms the BCDH with nation dummies (0.969 vs. 0.557). However, only the WLS P-test fails to reject the null of PPML when the BCDH with nation dummies is specified under the alternative. Overall, the WLS procedure appears to provide more reliable test results compared to the other versions of the P-test.

3.7 A Two-step Model for Endogenous Trade Agreements

While the PPML model used by Sun and Reed (2010) appears to fit the data well, it is still unclear if the PTA estimates are impacted by endogeneity bias. The two-step estimator for endogenous treatment effects proposed by Treza (1998) appears to be well suited for addressing this issue. Treza (1998) develops a two-step estimator for an exponential model, which can easily be estimated using pseudo maximum likelihood under a variance equal to conditional mean assumption. This yields an estimator similar to PPML, which includes a control function that can correct for endogeneity bias. A related approach developed by Smith and Blundell (1986) for the Tobit model could be adapted for the BCDH. However, the incidental parameter problem is of concern given the short panels available. Despite to comparable fit of the BCDH model to PPML, only the control function approach for PPML is considered in this study.

The problem at hand is one of self-selected treatment, where countries enter into agreements for reasons that are also related to the level of trade. Suppose we can model the agreement decision as

$$PTA_i^* = \boldsymbol{w_i z} + e_i \tag{6}$$

$$PTA_i = \begin{cases} 1 \;\; if\ PTA_i^* > 0 \\ 0 \;\; otherwise \end{cases}$$

where $\boldsymbol{w_i}$ is a set of regressors that influence the likelihood of a PTA, and $\boldsymbol{z}$ is vector of parameter estimates. If the model of primary interest (e.g. a gravity model) is an exponential function, Treza (1998) shows that with normally distributed errors a two-step estimator can be formed based on (6) and the following second step regression equation[18]:

$$y_i = exp\left(\boldsymbol{x_i B} + \delta PTA_i + ln(\lambda_i)\right) + e_i \tag{7}$$

$$\lambda_i = PTA_i \frac{\Phi(\theta + \boldsymbol{w_i z})}{\Phi(\boldsymbol{w_i z})} + (1 - PTA_i) \frac{1 - \Phi(\boldsymbol{\theta} + \boldsymbol{w_i z})}{1 - \Phi(\boldsymbol{w_i z})}$$

where $\Phi(\cdot)$ is the standard normal distribution function. The usual t-statistic for the estimated parameter θ provides a test of endogeneity for PTA_i.

Santos Silva and Tenreyro (2006) are critical of nonlinear least squares (NLS) estimators of exponential models like (7), because NLS gives more weight to observations with larger values for the dependent variable. They recommend an estimator that gives equal weight to all observations. For equation (7) this amounts to the following first order conditions:

$$\sum \begin{pmatrix} \boldsymbol{x_i} \\ PTA_i \\ \partial ln(\lambda_i)/\partial\theta \end{pmatrix} \left(y_i - exp\left(\boldsymbol{x_i B} + \delta PTA_i + ln(\lambda_i)\right)\right) = \mathbf{0} \tag{8}$$

A pseudo maximum likelihood estimator can be derived from the conditions (8), which reduces to PPML when $\theta = 0$. Because the log-likelihood is both misspecified and conditioned on $\boldsymbol{w_i z}$, inference should be based on a robust Murphy-Topel covariance matrix (Hardin, 2002).

The estimator outlined above attempts to address endogenous treatment effects through the control function λ_i. Egger et al (2011) propose the same estimator as (7), but use a bivariate probit for the probability of trade and the probability of a PTA. Similar to

[18] It should be noted that the conditional mean in an exponential sample selection model is not the function plus an inverse Mill's ratio, as was suggested by Xiong and Chen (2014). The exponential sample selection model would be similar to equation (7), but with only the first term on the right hand side of λ_i (see e.g. Yen and Rosinski (2008), or Treza (1998)).

the Heckman two-step estimator, the nonlinearity of λ_i can allow for identification of the parameters in (7), even if the same regressors are in $\boldsymbol{x_i}$ and $\boldsymbol{w_i}$.[19] This offers a great advantage over an instrumental variable (IV) approach, since it would be difficult to identify an instrument that is correlated with the probability of a PTA, but not the error term in the trade flow equation (Baier and Bergstrand, 2007).

When one is trying to estimate the effects of different PTAs, it might be more practical to aggregate all PTAs into a single dependent variable in (6), given that some PTAs might have very few observations (e.g. NAFTA). To estimate a PPML model with a correction for endogeneity, the following probit model is estimated in this study:

$$\begin{aligned} PTA_{ijt}^{*} = a_0 + a_i + a_j + a_t + z_1 lnGDP_{it} + z_2 lnGDP_{jt} \\ + z_3 lnPOP_{it} + z_4 lnPOP_{jt} + z_5 lnDist_{ij} \\ + z_6 Border_{ij} + z_7 ComLang_{ij} + z_8 Colony_{ij} + u_{ijt} \end{aligned}$$

where PTA_{ijt}^{*} is a latent variable that measures the net gains of a preferential trade agreement, and

$$Prob\left(\textstyle\sum_m PTA_{ijt}^{m} > 0\right) = \Phi\begin{pmatrix} a_0 + a_i + a_j + a_t + z_1 lnGDP_{it} + z_2 lnGDP_{jt} \\ + z_3 lnPOP_{it} + z_4 lnPOP_{jt} + z_5 lnDist_{ij} \\ + z_6 Border_{ij} + z_7 ComLang_{ij} + z_8 Colony_{ij} \end{pmatrix} \qquad (9)$$

3.8 Sun and Reed Revisited – Part 2: PPML with Endogenous PTAs

Table 3.2 presents the results of the two probit models estimated in this study. For brevity, the nation and time dummies are omitted from the table. As expected, less distance increases the likelihood of a shared PTA. The estimated negative effects for a shared border and GDP are surprising. Results from Egger et al (2011) also show that a

[19] A similar argument was given by Xiong and Chen (2014) for the exponential functional form in their sample selection model. For an example where nonlinearities in controls achieves identification without exclusions see Klein and Vella (2010)

shared border decreases the likelihood of a PTA. However, the estimates for GDP from this study conflict with the findings of Baier and Bergstrand (2004), which show that larger economies are more likely to have a PTA. This discrepancy can be reconciled if the model is arranged as a function of GDP and per-capita GDP. Rewriting the equation this way reveals a positive effect for GDP and a negative effect for per-capita GDP. Therefore, the results do suggest that larger economies are generally more open to PTAs, but that a larger per-capita income tends to reduce the likelihood of forming an agreement.

The results for the likelihood of trade suggest that greater importer GDP and exporter population increase the likelihood of agricultural trade, while greater distance and a shared border appear to decrease the probability of trade. The negative effect for a shared border seems counterintuitive, and conflicts with the results of the agri-food trade study by Haq et al (2011). However, the findings in the previous chapter also suggest that a shared border decreases the likelihood of trade. The results also show that the EU15 agreement increases the likelihood of trade with non-members. This positive effect on observing non-member agricultural trade is fairly consistent across time, and across imports and exports. The remaining PTA coefficients are generally insignificant, and suggest no stable effect of these PTAs on the likelihood of extra-block agriculture trade over the sample period. However, there is some evidence to suggest that COMESA has reduced the likelihood of agriculture imports from non-members, and that the countries that signed into the EU25 agreement are now less likely to trade in agriculture products with other countries.

Table 3.2 Probit Model Estimates for the Probability of Trade and PTAs

Variable	PTA	Trade	Variable	Trade
$lnGDP_{it}$	-0.444 (0.055) *	0.058 (0.052)	$EU15_{i,2007}$	0.623 (0.169) *
$lnGDP_{jt}$	-0.444 (0.055) *	0.199 (0.053) *	$EU15_{j,1996}$	0.234 (0.096) *
$lnPOP_{it}$	0.471 (0.055) *	0.244 (0.122) *	$EU15_{j,1999}$	0.532 (0.114) *
$lnPOP_{jt}$	0.471 (0.055) *	-0.085 (0.062)	$EU15_{j,2002}$	0.466 (0.115) *
$lnDist_{ij}$	-1.470 (0.044) *	-0.589 (0.019) *	$EU15_{j,2005}$	0.503 (0.188) *
$Border_{ij}$	-0.894 (0.113) *	-0.328 (0.100) *	$EU15_{j,2007}$	0.556 (0.198) *
$ComLang_{ij}$	0.359 (0.069) *	0.365 (0.032) *	$EU25_{i,2005}$	-0.297 (0.132) *
$Colony_{ij}$	-1.331 (0.308) *		$EU25_{i,2007}$	-0.229 (0.132)
$ASEAN_{i,2005}$		-0.041 (0.128)	$EU25_{j,2005}$	-0.371 (0.161) *
$ASEAN_{i,2007}$		-0.087 (0.121)	$EU25_{j,2007}$	-0.288 (0.168)
$ASEAN_{j,2005}$		-0.066 (0.107)	$NAFTA_{i,1996}$	0.264 (0.205)
$ASEAN_{j,2007}$		0.020 (0.113)	$NAFTA_{i,1999}$	0.173 (0.208)
$COMESA_{i,1996}$		0.021 (0.081)	$NAFTA_{i,2002}$	0.003 (0.206)
$COMESA_{i,1999}$		-0.120 (0.087)	$NAFTA_{i,2005}$	0.556 (0.494)
$COMESA_{i,2002}$		-0.066 (0.092)	$NAFTA_{i,2007}$	0.286 (0.285)
$COMESA_{i,2005}$		-0.147 (0.098)	$NAFTA_{j,1996}$	0.052 (0.198)
$COMESA_{i,2007}$		-0.176 (0.102)	$NAFTA_{j,1999}$	-0.064 (0.203)
$COMESA_{j,1996}$		-0.141 (0.083)	$NAFTA_{j,2002}$	-0.060 (0.208)
$COMESA_{j,1999}$		-0.271 (0.088) *	$NAFTA_{j,2005}$	0.343 (0.277)
$COMESA_{j,2002}$		-0.257 (0.093) *	$NAFTA_{j,2007}$	0.179 (0.250)
$COMESA_{j,2005}$		-0.192 (0.101)	$SADC_{i,2002}$	-0.029 (0.107)
$COMESA_{j,2007}$		-0.105 (0.101)	$SADC_{i,2005}$	-0.052 (0.109)
$EU15_{i,1996}$		0.303 (0.098) *	$SADC_{i,2007}$	0.356 (0.125) *
$EU15_{i,1999}$		0.697 (0.114) *	$SADC_{j,2002}$	-0.146 (0.112)
$EU15_{i,2002}$		0.675 (0.114) *	$SADC_{j,2005}$	-0.013 (0.118)
$EU15_{i,2005}$		0.627 (0.166) *	$SADC_{j,2007}$	-0.007 (0.124)

[1] Standard errors in parentheses are calculated using White's robust covariance matrix (White, 1980).
* indicates significance at the 5% level or higher.

The main results of this study are presented in Table 3.3, which gives the estimated shared PTA effects on agricultural trade using Sun and Reed's (2010) three-way fixed effects model before and after addressing endogeneity.[20] The first set of estimates are from the standard PPML model, which match the estimates in Sun and Reed (2010) under the exogenous PTA assumption[21]. The relatively small and insignificant coefficients on the NAFTA dummies suggest that the agreement had no effect on agricultural trade over the 1996-2007 period. Sun and Reed (2010) conclude that GDP growth and natural trade relations between these countries were responsible for the growth in trade. Other unusual results in Table 3.3 include the loss of significance for the SADC variable in 2007, and although insignificant, the negative estimates for COMESA in 1999 and 2002.

The second set of estimates in Table 3.3 give the shared PTA estimates after adding the control function for endogenous trade agreements. The estimate for θ is statistically significant at the 5% level, and surprisingly close to the estimate produced by Egger et al (2011) (-0.371 vs. -0.335). A negative θ suggests that unobserved factors with a positive effect on PTA creation are generally associated with unobserved factors that have a negative impact on trade. Therefore, the unobserved factors in the gravity equation are generating a downward bias in the PTA estimates, which supports Baier and Bergstrand's (2007) hypothesis.

[20] The estimated effects for the diversion variables using PPML do not change substantially after adding the endogeneity correction term. Results are available from the author upon request.

[21] However, the standard errors could not be reproduced. It is possible that Sun and Reed's (2010) estimator did not fully converge. The STATA procedure used by the authors appears to base convergence on the change in parameter estimates. Analysis of the estimation output in this study shows that the change in parameters slows considerably near the optimum, even though the gradient vector is not within the tolerance level. This study bases convergence on the sum of squared gradients being less than 1e-5.

Table 3.3 Estimated Effects of PTAs on Agriculture Trade

Variable	PPML: Exogenous PTAs [1]	PPML: Endogenous PTAs [2]	Two-part PPML: Exogenous PTAs [1]	Two-part PPML: Endogenous PTAs [2]
$ASEAN_{ij,2005}$	0.183 (0.091) *	0.938 (0.362) *	0.240 (0.090) *	-0.013 (0.683)
$ASEAN_{ij,2007}$	0.427 (0.105) *	1.209 (0.377) *	0.477 (0.102) *	0.214 (0.713)
$COMESA_{ij,1996}$	0.328 (0.367)	0.966 (0.469) *	0.002 (0.343)	-0.241 (0.765)
$COMESA_{ij,1999}$	-0.120 (0.343)	0.403 (0.412)	-0.519 (0.300)	-0.728 (0.669)
$COMESA_{ij,2002}$	-0.330 (0.322)	0.116 (0.380)	-0.697 (0.283) *	-0.878 (0.594)
$COMESA_{ij,2005}$	0.259 (0.283)	0.709 (0.338) *	-0.151 (0.246)	-0.339 (0.596)
$COMESA_{ij,2007}$	0.643 (0.279) *	1.126 (0.344) *	0.238 (0.241)	0.037 (0.625)
$EU15_{ij,1996}$	0.302 (0.116) *	0.948 (0.325) *	0.323 (0.107) *	0.101 (0.610)
$EU15_{ij,1999}$	0.394 (0.104) *	0.868 (0.228) *	0.524 (0.099) *	0.359 (0.458)
$EU15_{ij,2002}$	0.471 (0.102) *	0.913 (0.212) *	0.605 (0.097) *	0.450 (0.432)
$EU15_{ij,2005}$	0.571 (0.106) *	0.974 (0.193) *	0.718 (0.101) *	0.571 (0.414)
$EU15_{ij,2007}$	0.472 (0.134) *	0.874 (0.207) *	0.621 (0.129) *	0.474 (0.423)
$EU25_{ij,2005}$	0.457 (0.085) *	1.061 (0.276) *	0.559 (0.081) *	0.349 (0.571)
$EU25_{ij,2007}$	0.581 (0.094) *	1.196 (0.281) *	0.684 (0.090) *	0.470 (0.587)
$NAFTA_{ij,1996}$	0.050 (0.081)	0.717 (0.335) *	0.067 (0.058)	-0.157 (0.602)
$NAFTA_{ij,1999}$	0.001 (0.066)	0.530 (0.252) *	0.156 (0.053) *	-0.021 (0.476)
$NAFTA_{ij,2002}$	0.067 (0.069)	0.589 (0.248) *	0.215 (0.057) *	0.039 (0.474)
$NAFTA_{ij,2005}$	0.029 (0.068)	0.457 (0.206) *	0.188 (0.054) *	0.040 (0.396)
$NAFTA_{ij,2007}$	0.098 (0.069)	0.505 (0.196) *	0.258 (0.055) *	0.117 (0.379)
$SADC_{ij,2002}$	1.175 (0.375) *	1.481 (0.418) *	0.730 (0.361) *	0.635 (0.429)
$SADC_{ij,2005}$	1.048 (0.342) *	1.243 (0.357) *	0.645 (0.338)	0.584 (0.368)
$SADC_{ij,2007}$	0.561 (0.351)	0.750 (0.371) *	0.075 (0.403)	0.012 (0.425)
θ	--	-0.335 (0.150) *	--	0.120 (0.330)

[1] Standard errors for the exogenous PTA model are calculated using White's robust covariance matrix (White, 1980). * indicates significance at the 5% level or higher.
[2] Standard errors for the endogenous PTA model are calculated Hardin's (2002) robust Murphy-Topel covariance matrix. * indicates significance at the 5% level or higher.

Overall, there is a significant increase in most of the estimated shared PTA effects after addressing the endogeneity issue, especially for NAFTA. The estimates show that NATFA increased member trade by 58% to 105% over the time period studied. However, these estimates for NAFTA fall short of the 145% and 137% increases estimated by Lambert and McKoy (2009) and Grant and Lambert (2008), respectively. After addressing the endogeneity issue, the estimated impacts of other PTAs in 2007 increase from 53% to 235% for ASEAN, 90% to 208% for COMESA, 60% to 140% for the EU15 countries, and 79% to 230% for the EU25 countries[22]. The effect of ASEAN on agriculture trade is notably higher than the estimate produced by Grant and Lambert (2008). The negative coefficients for COMESA also turn positive, the COMESA estimates for 1996 and 2005 become significant, and the SADC estimate for 2007 becomes significant. It is also worth noting that the estimated effects on trade between PTA members appears more stable across time after adding the control function.

The final two sets of estimates are for the two-part PPML model. The most striking result for the two-part model is that endogeneity bias does not appear to be present. The estimate for θ is highly insignificant, and the correction term appears to inflate the standard errors of the shared PTA variables. These results are concerning, give that an exclusion restriction was not used. It is likely that identification issues are present for the estimates obtained when estimating the model over only positive observations. If correct, it appears that identification of the gravity model parameters and the endogeneity correction term was aided by estimating the PPML model over the zero observations. This aspect of the analysis has not been identified in any previous trade study.

3.9 Conclusion

Overall, three important conclusions can be drawn from this study. First, the results cast doubt on the assumption that PTAs are exogenous in agricultural trade models. Secondly, Sun and Reed's (2010) three-way fixed effect model does not address

[22] This dummy variable captures the impact of EU membership on the countries that were not in the EU15.

the endogeneity bias, and there appears to be substantial downward bias in their estimates. The results of this study challenge their conclusion that NAFTA had no effect on intra-block agricultural trade. Finally, the results show that endogeneity bias in agricultural trade models can be substantial. The fact that most agricultural trade studies using the gravity model have not addressed the endogeneity of PTAs is of concern. The findings of this study suggest that future research on agriculture trade should investigate ways to address this endogeneity bias.

CHAPTER 4: CONCLUSION

Interest in the econometrics of the gravity model will likely remain high for the foreseeable future, as the model is regarded as a useful analytical tool for studying various factors related to international trade. The research results presented in this thesis support future analysis, by addressing several important econometric issues, namely the handling of zero trade flows, heteroskedasticity, and the endogeneity of preferential trade agreements (PTAs). This thesis also explores a WLS version of the Davidson and MacKinnon (1981) model selection test, which appears to be an improvement over the original version and the modification suggested by Santos Silva et al (2015). Many of the econometric methods utilized in this research appear to be new to the area of agricultural trade analysis. Consequently, this research should contribute significantly to future study on agricultural trade flows.

In both empirical studies the BCDH appears to fit the trade data as well as the popular PPML model. Based on a review of the literature, no other proposed limited dependent variable model has been shown to perform as well as PPML when applied to real trade data. Unfortunately, the estimated effects of explanatory variables can differ substantially across the two models. Despite the evidence favouring BCDH over PPML, it is still unclear if the estimated effects from the BCDH are better. Overall, it appears that model specification remains an important consideration for future gravity trade analysis.

There are two notable extensions to the BCDH model worth considering in future study. First, there are alternative data-dependent transformations that could be used, including several versions of the Box-Cox transformation. The transformation used in this study was originally considered because it nested the log-linear sample selection model.

Secondly, if heteroskedasticity is suspected in the BCDH, one possible solution would be to include a skedastic function in the log-likelihood. This approach could address heteroskedasticity, and assuming that normality holds, the log transformation of the dependent variable might be a valid restriction.

The results in Chapter 3 show that PTAs can introduce substantial endogeneity bias in agricultural trade models. Unfortunately, nearly all agricultural trade studies using the gravity model have not explored this issue. The evidence produced in this study suggests that the estimated effects of PTAs found in recent papers could be substantially biased. It appears that there is a need to revaluate most of the current gravity equation estimates on the effects of PTAs for agricultural trade, in order to better inform policy discussions. Future research on agriculture trade should also investigate alternative ways to address endogeneity bias. A comparison of the control function and the country pair with time varying nation dummies approach might be of interest, results of which could help identify the best method for correcting the endogeneity of trade agreements.

BIBLIOGRAPHY

Anderson, J. E. and E. van Wincoop. 2003. Gravity with gravitas: a solution to the border puzzle. *The American Economic Review* 93(1): 170-192.

Baldwin, R. and J. Harrigan. 2011. Zeros, quality, and space: trade theory and trade evidence. *American Economic Journal: Microeconomics* 3(2): 60-88.

Baldwin, R. and D. Taglioni. 2006. Gravity for dummies and dummies for gravity equations. National Bureau of Economic Research. Working Paper No. 12516.

Baier, S. L. and J. H. Bergstrand. 2007. Do free trade agreements actually increase members' international trade? *Journal of International Economics* 71(1): 72-95.

Baier, S. L. and J. H. Bergstrand. 2004. Do free trade agreements actually increase members' international trade? Working paper available from: http://www.nd.edu/~jbergstr/Working_Papers/BaierBergstrandFTA2Oct2004.pdf.

Blaylock, J. R., and D. M. Smallwood. 1985. Box-Cox transformations and a heteroscedastic error variance: import demand equations revisited. *International Statistical Review* 53(1): 91-97.

Bergstrand, J. H. 1989. The generalized gravity equation, monopolistic competition, and the factor-proportions theory in international trade. *The review of economics and statistics* 71(1): 143-153.

Box, G. E. P. and D. R. Cox. 1964. An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)* 26(2): 211-252.

Celik, H. M. and J. M. Guldmann. 2007. Spatial interaction modeling of interregional commodity flows. *Socio-economic Planning Sciences* 41(2): 147-162.

Cipollina, M. and L. Salvatici. 2010. The trade impact of European Union agricultural preferences. *Journal of Economic Policy Reform* 13(1): 87-106.

Davidson, R. and J. G. MacKinnon 1981. Several tests for model specification in the presence of alternative hypotheses. *Econometrica* 49(3), 781-793.

Disdier, A. C. and K. Head. 2008. The puzzling persistence of the distance effect on bilateral trade. *The Review of Economics and Statistics* 90(1): 37-48.

Disdier, A. C. and S. Marette. 2010. The combination of gravity and welfare approaches for evaluating nontariff measures. *American Journal of Agricultural Economics* 92(3): 713-726.

Dutt, P. and D. Traca. 2010. Corruption and bilateral trade flows: extortion or evasion? *The Review of Economics and Statistics* 92(4): 843-860.
Eaton, J. and S. Kortum. 2002. Technology, geography, and trade. *Econometrica* 70(5): 1741-1779.

Egger, P., M. Larch, K. E. Staub and R. Winkelmann. 2011. The trade effects of endogenous preferential trade agreements. *American Economic Journal: Economic Policy* 3(3): 113-143

Feenstra, R. C., J. R. Markusen and A. K. Rose. 2001. Using the gravity equation to differentiate among alternative theories of trade. *Canadian Journal of Economics* 34(2): 430-447.

Felbermayr, G. and W. Kohler. 2010. Modelling the extensive margin of world trade: new evidence on GATT and WTO membership. *The World Economy* 33(11): 1430-1469.

Fisher, G. and M. McAleer. 1979. On the interpretation of the Cox test in econometrics. *Economics Letters* 4(2): 145-150.

Gourieroux, C., A. Monfort and A. Trognon. 1984. Pseudo maximum likelihood methods: applications to poisson models. *Econometrica* 52(3): 701-720.

Grant, J. H. and K. A. Boys. 2011. Agricultural trade and the GATT/WTO: does membership make a difference? *American Journal of Agricultural Economics* 94(1): 1-24.

Grant, J. H. and D. M. Lambert. 2008. Do regional trade agreements increase members' agricultural trade? *American Journal of Agricultural Economics* 90(3): 765-782.

Greene, W. 2004. Fixed effects and bias due to the incidental parameters problem in the Tobit model. *Econometric Reviews* 23(2): 125-147.

Greene, W. H. 2003. *Econometric Analysis* Prentice Hall, New Jersey.

Hanson, G. H. and C. Xiang. 2002. The home market effect and bilateral trade patterns. National Bureau of Economic Research. Working Paper No. 9076

Haq Z., K. Meilke and J. Cranfield. 2011. The gravity model and the problem of zeros in agrifood trade. Canadian Agricultural Trade Policy Research Network. Working Paper 2011-02.

Hardin, J. W. 2002. The robust variance estimator for two-stage models. *Stata Journal* 2(3): 253-266.

Heckman, J. J. 1979. Sample selection bias as a specification error. *Econometrica* 47(1): 153-161.

Hertel, T., D. Hummels, M. Ivanic and R. Keeney. 2007. How confident can we be of CGE-based assessments of free trade agreements? *Economic Modelling* 24(4): 611-635.

Hoekman, B. and A. Nicita. 2011. Trade policy, trade costs, and developing country trade. *World Development* 39(12): 2069-2079.

Jayasinghe, S., J. C. Beghin, and G. Moschini. 2010. Determinants of world demand for US corn seeds: the role of trade costs. *American Journal of Agricultural Economics* 92(4): 999-1010.

Jayasinghe, S. and R. Sarker. 2008. Effects of regional trade agreements on trade in agrifood products: evidence from gravity modeling using disaggregated data. *Applied Economic Perspectives and Policy* 30(1): 61-81.

Jones, A. M. and S. T. Yen. 2000. A Box-Cox double-hurdle model. *The Manchester School* 68(2): 203-221.

Klein, R. and F. Vella. 2010. Estimating a class of triangular simultaneous equations models without exclusion restrictions. *Journal of Econometrics* 154(2): 154-164.

Koo, W. W., P. L. Kennedy and A. Skripnitchenko. 2006. Regional preferential trade agreements: trade creation and diversion effects. *Applied Economic Perspectives and Policy* 28(3): 408-415.

Krueger, A. O. 1999. Are preferential trading arrangements trade-liberalizing or protectionist? *The Journal of Economic Perspectives* 13(4): 105-124.

Lambert, D. and S. McKoy. 2009. Trade creation and diversion effects of preferential trade associations on agricultural and food trade. *Journal of Agricultural Economics*. 60(1): 17-39.

Liu, X. 2009. GATT/WTO promotes trade strongly: sample selection and model specification. *Review of International Economics* 17(3): 428-446.

Magee, C. S. 2003. Endogenous preferential trade agreements: an empirical analysis. *Contributions in Economic Analysis & Policy*, 2(1): 1-17.

Manning, W. G. 1998. The logged dependent variable, heteroscedasticity, and the retransformation problem. *Journal of Health Economics* 17(3): 283-295.

Martin, W. and C. S. Pham. 2008. Estimating the gravity model when zero trade flows are frequent. Deakin University, Faculty of Business and Law, School of Accounting, Economics and Finance. Working Paper No. 2008_03.

Martin, W. and C. S. Pham. 2007. Estimating the gravity model when zero trade flows are important. World Bank. Mimeo.

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

Melitz, M. J. and G. I. Ottaviano. 2008. Market size, trade, and productivity. *The Review of Economic Studies* 75(1): 295-316.

Mullahy, J. 1998. Much ado about two: reconsidering retransformation and the two-part model in health econometrics. *Journal of Health Economics* 17(3): 247-281.

Olper, A. and V. Raimondi. 2008. Agricultural market integration in the OECD: a gravity-border effect approach. *Food Policy* 33(2): 165-175.

Portugal-Perez, A. and J. S. Wilson. 2012. Export performance and trade facilitation reform: hard and soft infrastructure. *World Development* 40(7): 1295-1307.

Sanso, M., R. Cuairan and F. Sanz. 1993. Bilateral trade flows, the gravity equation, and functional form. *The Review of Economics and Statistics* 75(2): 266-275.

Santos Silva, J. M. C., S. Tenreyro and F. Windmeijer. 2015. Testing competing models for non-negative data with many zeros. *Journal of Econometric Methods* 4(1): 29-46.

Santos Silva, J. M. C. and S. Tenreyro. 2006. The log of gravity. *The Review of Economics and Statistics* 88(4): 641-658.

Siliverstovs, B. and D. Schumacher. 2009. Estimating gravity equations: to log or not to log? *Empirical Economics* 36(3): 645-669.

Smith, M. D. 2003. On dependency in double-hurdle models. *Statistical Papers* 44(4): 581-595.

Smith, R. J., and R. W. Blundell. 1986. An exogeneity test for a simultaneous equation Tobit model with an application to labor supply. *Econometrica* 54(3): 679-685.

Sun, L. and M R. Reed. 2010. Impacts of free trade agreements on agricultural trade creation and trade diversion. *American Journal of Agricultural Economics* 92(5): 1351-1363.

Terza, J. V. 1998. Estimating count data models with endogenous switching: sample selection and endogenous treatment effects. *Journal of Econometrics* 84(1): 129-154.

Vuong, Q. H. 1989. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* 57(2): 307-333.

White, H. 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 48(4): 817-838.

WTO, 2015. Regional Trade Agreements. Available From https://www.wto.org/english/tratop_e/region_e/region_e.htm [Accessed April, 2015]

Xiong, B. and S. Chen. 2014. Estimating gravity equation models in the presence of sample selection and heteroscedasticity. *Applied Economics* 46(24): 2993-3003.

Yen, S. T. and J. Rosinski. 2008. On the marginal effects of variables in the log-transformed sample selection models. *Economics Letters* 100(1): 4-8.

APPENDIX A – BACKGROUND ON THE OPTIMIZATION PROCEDURE

All maximum-likelihood estimators were coded and solved in the Gauss mathematical and statistical system using the Broyden–Fletcher–Goldfarb–Shanno (BFGS) method. The BFGS method is a Quasi-Newton algorithm similar to the Davidson–Fletcher–Powell (DFP) method, which eliminates calculation of the second derivatives at each iteration. Starting with an initial estimate of the inverse of the second derivative matrix (i.e. the Hessian), the BFGS method provides a rank three update of the matrix at each iteration. The procedure can be more efficient than the rank two update performed using the DFP method, and has been shown to have excellent convergence properties, even for ill-behaved optimization problems (Greene, 2003).

Several practical issues on maximum-likelihood estimation are worth mentioning. First, convergence of the estimators in this study were based on the squared sum of the derivatives being less than 1e-5. Second, start values for the BCDH model were obtained by working from the simplest nested model (i.e. the two-part log-linear model) to the most complicated nested model (i.e. the Independent BCDH). Final estimates were randomly shocked to test whether the procedure converged back to the original solution. Third, the optimization algorithm for limited dependent variable models, such as the BCDH, began with a numeric approximation of the Hessian calculated from the first derivative vector. This numeric approximation was also used to calculate the covariance matrix of the parameter estimates, since in practice Quasi-Newton methods rarely converge to an accurate estimate of the inverse Hessian (Greene, 2003).