# CHAPTER 1: Partial equilibrium trade policy analysis with structural gravity[1]

## TABLE OF CONTENTS

## LIST OF FIGURES



## LIST OF TABLES



## LIST OF BOXES

## A. Overview and learning objectives

Despite solid theoretical foundations and remarkable empirical success, the empirical gravity equation is still often applied a-theoretically and without account for important estimation challenges that may lead to biased and even inconsistent gravity estimates. The objective of this chapter is to serve as a practical guide for estimating the effects of trade policies (and other determinants of bilateral trade) with the structural gravity model.

The first part of this chapter will present a brief overview of the evolution of gravity theory over time and review the theoretical foundations of the Armington-Constant Elasticity of Substitution (CES) version of the structural gravity model. Importantly, the Armington-CES framework is used as a representative theoretical setting for a wide family of trade models that all lead to the same empirical gravity specification. Next, the main challenges faced when estimating the gravity model will be discussed along with the solutions that have been proposed in the trade literature to address them. Drawing from the latest developments in the empirical gravity literature, six recommendations will be formulated to obtain reliable partial equilibrium estimates of the effects of bilateral and non-discriminatory trade policies within the same comprehensive and theoretically-consistent econometric specification. Interpretation of the partial equilibrium gravity estimates and methods to consistently aggregate bilateral trade costs will then be discussed. Finally, data sources for gravity analysis, including bilateral trade flows and trade costs, will be provided.

Once familiarized with these theoretical concepts and analytical tools, a series of empirical applications, demonstrating the usefulness, validity and applicability of the recommendations proposed will be presented. Specifically, instructions will be provided on how to estimate a structural gravity model in order to assess the partial equilibrium effects of traditional gravity variables (e.g. distance, common language …), globalization, and regional trade agreements (RTAs) (as a representative form of bilateral trade policy).

Two exercises are provided at the end of the chapter. Data and STATA do-files for the solution of these exercises can be downloaded from the website.

In this chapter, you will learn:

- How the structural gravity model is derived;
- Where to find the data needed to estimate econometrically the structural gravity model;
- What are the main measurement issues associated with gravity data;
- What are the main econometric issues associated with the estimation of the structural gravity model and how to address them;
- How to econometrically estimate the structural gravity model;
- How to interpret and consistently aggregate gravity estimates.

After reading this chapter, with good econometric knowledge, and familiarity with STATA, you will be able to estimate using STATA software a theoretically-consistent structural gravity model and assess the effects of trade policies (and other determinants) on bilateral trade, while interpreting the econometric results with key caveats in mind.

# B. Analytical tools

## 1. Structural gravity: from theory to empirics

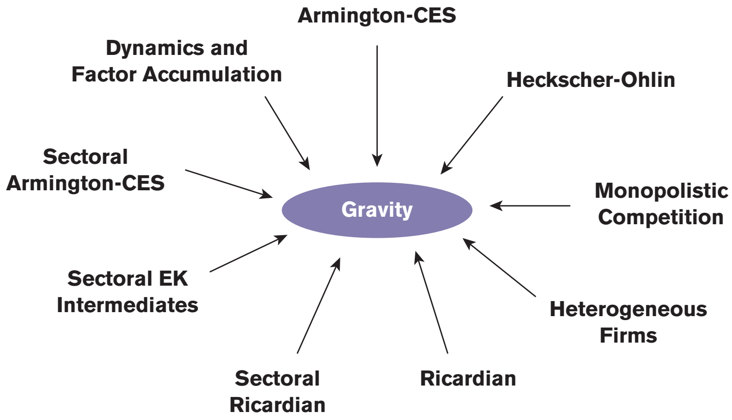
### (a) Evolution of gravity theory over time

According to Newton's Law of Universal Gravitation, any particle in the universe attracts any other particle thanks to a force that is directly proportional to the product of their masses and inversely proportional to the square of the distance between them. Applied to international trade, Newton's Law of Gravity implies that, just as particles are mutually attracted in proportion to their sizes and proximity, countries trade in proportion to their respective market size (e.g. gross domestic products) and proximity.

The initial applications of Newton's Law of Gravitation to economics are *a-theoretical*. Prominent examples include Ravenstein (1885) and Tinbergen (1962), who used gravity to study immigration and trade flows, respectively. Anderson (1979) is the first to offer a *theoretical* economic foundation for the gravity equation under the assumptions of product differentiation by place of origin and Constant Elasticity of Substitution (CES) expenditures. Another early contribution to gravity theory is Bergstrand (1985).

Despite these theoretical developments and its solid empirical performance, the gravity model of trade struggled to make much impact in the profession until the late 1990s and early 2000s. Arguably, the most influential structural gravity theories in economics are those of Eaton and

**Figure 1 Gravity model's strong theoretical foundations**



Kortum (EK) (2002), who derived gravity on the supply side as a Ricardian structure with intermediate goods, and Anderson and van Wincoop (2003), who popularized the Armington-CES model of Anderson (1979) and emphasized the importance of the general equilibrium effects of trade costs.

The academic interest in the gravity model was recently stimulated by the influential work of Arkolakis et al. (2012), who demonstrated that a large class of models generate *isomorphic* gravity equations which preserve the gains from trade. As depicted in Figure 1, the gains from trade are invariant to a series of alternative micro-foundations including a single economy model with monopolistic competition (Anderson, 1979; Anderson and van Wincoop, 2003); a Heckscher-Ohlin framework (Bergstrand, 1985; Deardoff, 1998); a Ricardian framework (Eaton and Kortum, 2002); entry of heterogeneous firms, selection into markets (Chaney, 2008; Helpman et al., 2008); a sectoral Armington-model (Anderson and Yotov, 2016); a sectoral Ricardian model (Costinot et al., 2012; Chor, 2010); a sectoral input-output linkages gravity model based on Eaton and Kortum (2002) (Caliendo and Parro, 2015), and a dynamic framework with asset accumulation (Olivero and Yotov, 2012, Anderson et al. 2015C, and Eaton et al., 2016). Most recently, Allen et al. (2014) established the universal power of gravity by deriving sufficient conditions for the existence and uniqueness of the trade equilibrium for a wide class of general equilibrium trade models.

## (b) Review of the structural gravity model

One of the main advantages of the structural gravity model is that it delivers a tractable framework for trade policy analysis in a multi-country environment. Accordingly, the model reviewed in this *Advanced Guide* considers a world that consists of $N$ countries, where each economy produces a variety of goods (i.e. goods are differentiated by place of origin (Armington, 1969)) that is traded with the rest of the world. The supply of each good is fixed to $Q_i$, and the factory-gate price for each variety is $p_i$. Thus, the value of domestic production in a representative economy is defined as $Y_i = p_i Q_i$, where $Y_i$ is also the nominal income in country *i*. Country *i*'s aggregate expenditure is denoted by $E_i$. Aggregate expenditure can also be expressed in terms of nominal income by $E_i = \phi_i Y_i$, where $\phi_i > 1$ shows that country *i* runs a trade deficit, while $1 > \phi_i > 0$ reflects a trade surplus. Similar to Dekle et al. (2007; 2008), trade deficits and surpluses are treated as exogenous. For brevity's sake, the time dimension *t* is omitted in the derivation of the structural gravity model. In addition, the structural gravity model presented below is derived from the demand side. However, as demonstrated in Appendix A, the same gravity system can be derived from the supply side.

On the demand side, consumer preferences are assumed to be homothetic, identical across countries, and given by a CES-utility function for country *j*:[2]

$$\left\{ \sum_i \alpha_i^{\frac{1-\sigma}{\sigma}} c_{ij}^{\frac{\sigma-1}{\sigma}} \right\}^{\frac{\sigma}{\sigma-1}} \tag{1-1}$$

where $\sigma > 1$ is the elasticity of substitution among different varieties, i.e. goods from different countries, $\alpha_i > 0$ is the CES preference parameter, which will remain treated as an exogenous taste parameter and $c_{ij}$ denotes consumption of varieties from country *i* in country *j*.

Consumers maximize equation (1-1) subject to the following standard budget constraint:

$$\sum_i p_{ij} c_{ij} = E_j \tag{1-2}$$

Equation (1-2) ensures that the total expenditure in country $j$, $E_j$, is equal to the total spending on varieties from all countries, including $j$, at delivered prices $p_{ij} = p_i t_{ij}$, which are defined conveniently as a function of factory-gate prices in the country of origin, $p_i$, marked up by bilateral trade costs, $t_{ij} \geq 1$, between trading partners $i$ and $j$. Throughout the analysis, the bilateral trade costs are defined as iceberg costs, as is standard in the trade literature (Samuelson, 1952). In order to deliver one unit of its variety to country $j$, country $i$ must ship $t_{ij} \geq 1$ units, i.e. $1/t_{ij}$ of the initial shipment melts "en route". While the Armington model presumes that all bilateral trade costs are variable, in principle, structural gravity can also accommodate fixed trade costs (Melitz, 2003). The iceberg trade costs metaphor can also be extended to accommodate fixed costs with the interpretation that "a chunk of the iceberg breaks off as it parts from the mother glacier" (Anderson, 2011).

Solving the consumer's optimization problem yields the expenditures on goods shipped from origin $i$ to destination $j$ as:[3]

$$X_{ij} = \left( \frac{\alpha_i p_i t_{ij}}{P_j} \right)^{(1-\sigma)} E_j \tag{1-3}$$

where $X_{ij}$ denotes trade flows from exporter $i$ to destination $j$ and, for now, $P_i$ can be interpreted as a CES consumer price index:

$$P_j = \left[ \sum_i \left( \alpha_i p_i t_{ij} \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \tag{1-4}$$

Given that the elasticity of substitution is greater than one, $\sigma > 1$, equation (1-3) captures several intuitive relationships. In particular, expenditure in country $j$ on goods from source $i$, $X_{ij}$, is:

(i) proportional to total expenditure, $E_j$, in destination $j$. The simple intuition is that, all else equal, larger/richer markets consume more of all varieties, including goods from $i$.

(ii) inversely related to the (delivered) prices of varieties from origin $i$ to destination $j$, $p_{ij} = p_i t_{ij}$. This is a direct reflection of the *law of demand*, which depends not only on factory-gate price $p_i$ but also on bilateral trade cost $t_{ij}$ between partners $i$ and $j$. The ideal combination that favours bilateral trade is an efficient producer, characterized by low factory-gate price, and low bilateral trade cost between countries $i$ and $j$.

(iii) directly related to the CES price aggregator $P_j$. This relationship reflects the substitution effects across varieties from different countries. All else equal, the relatively more expensive the rest of the varieties in the world are, the more consumers in country *j* will substitute away from them and toward the goods from country *i*.

(iv) contingent on the elasticity of substitution $\sigma_i$ when factory-gate prices or the aggregate CES prices (or in the combination of those as a relative price) change. All else equal, a higher elasticity of substitution will magnify the trade diversion effects from the more expensive commodities to the cheaper ones.

The final step in the derivation of the structural gravity model is to impose market clearance for goods from each origin:

$$Y_i = \sum_j \left( \frac{\alpha_i p_i t_{ij}}{P_j} \right)^{1-\sigma} E_j \tag{1-5}$$

Equation (1-5) states that, at delivered prices (because part of the shipments melt “en route”), the value of output in country *i*, $Y_i$, should be equal to the total expenditure of this country’s variety in all countries in the world, including *i* itself. To see this intuition more clearly, note that the right-hand-side expression in equation (1-5) can be replaced with the sum of all bilateral shipments from *i* as defined in equation (1-3), so that $Y_i \equiv \sum_j X_{ij} \ \forall \ i$.

Defining $Y \equiv \sum_i Y_i$ and dividing equation (1-5) by *Y*, the terms can be rearranged to obtain:

$$(\alpha_i p_i)^{1-\sigma} = \frac{\frac{Y_i}{Y}}{\sum_j \left( \frac{t_{ij}}{P_j} \right)^{1-\sigma} \frac{E_j}{Y}} \tag{1-6}$$

Following Anderson and van Wincoop (2003), the term in the denominator of equation (1-6) can be defined as $\Pi_i^{1-\sigma} \equiv \sum_j \left( t_{ij}/P_j \right)^{1-\sigma} E_j/Y$, and be substituted into equation (1-6):

$$(\alpha_i p_i)^{1-\sigma} = \frac{Y_i/Y}{\Pi_i^{1-\sigma}} \tag{1-7}$$

Using equation (1-7) to substitute for the power transform $(\alpha_i p_i)^{1-\sigma}$ in equations (1-3) and (1-4), and combining the definition of $\Pi_i^{1-\sigma}$ with the resulting expressions that correspond to equations (1-3) and (1-4), the structural gravity system is given by:

$$X_{ij} = \frac{Y_i E_j}{Y} \left( \frac{t_{ij}}{\Pi_i P_j} \right)^{1-\sigma} \tag{1-8}$$

$$\Pi_i^{1-\sigma} = \sum_j \left(\frac{t_{ij}}{P_j}\right)^{1-\sigma} \frac{E_j}{Y} \tag{1-9}$$

$$P_j^{1-\sigma} = \sum_i \left(\frac{t_{ij}}{\Pi_i}\right)^{1-\sigma} \frac{Y_i}{Y} \tag{1-10}$$

### (c) Structural decomposition of gravity: size *vs.* trade cost

Equation (1-8), representing the theoretical gravity equation that governs bilateral trade flows, can be conveniently decomposed into two terms: (i) a *size* term, $Y_i E_j / Y$, and a *trade cost* term, $\left(t_{ij}/\left(\Pi_i P_j\right)\right)^{1-\sigma}$:

(i) The intuitive interpretation of the *size term*, $Y_i E_j / Y$, is as the hypothetical level of frictionless trade between partners *i* and *j* if there were no trade costs.[4] Mechanically, this can be shown by eliminating bilateral trade frictions (i.e. setting $t_{ij} = 1$), and re-deriving the gravity system. Intuitively, a frictionless world implies that consumers will face the same price for a given variety regardless of their physical location and that their expenditure share on goods from a particular country will be equal to the share of production in the source country in the global economy (i.e. $X_{ij}/E_j = Y_i/Y$). Overall, the size term already carries some very useful information regarding the relationship between country size and bilateral trade flows:[5] namely, large producers will export more to all destinations; big/rich markets will import more from all sources; and trade flows between countries *i* and *j* will be larger the more similar in size the trading partners are.

(ii) The natural interpretation of the *trade cost* term, $\left(t_{ij}/(\Pi_i P_j)\right)^{1-\sigma}$, is that it captures the total effects of trade costs that drive a wedge between realized and frictionless trade. The trade cost term consists of three components:

   (1) Bilateral trade cost between partners *i* and *j*, $t_{ij}$, is typically approximated in the literature by various geographic and trade policy variables, such as bilateral distance, tariffs and the presence of regional trade agreements (RTAs) between partners *i* and *j*.
   (2) The structural term $P_j$, coined by Anderson and van Wincoop (2003) as inward multilateral resistance represents importer *j*'s ease of market access.
   (3) The structural term $\Pi_i$, defined as outward multilateral resistances by Anderson and van Wincoop (2003), measures exporter *i*'s ease of market access.

As will be discussed in more details in section B.1 of Chapter 2, the multilateral resistances are the vehicles that translate the initial, partial equilibrium effects of trade policy at the bilateral level to country-specific effects on consumer and producer prices. The direct effects do give the initial impact effects of trade costs on trade flows, while the general equilibrium trade costs also take into account the changes in prices, incomes and expenditures induced by trade cost changes. While this chapter focuses on the direct, partial effects of trade costs, chapter 2 deals with the general equilibrium trade costs.

**Box 1 Analogy between the Newtonian theory of gravitation and the gravity trade model**

To see the remarkable resemblance between the trade gravity equation and the corresponding equation from physics, two terms, $T_{ij}^{\theta}$ and $\tilde{G}$ have to be defined in equation (1-8) as reported in the right-hand side of the table below.

| *Newton's Law of Universal Gravitation* | *Gravity Trade Model* |
|---|---|
| $F_{ij} = G\frac{M_i M_j}{D_{ij}^2}$ | $X_{ij} = \tilde{G}\frac{Y_i E_j}{T_{ij}^{\theta}}$ |
| where: | where: |
| – $F_{ij}$: gravitational force between objects *i* and *j* | – $X_{ij}$: exports from countries *i* and *j* |
| – $G$: gravitational constant | – $\tilde{G}$: inverse of world production $\tilde{G} \equiv 1/Y$ |
| – $M_i$: object *i*'s mass | – $Y_i$: country *i*'s domestic production |
| – $M_j$: object *j*'s mass | – $E_j$: country *j*'s aggregate expenditure |
| – $D_{ij}$: distance between objects *i* and *j* | – $T_{ij}^{\theta}$: total trade costs between countries *i* and *j* $T_{ij}^{\theta} \equiv \left(t_{ij}/\left(\Pi_i P_j\right)\right)^{\sigma-1}$ |

Based on the metaphor of Newton's Law of Universal Gravitation, the gravity model of trade predicts that international trade (gravitational force) between two countries (objects) is directly proportional to the product of their sizes (masses) and inversely proportional to the trade frictions (the square of distance) between them.

## 2. Gravity estimation: challenges, solutions and best practices

Given the multiplicative nature of the structural gravity equation (1-8), and assuming that it holds in each period of time $t$, it is possible to log-linearize it and expand it with an additive error term, $\varepsilon_{ij,t}$:

$$\ln X_{ij,t} = \ln E_{j,t} + \ln Y_{i,t} - \ln Y_t + (1-\sigma)\ln t_{ij,t} - (1-\sigma)\ln P_{j,t} - (1-\sigma)\ln \Pi_{i,t} + \varepsilon_{ij,t} \quad (1\text{-}11)$$

Specification (1-11) is the most popular version of the empirical gravity equation, and it has been used routinely in the trade literature to study the effects of various determinants of bilateral trade. Hundreds of papers have used the gravity equation to study the effects of geography, demographics, RTAs, tariffs, exports subsidies, embargoes, trade sanctions, the World Trade Organization membership, currency unions, foreign aid, immigration, foreign direct investment, cultural ties, trust, reputation, mega sporting events (Olympic Games and World Cup), melting ice caps, etc. on international trade.

Despite the numerous applications of the gravity model and despite the great progress in the empirical gravity literature, many of the gravity estimates found in the existing literature still suffer biases and even inconsistency, which, as demonstrated in this section, can be avoided with some simple steps and stricter adherence to gravity theory.

This section begins with a discussion of the main challenges that need to be addressed in order to obtain reliable estimates with the structural gravity model. In addition, the solutions that have been proposed in the literature to address each of those challenges are reviewed and discussed. Capitalizing on the latest developments in the gravity literature, six recommendations to obtain reliable estimates of the structural gravity model are formulated. Finally, a comprehensive and theoretically-consistent estimating gravity specification that simultaneously identifies the effects of bilateral and unilateral non-discriminatory trade policy is proposed. Relevant examples of STATA commands are also presented throughout the section.

## (a) Challenges and solutions for estimating structural gravity models

Estimating the gravity model is subject to a number of modelling and econometric issues. This section reviews the eight main issues and discusses the relevant solutions that have been proposed in the literature to address them.

### *Challenge 1: Multilateral resistances*

One obvious challenge with the estimation of gravity equation (1-11) is that the multilateral resistance terms $P_{j,t}$ and $\Pi_{i,t}$ are theoretical constructs and, as such, they are not directly observable by the researcher and/or by the policy maker. Baldwin and Taglioni (2006) emphasize the importance of proper control for the multilateral resistance terms by characterizing studies that fail to do that as committing the "Gold Medal Mistake".

***Solutions to challenge* 1:** The treatment of the multilateral resistance terms in gravity estimations has evolved over the years and researchers have proposed various solutions to this challenge.

(i) In their original paper, Anderson and van Wincoop (2003) use iterative custom nonlinear least squares programming to account for the multilateral resistances in a static setting. Specifically, they first estimate the trade cost parameters without controlling for the multilateral resistances. Then, they use the estimated trade costs to construct an initial set of multilateral resistances. Then, they reestimate the gravity model using the initial multilateral resistances in the regression to obtain a new set of trade costs, which are used to construct a new set of multilateral resistances. The process is repeated until convergence, i.e. until the gravity estimates stop changing.

(ii) Many researchers have used a reduced-form version of the custom treatment from Anderson and van Wincoop (2003), where the multilateral resistance terms are approximated by the so-called "remoteness indexes" constructed as functions of bilateral distance, and Gross Domestic Products (GDPs) (Wei, 1996; Baier and Bergstrand, 2009). Head and Mayer (2014) criticize such reduced-form approaches as they bear little resemblance to the theoretical counterpart of multilateral terms.

(iii) An alternative approach to handle the multilateral resistances is to simply eliminate these terms by using appropriate ratios based on the structural gravity equation. Notable examples include Head and Ries (2001), Head et al. (2010), and Novy (2013) as discussed in Chapter 2.

(iv) Another approach, advocated by Hummels (2001) and Feenstra (2016), that is able to overcome the computational difficulties of the custom programming from Anderson and van Wincoop (2003), while at the same time fully accounting for the multilateral resistance terms, consists in using *directional* (*exporter* and *importer*) fixed effects in cross-section estimations. More recently, Olivero and Yotov (2012) extend the cross-section recommendations from Hummels (2001) and Feenstra (2016), and demonstrate that the multilateral resistance terms should be accounted for by *exporter-time* and *importer-time* fixed effects in a dynamic gravity estimation framework with panel data. It should be noted that in addition to accounting for the unobservable multilateral resistance terms, the *exporter-time* and *importer-time* fixed effects will also absorb the size variables ($E_{j,t}$ and $Y_{i,t}$) from the structural gravity model as well as all other observable and unobservable country-specific characteristics, which vary across these dimensions, including various national policies, institutions, and exchange rates.

### *Challenge 2: Zero trade flows*

Starting with Tinbergen (1962) and continuing today, the ordinary least-squares (OLS) estimator has been the most widely used technique to estimate various versions of the gravity equation (1-11). A clear drawback of the OLS approach, however, is that it cannot take into account the information contained in the zero trade flows, because these observations are simply dropped from the estimation sample when the value of trade is transformed into a logarithmic form. The problem with the zeroes becomes more pronounced the more disaggregated the trade data are. It is especially severe for sectoral services trade due to the highly localized consumption and highly specialized production.

***Solutions to challenge 2*:** Researchers have, over the years, proposed several approaches to handle the presence of zero trade flows.

(i) One frequently applied and very convenient – but theoretically inconsistent – method is to just add a very small, and in fact completely arbitrary, value to replace the zero trade flows. As noted in Head and Mayer (2014), however, this approach should be avoided because the results depend on the units of measurement and the interpretation of the gravity coefficients as elasticities is lost.[6]

(ii) Eaton and Tamura (1995) and Martin and Pham (2008) propose the use of the Tobit estimator as an econometric solution to the presence of zeroes. However, gravity theory is silent about the determination of the Tobit thresholds, causing a disconnect between estimation and theory. In practice, the Tobit model would apply to a situation where small values of trade are rounded to zero or actual zero trade might reflect desired negative trade.

(iii) The difficulty associated with the Tobit model is overcome by Helpman et al. (2008) who propose a theoretically-founded two-step selection process, where exporters must absorb some fixed costs to enter a market. Thus, fixed costs provide an intuitive economic explanation for the zero trade flows to bridge theory and empirics. The Helpman, Melitz and Rubinstein (HMR) model is estimated in two stages: (i) a first-stage Probit estimation, which determines the probability to export, and (ii) a second-stage OLS estimation based on the positive sample of trade flows that also accounts for selection into exporting due to fixed costs of exporting. Some challenges with the HMR estimation are that it is hard to find good exclusion restrictions for the first-stage Probit estimation and/or the need for custom programming when identification relies on functional form. Additional difficulties with the HMR approach arise for panel data estimations and when dynamic considerations are taken into account.

(iv) Egger et al. (2011) suggest a two-part gravity model that enables to decompose the effects of the explanatory variables on exports into an effect on the extensive country margin, i.e. the decision to export to a country at all, and on the intensive margin, i.e. the value of exports conditional on positive exports. Additionally, and contrary to Helpman et al. (2008), their approach also takes care of potential endogenous regressors such as RTAs in the estimating equation for the extensive and intensive margin (see Challenge 5).

(v) An easy and convenient solution to the presence of zero trade flows is to estimate the gravity model in multiplicative form instead of logarithmic form. This approach, advocated by Santos Silva and Tenreyro (2006), consists in applying the Poisson Pseudo Maximum Likelihood (PPML) estimator to estimate the gravity model.[7] Monte Carlo simulations show that the PPML estimator performs very well even when the proportion of zeroes is large.

### *Challenge 3: Heteroscedasticity of trade data*

It is well known that trade data are plagued by heteroscedasticity. The problem is important because, as pointed out by Santos Silva and Tenreyro (2006), in the presence of heteroscedasticity (and owing to Jensen's inequality), the estimates of the effects of trade costs and trade policy are not only biased but also inconsistent when the gravity model is estimated in log-linear form with the OLS estimator (or any other estimator that requires non-linear transformation).

***Solutions to challenge* 3:** The literature proposes at least two solutions to address the issue of heteroscedasticity in the gravity equation.

(i) Equation (1-11) can be estimated after transforming the dependent variable into size-adjusted trade, which is defined as the ratio between trade and the product of the sizes of the two markets, $X_{ij,t}/(E_{j,t}Y_{i,t})$, (Anderson and van Wincoop, 2003). The intuition behind this adjustment is that, arguably, the variance of the error term $\varepsilon_{ij,t}$ is proportional to the product of the sizes of the two markets. A potential drawback of this approach is that it accounts for (the product of) country size as the only source of heteroscedasticity. Furthermore, using the proposed size-adjusted trade as dependent variable would not eliminate the issue of "zero trade flows" highlighted in Challenge 2.

(ii) An alternative and more comprehensive approach, proposed by Santos Silva and Tenreyro (2006), is to apply the PPML estimator.[8] In addition, as discussed above, the PPML estimator also effectively handles the presence of zero trade flows, making it a very attractive choice for empirical gravity analysis.

### *Challenge 4: Bilateral trade costs*

Proper specification of bilateral trade costs is crucial for partial equilibrium as well as for general equilibrium trade policy analysis.

***Solutions to challenge* 4:** The standard practice suggested in the literature is to proxy for the bilateral trade cost term appearing in the structural gravity specification (1-11), $(1\text{-}\sigma)\ln t_{ij,t}$, by using a series of observable variables most of which have become standard covariates in empirical gravity specifications, namely:

$$(1-\sigma)\ln t_{ij,t} = \beta_1 \ln DIST_{ij} + \beta_2 CNTG_{ij} + \beta_3 LANG_{ij} + \beta_4 CLNY_{ij} + \beta_5 RTA_{ij,t} + \beta_6 \tilde{\tau}_{ij,t} \quad (1\text{-}12)$$

The first two variables in equation (1-12) are the most widely used and robust gravity proxies for trade costs. In $DIST_{ij}$ is the logarithm of bilateral distance between trading partners *i* and *j*, and $CNTG_{ij}$ is an indicator variable that captures the presence of contiguous borders between countries *i* and *j*.[9] $LANG_{ij}$ and $CLNY_{ij}$ are dummy variables that take the value of one for common official language and for the presence of colonial ties, respectively. Finally, $RTA_{ij,t}$ and $\tilde{\tau}_{ij,t}$ are both trade policy variables. $RTA_{ij,t}$ is a dummy variable that accounts for the presence of a RTA between trading partners *i* and *j* at time *t* by taking the value of one, and zero otherwise. The term $\tilde{\tau}_{ij,t}$ accounts for bilateral tariffs and is defined as $\tilde{\tau}_{ij,t} = \ln(1 + tariff_{ij,t})$, where $tariff_{ij,t}$ is the tariff that country *j* imposes on imports from country *i* at time *t*. Importantly, since tariffs act as direct price shifters, the coefficient on $\tilde{\tau}_{ij,t}$ can be expressed only in terms of the trade elasticity of substitution $\beta_6 = -\sigma$, which means that the trade elasticity itself can be recovered directly from the estimate on $\tilde{\tau}_{ij,t}$ as $\hat{\sigma} = -\hat{\beta}_6$. Appendix A.2 provides the derivation and implications of the structural gravity model with tariffs.

### *Challenge 5: Endogeneity of trade policy*

One of the biggest challenges in obtaining reliable estimates of the effects of trade policy within the gravity model is that the trade policy variables $RTA_{ij,t}$ and $\tilde{\tau}_{ij,t}$ are endogenous, because it is possible that trade policy may be correlated with unobservable cross-sectional trade costs. For instance, trade policy variables may suffer from "reverse causality", because, all else equal, a given country is more likely to liberalize its trade with another country that is already a significant trade partner.

***Solutions to challenge* 5:** The issue of endogeneity of trade policy is well-known in the trade literature (Trefler, 1993). However, primarily due to the lack of reliable instruments, early attempts to account for endogeneity with standard instrumental variable (IV) treatments in cross-sectional settings have not been successful in addressing the problem.[10]

Baier and Bergstrand (2007) summarize the findings from existing IV studies as "at best mixed evidence" of isolating the effect of RTAs on trade flows. The same authors propose applying the average treatment effect (ATE) methods described in Wooldridge (2010) in order to address the endogeneity of RTAs in panel trade data. In particular, *first-differencing* bilateral trade flows or using *country-pair* fixed effects eliminates or accounts for, respectively, the unobservable linkages between the endogenous trade policy covariate and the error term in gravity regressions. It should be noted that the set of pair fixed effects will absorb all bilateral time-invariant covariates (e.g. bilateral distance) that are used standardly in gravity regressions. However, the pair fixed effects will not prevent the estimation of the effects of bilateral trade policy, since trade policies are time-varying by definition. In addition the pair fixed effects will also account for any unobservable time invariant trade cost components. Egger and Nigai (2015) and Agnosteva et al. (2014) show that the pair-fixed effects are a better measure of bilateral trade costs than the standard set of gravity variables.

### *Challenge 6: Non-discriminatory trade policy*

Despite the importance of unilateral and non-discriminatory trade policies, such as export subsidies or most-favoured-nation (MFN) tariffs, and the natural interest to gauge their effects on bilateral trade flows, researchers and policy makers have struggled to estimate the effects of non-discriminatory trade policy within the structural gravity model. The issue with non-discriminatory trade policy covariates is that they are exporter- and/or importer-specific, and therefore they will be absorbed, respectively, by the *exporter-time* and by the *importer-time* fixed effects that need to be used in order to control for the multilateral resistances in the structural gravity model. More generally, in the presence of importer and exporter fixed effects, the gravity model can no longer estimate the impact of any variable (i) affecting exporters' propensity to export to all destinations (e.g. being an island); (ii) affecting imports without regard to origin (e.g. country-level average applied tariff); and (iii) representing sums, averages, and differences of country-specific variables (Head and Mayer, 2014).

***Solutions to challenge* 6:** *Several approaches have been proposed in the literature to be able to estimate the impact of non-discriminatory trade policy in a gravity setting.*

(i) One possible solution is to approximate the multilateral resistances with the "remoteness indexes" rather than including directional (exporter and importer) fixed effects. Renouncing exporter and importer fixed effects enables to identify separately the effects of country-specific policies of interest. However, this approach is not recommended because it does not account properly for the multilateral resistance terms, and is therefore likely to produce biased gravity estimates (including the effects of trade policy), as forcefully argued by Anderson and van Wincoop (2003).

(ii) Another solution is to employ a two-stage estimation, where the estimates of the multilateral resistances from the first-stage gravity regression are explained in an auxiliary regression that includes the non-discriminatory covariate of interest (Anderson and Yotov, 2016; Head and Mayer, 2014).

(iii) An alternative approach, proposed by Heid et al. (2015), consists in estimating the structural gravity model with *international* and *intra-national* trade flows by capitalizing on the fact that while non-discriminatory trade policies are country-specific, they do not apply to *intra-national* trade. As a result, the inclusion of *intra-national* trade implies that non-discriminatory variables become bilateral in nature, making their identification and estimation possible. As noted by Heid et al. (2015), the estimates of non-discriminatory trade policies in the structural gravity model are less likely to be subject to endogeneity concerns as compared to their bilateral counterparts for two reasons. First, it is unlikely that a non-discriminatory trade policy will be influenced by any bilateral trade flow. Second, the directional fixed effects in the structural gravity model will absorb much of the unobserved correlation between the non-discriminatory trade policy covariates and the gravity error term.

### *Challenge 7: Adjustment to trade policy changes*

It is natural to expect that the adjustment of trade flows in response to trade policy changes will not be instantaneous. For that reason, Trefler (2004) criticizes trade estimations pooled over consecutive years. The challenge of adjustment is even more pronounced in econometric specifications with fixed effects such as the ones described in this section. As noted in Cheng and Wall (2005), fixed-effects estimation applied to data pooled over consecutive years is sometimes criticized on the grounds that dependent and independent variables cannot fully adjust in a single year's time.

***Solutions to challenge* 7:** In order to avoid this critique, researchers have used panel data with intervals instead of data pooled over consecutive years. For example, Trefler (2004) uses 3-year intervals, Anderson and Yotov (2016) use 4-year intervals, and Baier and Bergstrand (2007) use 5-year intervals. Olivero and Yotov (2012) provide empirical evidence that gravity estimates obtained with 3-year and 5-year interval trade data are very similar, while estimations performed with panel samples pooled over consecutive years produce suspicious estimates of the trade cost elasticity parameters.

### *Challenge 8: Gravity with disaggregated data*

Many trade policies are negotiated and applied at the sectoral level, such as tariffs. While it is in principle possible to aggregate trade policy and still use the aggregate gravity model, such aggregation practices should be avoided and, whenever possible, gravity should be estimated at the level of aggregation which is the target of the specific policy. Furthermore, even for policies that are negotiated at the aggregate level (e.g. some RTAs), it may be desirable to also obtain sectoral effects because the effects of these non-discriminatory policies may actually be quite heterogeneous across sectors.

***Solutions to challenge* 8:** Fortunately, one of the most attractive properties of the structural gravity theory is that the model is separable. In other words, bilateral expenditures across countries both at the aggregate and at the sectoral level are separable from output and expenditure at the country level (Larch and Yotov, 2016b). As demonstrated by Anderson and van Wincoop (2004), one nice implication of separability is that for a given set of country-level output ($Y_{i,t}^k$) and expenditure ($E_{j,t}^k$) values, where $k$ denotes a class of goods/sector, theory delivers the familiar sectoral gravity equation:

$$X_{ij,t}^k = \frac{Y_{i,t}^k E_{j,t}^k}{Y_t^k} \left( \frac{t_{ij,t}^k}{P_{j,t}^k \Pi_{i,t}^k} \right)^{1-\sigma_k} \tag{1-13}$$

Two properties of equation (1-13) deserve a note. First, by definition, the bilateral trade costs $t_{ij,t}^k$, including the effects of trade policy, are sector-specific. Second, the multilateral resistances are sector-specific as well. From an empirical perspective, trade separability implies that equation (1-13) can be estimated for each sector as if the data were aggregate. Alternatively, the gravity model can be estimated with data pooled across sectors, in which case the proper treatment of the multilateral resistance requires *exporter-product-time* and *importer-product-time* fixed effects, and the effects of trade policy should be allowed to vary by sector. Depending on the question of interest, the estimates of the trade policy variables in gravity estimations that are pooled across sectors can be sector-specific or constrained to be common across sectors.

## (b) Practical recommendations for estimating structural gravity model

Taking into account all of the above considerations and combining the best solutions suggested in the literature to address the challenges with the estimation of the gravity model, the following best practices for estimating structural gravity equations are highly recommended:

***Recommendation 1:*** *Whenever available, panel data should be used to obtain structural gravity estimates.*

Various reasons motivate this recommendation. First, using panel data leads to improved estimation efficiency. Second, the panel dimension enables to apply the pair-fixed-effects methods to address the issue of endogeneity of trade policy variables (Baier and Bergstrand, 2007). Third, on a related note, the use of panel data allows for a flexible and comprehensive treatment and estimation of the effects of time-invariant bilateral trade costs with pair fixed effects. The downside is that, as discussed in Box 2, panel data may not always be available.

***Recommendation 2:*** *Panel data with intervals should be used instead of data pooled over consecutive years in order to allow for adjustment in trade flows.*

Interval panel data should be employed in order to allow for adjustment in bilateral trade flows in response to trade policy or other changes in trade costs. Olivero and Yotov (2012) build a dynamic gravity model and experiment with alternative interval specifications and find that gravity estimates obtained with 3-, 4-, and 5-year lags deliver similar results with respect to the estimates of the standard gravity variables. It is recommended to experiment with alternative intervals while keeping estimation efficiency in mind.

***Recommendation 3:*** *Gravity estimations should be performed with intra-national and international trade flows data.*

The inclusion of *intra-national* trade data in structural gravity estimations is desirable for several reasons. First, it ensures consistency with gravity theory, where consumers choose among and consume domestic as well as foreign varieties. Second, it leads to the theoretically consistent identification of the effects of bilateral trade policies (Dai et al., 2014). Third, it also enables to identify and estimate the effects of non-discriminatory trade policies (Heid et al., 2015). Fourth, it resolves the "distance puzzle" in trade, by measuring the effects of distance on international trade relative to the effects of distance on internal trade (Yotov, 2012). Finally, it enables to capture the effects of globalization on international trade and to correct for biases in the estimation of the impact of RTAs on trade (Bergstrand et al., 2015). Importantly, *intra-national* trade data has to be constructed consistently as the difference between *gross* production value data and total exports. Section 4 provides further discussion on the construction and sources of *intra-national* trade data.

***Recommendation 4:*** *In accordance with gravity theory, directional time-varying (importer and exporter) fixed effects should be included in panel trade data.*

The use of *exporter-time* and *importer-time* fixed effects enables to control for the unobservable multilateral resistances, and potentially for any other observable and unobservable *characteristics* that vary over time for each exporter and importer, respectively (Anderson and van Wincoop, 2003). In addition, as will be discussed in detail in Chapter 2, the estimates of the fixed effects of the gravity model can be used directly to recover the estimates of the general equilibrium effects of trade policy changes as well as to construct a series of useful general equilibrium indexes summarizing and aggregating consistently the effects of trade policy and trade costs (Anderson et al., 2015b; Larch and Yotov, 2016b).

```
* STATA commands to create importer- and exporter-time fixed effects:
     egen exp_time = group(exporter year)
          tabulate exp_time, generate(EXPORTER_TIME_FE)
     egen imp_time = group(importer year)
          tabulate imp_time, generate(IMPORTER_TIME_FE)
```

***Recommendation 5:*** *Pair fixed effects should be included in gravity estimation with panel trade data.*

Two major benefits are associated with using pair fixed effects in gravity estimations. First, the pair fixed effects are able to account for the endogeneity of trade policy variables (Baier and Bergstrand, 2007). Second, on a related note, the pair fixed effects provide a flexible and comprehensive account of the effects of all time-invariant bilateral trade costs, because pair fixed effects have been shown to carry systematic information about *trade* costs in addition to the information captured by the standard gravity variables (Egger and Nigai, 2015; Agnosteva et al., 2014). The downside of using pair fixed effects is that one cannot identify the effects of any time-invariant bilateral determinants of trade flows, because the latter will be absorbed by the pair fixed effects. One way to address this issue is to apply a two-stage procedure, where the estimates of the pair fixed effects from the first-stage structural gravity equation are regressed on standard gravity variables in a second-stage estimation (Agnosteva et al., 2014). This two-step approach also enables to recover estimates of the pair fixed effects that cannot be identified directly in the first stage, due to missing or zero trade flows, and then the complete set of pair fixed effects can be used to construct the full matrix of bilateral trade costs and to perform counterfactual experiments (Anderson and Yotov, 2016).

```
* STATA commands to compute country-pair fixed effects:
     * Asymmetric country-pair fixed effects
     egen pair_id = group(exporter importer)
          tabulate pair_id, generate(PAIR_FE)

     * Symmetric country-pair fixed effects
     * Short-cut code valid if none of the pairs has identical distances.
     egen pair_id = group(DIST)
          tabulate pair_id, generate(PAIR_FE)
```

***Recommendation 6:*** *Estimate gravity with the Poisson Pseudo Maximum Likelihood (PPML) estimator.*

The use of the PPML estimator is justified on various grounds. First, the PPML estimator, applied to the gravity model expressed in a multiplicative form, accounts for heteroscedasticity, which often plagues trade data (Santos Silva and Tenreyro, 2006). Second, for the same reason, the PPML estimator is able to take advantage of the information contained in the zero trade flows. Third, the additive property of the PPML estimator ensures that the gravity fixed effects are identical to their corresponding structural terms (Arvis and Shepherd, 2013; Fally, 2015). Finally, as will be reviewed in greater details in Chapter 2, the PPML estimator can also be used to *calculate* theory-consistent general equilibrium effects of trade policies (Anderson et al., 2015b; Larch and Yotov, 2016b). As a robustness check, the gravity model can be estimated by applying the Gamma Pseudo Maximum Likelihood (GPML) and the OLS estimators (Head and Mayer, 2014).[11]

```
* STATA commands to estimate gravity model with the PPML estimator:
     ppml trade EXPORTER_TIME_FE* IMPORTER_TIME_FE* PAIR_FE* RTA, cluster(pair_id)

     * Alternative command: glm
     glm  trade EXPORTER_TIME_FE* IMPORTER_TIME_FE* PAIR_FE* RTA, cluster(pair_id) ///
          family(poisson) diff iter(30)
```

**Box 2 In the absence of panel trade data**

When panel data are not available, the gravity model can still be estimated with cross-section samples:

$$\ln X_{ij} = \ln E_j + \ln Y_i - \ln Y + (1-\sigma)\ln t_{ij} - (1-\sigma)\ln P_j - (1-\sigma)\ln \Pi_i + \varepsilon_{ij}$$

In a cross-section setting, recommendations 3, 4, and 6 mentioned above continue to hold, namely gravity specification should include *intra-national* trade and directional (importer and exporter) fixed effects, and be estimated applying the PPML estimator.

However, the recommendations 2 and 5 to allow for adjustment in trade flows by using interval data and to include pair fixed effects are no longer applicable. The gravity specification in cross-section should include the standard set of gravity variables (e.g., bilateral distance, contiguity …) instead of pair fixed effects, in order to proxy for bilateral trade costs. That being said, as pointed out by Egger and Nigai (2015) and Agnosteva et al. (2014), the error term should be interpreted with caution, because it may capture systematic effects of unobserved trade costs. In order to address the endogeneity of bilateral trade policy, IV treatment is highly recommended (Baier and Bergstrand, 2004; Egger et al., 2011).

## (c) A theoretically-consistent estimating structural gravity model

The best practices and recommendations proposed in the previous section are reflected in the following generic and comprehensive econometric version of the structural gravity model, which can be modified and adjusted by researchers and policy makers depending on their specific needs:

$$X_{ij,t} = \exp\left[\pi_{i,t} + \chi_{j,t} + \mu_{ij} + \eta_1 BTP_{ij,t} + \eta_2 NES_{i,t} \times INTL_{ij} + \eta_3 NIP_{j,t} \times INTL_{ij}\right] \times \varepsilon_{ij,t} \qquad (1\text{-}14)$$

The variable $X_{ij,t}$ denotes nominal trade flows, which include *international* and *intra-national* trade, at non-consecutive year *t*. The term $\pi_{i,t}$ denotes the set of time-varying source-country dummies, which control for the outward multilateral resistances, countries' output shares and, potentially any other observable and unobservable exporter-specific factors that may influence bilateral trade. The term $\chi_{j,t}$ encompasses the set of time-varying destination-country dummy variables that account for the inward multilateral resistances, total expenditure, and any other observable and unobservable importer-specific characteristics that may influence trade. The term $\mu_{ij}$ denotes the set of country-pair fixed effects, which serve two main purposes as highlighted in the previous section. First, the pair fixed effects are the most flexible and comprehensive measure of time-invariant bilateral trade costs because they will absorb all time-invariant gravity covariates from equation (1-12) along with any other time-invariant bilateral determinants of trade costs that are not observable by the researcher and/or the policy maker. Second, the pair fixed effects will absorb most of the linkages between the endogenous trade policy variables and the remainder error term $\varepsilon_{ij,t}$ in order to control for potential endogeneity of the former. In principle, it is possible that the error term in gravity equations may carry some systematic information about trade costs. However, due to the rich fixed effects structure in equation (1-14), researchers should be more confident to treat and interpret $\varepsilon_{ij,t}$ as a true measurement error. Finally, whether the error term $\varepsilon_{ij,t}$ in equation (1-14) is introduced as additive or multiplicative does not matter for the PPML estimator (Santos Silva and Tenreyro, 2006).

The term $BTP_{ij,t}$ represents the vector of any time-varying bilateral determinants of trade flows, such as RTAs, bilateral tariffs and currency unions. In principle, the $BTP_{ij,t}$ vector may include any time-varying covariates, however, given the focus on trade policy of this *Advanced Guide*, the expression *BTP* stands for Bilateral Trade Policy. The expression $NES_{i,t} \times INTL_{ij}$ corresponds to the product between $NES_{i,t}$ and $INTL_{ij}$. The term $NES_{i,t}$ denotes the vector of any Non-discriminatory Export Support (NES) policies, such as export subsidies, while $INTL_{ij}$ is a dummy variable taking a value of one for international trade between countries *i* and *j*, and zero otherwise. Importantly, the interaction between the country-specific NES variables and the bilateral dummy for international trade flows results in a new bilateral term, i.e. $NES_{i,t} \times INTL_{ij}$, which will enable to identify the effects of any non-discriminatory export support policies, even in the presence of exporter-time fixed effects as required by gravity theory (Heid et al., 2015). In addition, with appropriate data on export support measures that act as direct price-shifters, the estimate of the coefficient(s) associated with the variable(s) $NES_{i,t} \times INTL_{ij}$ can be used to recover an estimate of the export supply elasticity, which plays a prominent role in theoretical trade policy analysis but has attracted little attention in the empirical trade literature. Similarly, the covariate $NIP_{j,t} \times INTL_{ij}$ is constructed as the product between the term $NIP_{j,t}$, which denotes the vector of any Non-discriminatory Import Protection (NIP) policies, such as MFN tariffs, and the dummy for bilateral international trade $INTL_{ij}$. Given its bilateral nature, the expression $NIP_{j,t} \times INTL_{ij}$ can be used to identify the effects of any non-discriminatory import protection policies.

```
* STATA commands to compute the term NES_{i,t} × INTL_{ij} and NIP_{j,t} × INTL_{ij}:
      generate INTL = 1
            replace INTL = 0 if exporter == importer
      generate NES_INTL = NES * INTL
      generate NIP_INTL = NIP * INTL
```

## 3. Gravity estimates: interpretation and aggregation

Once equation (1-14) has been estimated, the researcher may be interested in interpreting and/or aggregating some of the estimates of the gravity model. This section discusses various approaches to interpret the gravity estimates in terms of partial equilibrium effects on bilateral trade, including the tariff equivalents of non-tariff trade policy variables. In addition, theoretically-consistent methods to aggregate bilateral trade costs into a single figure are presented too.

### (a) Interpretation of gravity estimates

Two related methods are widely used to interpret the estimates from gravity regressions. The first approach is to use the gravity estimates to construct *trade volume effects*, while the second approach capitalizes on the theoretical foundations of gravity to convert the estimates of various trade policies and other determinants of trade flows into *tariff equivalent effects*. In order to demonstrate how the structural gravity estimates can be translated into trade volume effects and interpreted as tariff equivalent effects, the following simplified version of the empirical gravity model (1-14) is considered:

$$X_{ij,t} = \exp\left[\pi_{i,t} + \chi_{j,t} + \beta_{DIST} \ln DIST_{ij} + \beta_{RTA} RTA_{ij,t} + \beta_{TARIFF} \tilde{\tau}_{ij,t}\right] \times \varepsilon_{ij,t} \qquad (1\text{-}15)$$

The variable ln $DIST_{ij}$ denotes the logarithm of bilateral distance between countries *i* and *j*. The covariate $RTA_{ij,t}$ represents an indicator variable taking the value of one if there is a RTA between countries *i* and *j* at time *t*, and zero otherwise. For expositional purposes, both variables ln $DIST_{ij}$ and $RTA_{ij,t}$ will be used, respectively, as representative continuous variable and dummy variable in gravity regressions. Finally, $\tilde{\tau}_{ij,t} = \ln(1 + tariff_{ij,t})$ accounts for bilateral tariffs, where $tariff_{ij,t}$ is the ad-valorem tariff that country *j* imposes on imports from country *i* at time *t*. Importantly, as emphasized earlier, the coefficient on bilateral tariffs, $\tilde{\tau}_{ij,t}$, can be interpreted in the context of the structural gravity model as the trade elasticity of substitution, namely $\beta_{TARIFF} = -\sigma$. Overall, the interpretation of the coefficient on tariffs in gravity regressions depends on the trade flow data used to estimate the model, which here are assumed to be expressed at *cost, insurance and freight* (*c.i.f) prices*, but not tariffs. See Appendix B of this chapter for further details.

#### ***Trade volume effects***

The construction of trade volume effects from gravity estimates is straightforward but depends on the nature of the variable, namely whether it is a continuous or an indicator variable.

*Trade volume effect of continuous variables.* In the case of continuous variables, such as bilateral distance, the interpretation of the estimate of the coefficient on the logarithm of the continuous variable is simply the elasticity of (the value of trade flows) with respect to the continuous variable. For example, the standard empirical value for the distance variable estimate in gravity regressions of $\hat{\beta}_{DIST} = -1$ implies that a 10 percent increase in distance should be accompanied by a 10 percent decrease in trade flows (Disdier and Head, 2008; Head and Mayer, 2014).

*Trade volume effect of indicator variables.* The volume effects triggered by a change in an indicator gravity variable, such as the presence of RTAs, can be calculated in percentage terms as follows:

$$\left[e^{\hat{\beta}_{dummy}}-1\right]\times 100 \qquad (1\text{-}16)$$

where $\hat{\beta}_{dummy}$ is the estimate of the effects of any indicator gravity variable specified in the gravity model. For example, the benchmark estimate of the effects of RTAs in gravity regressions found in the empirical literature of $\hat{\beta}_{RTA}=0.76$ implies that the RTAs that entered into force between 1960 and 2000 on average have increased trade by $[e^{0.76} - 1] \times 100 = 114$ percent (Baier and Bergstrand, 2007).

With the exception of the direct price shifters, such as tariffs, the estimates of the remaining gravity covariates consist of two components: (i) a structural component and (ii) a trade cost component. For example, the structural interpretation of the estimate of the coefficient of distance is $\hat{\beta}_{DIST}=(1-\hat{\sigma})\rho$, where $\rho$ is the elasticity of trade costs with respect to distance. This decomposition is useful for two reasons. First, because it enables to recover the direct effects of distance, namely $\rho=\hat{\beta}_{DIST}/(1-\hat{\sigma})$. Empirical evidence based on this approach suggests that the distance variable in gravity estimations accounts for much more than just transportation costs (Head and Mayer, 2013). Second, because it can, as discussed next, be used to convert gravity estimates into tariff equivalent effects.

```
* STATA commands to compute trade volume effects:
    ppml trade IMPORTER_FE* EXPORTER_FE* LN_DIST CNTG RTA TARIFF
    scalar TradeVolumeEffectCNTG = (exp(_b[CNTG]) - 1) * 100
    scalar TradeVolumeEffectDIST = _b[LN_DIST] * 100
```

### *Tariff equivalent effects*

Quantifying the effects of tariffs is useful both from a policy and from a pedagogical perspective. However, the proliferation of non-tariff trade measures poses big challenges in quantifying their effects on international trade. Furthermore, it is often useful and desirable to be able to express the effects of alternative trade policies in a consistent measure. The structural gravity model offers a solution that enables researchers and policy makers to translate the effects of concluding any trade policy variable into a tariff equivalent effect, i.e. to find the ad-valorem tariff whose removal would have generated the same impact as the trade policy in question. In the context of equation (1-16), the tariff equivalent effect of RTAs would be equal to:[12]

$$\left[e^{\hat{\beta}_{RTA}/-\hat{\beta}_{TARIFF}}-1\right]\times 100 \qquad (1\text{-}17)$$

where $\hat{\beta}_{RTA}$ and $\hat{\beta}_{TARIFF}$ are the estimates of the coefficient associated with variables RTAs and tariffs specified in equation (1-15), respectively.

In an ideal situation, the effects of tariffs and all other determinants of trade could and should be obtained within the same theoretically-consistent empirical specification. Although, as discussed earlier, most gravity estimations do not include tariffs, this does not necessarily preclude the calculation of tariff equivalent effects by relying on the structural properties of the gravity model in order to construct them. In particular, capitalizing on the structural interpretation of the coefficient on tariffs as $\beta_{TARIFF} = -\sigma$, where $\sigma$ is the trade elasticity of substitution, equation (1-17) becomes:

$$\left[e^{\hat{\beta}_{RTA}/\hat{\sigma}} - 1\right] \times 100 \qquad (1\text{-}18)$$

An advantage of the structural specification (1-18) is that it demonstrates that, in principle, no data on tariffs are needed in order to obtain tariff equivalent effects of other gravity covariates as long as reliable estimates of the trade elasticity of substitution are available from outside studies. Returning to the example of the effects of RTAs from Baier and Bergstrand (2007), and taking a representative value for the elasticity of substitution from the literature[13], $\sigma = 5$, the average tariff-equivalent fall of the introduction of RTAs would amount to $[e^{0.76/5}-1] \times 100 = 16.4$ percent.

```
* STATA commands to compute tariff equivalent effects:
    ppml trade IMPORTER_FE* EXPORTER_FE* LN_DIST RTA TARIFF

    * If trade elasticity of substitution is taken from tariff estimates
    scalar TariffEquivalentRTA_1 = (exp(_b[RTA]/(-_b[TARIFF])) - 1) * 100

    * If trade elasticity of substitution is taken from literature
    scalar sigma = 5
    scalar TariffEquivalentRTA_1 = (exp(_b[RTA]/sigma) - 1) * 100
```

### (b) Consistent aggregation of bilateral trade costs

Aggregation of bilateral trade costs may be desirable for many policy purposes. For example, policy makers in a customs union or common market may wish to aggregate the effects of changes in bilateral trade costs of members to the level of the customs union or common market. Similarly decision makers may wish to aggregate interprovincial trade costs to the national level. Finally, national agencies may find it useful to consistently aggregate sectoral trade costs to the aggregate level of the economy. While a-theoretic weights are often used to form such indexes, the literature emphasizes the practical importance of theoretically consistent weights (Anderson and Neary, 2005). Different theoretically consistent aggregation methods have been proposed in the literature (Agnosteva et al., 2014). Although, for expositional purposes, the focus will be on the aggregation across regions within a customs union at a given point of time, similar principles apply for consistent aggregation over sectors.

The goal is to consistently aggregate bilateral trade costs $t_{ij}$ within a customs union (CU) so as to preserve the aggregate export volume from $i$ to destinations $j$ in the subset of countries that belong to the CU, $j \in CU(i)$, $j \neq i$. Following Agnosteva et al. (2014), the effect of changes in bilateral trade costs $t_{ij}$, $j \in CU(i)$ on the multilateral resistances $\Pi_i$ and $P_j$ are ignored. This assumption is particularly useful for practical purposes and it is justified for subsets with small trade volume shares. Alternatively, a more computationally intensive procedure should take into account the changes in the multilateral resistances that are driven by changes in bilateral trade costs.

Under the assumption of no change in bilateral trade costs, the volume equivalent uniform bilateral trade cost index $b_{CU(i)}$ is implicitly defined as:

$$\sum_{j \in CU(i)} X_{ij} = \sum_{j \in CU(i)} \frac{Y_i E_j}{Y} \left( \frac{t_{ij}}{\Pi_i P_j} \right)^{1-\sigma} = \sum_{j \in CU(i)} \frac{Y_i E_j}{Y} \left( \frac{b_{CU(i)}}{\Pi_i P_j} \right)^{1-\sigma} \tag{1-19}$$

Dividing the middle and rightmost expressions of equation (1-19) by $Y_i / (\Pi_i^{1-\sigma} Y)$ and solving for $b_{CU(i)}$ yields:

$$b_{CU(i)} = \left[ \sum_{j \in CU(i)} \frac{\frac{E_j}{P_j^{1-\sigma}}}{\sum_{k \in CU(i)} \frac{E_k}{P_k^{1-\sigma}}} t_{ij}^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \tag{1-20}$$

Equation (1-20) reveals that the CU regional trade cost aggregate is a weighted-average across the bilateral trade costs for the exporters in the CU region. The weights in equation (1-20) can be interpreted in the spirit of the market access and market potential indexes from the economic geography literature (Redding and Venables, 2004). From a practical perspective, the weights can be constructed directly from the importer fixed effects, $\chi_j$, in the estimating gravity equation (1-14), so that the aggregating equation becomes:

$$b_{CU(i)} = \left[ \sum_{j \in CU(i)} \frac{\chi_j}{\sum_{k \in CU(i)} \chi_k} t_{ij}^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \tag{1-21}$$

Applying the same principles and methods delivers a consistent aggregate of bilateral trade costs for the customs union on the demand side:

$$b_{CU(j)} = \left[ \sum_{i \in CU(j)} \frac{\frac{Y_i}{\Pi_i^{1-\sigma}}}{\sum_{k \in CU(j)} \frac{Y_k}{\Pi_k^{1-\sigma}}} t_{ij}^{1-\sigma} \right]^{\frac{1}{1-\sigma}} = \left[ \sum_{i \in CU(j)} \frac{\pi_i}{\sum_{k \in CU(j)} \pi_k} t_{ij}^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \tag{1-22}$$

where, as can be seen from the rightmost expression, the aggregating weights are now the exporter fixed effects, $\pi_i$, from the gravity equation (1-14).

Expressing equations (1-21) and (1-22) in terms of importer and exporter fixed effects, respectively, is important for two reasons. First, from a theoretical perspective, owing to the additive property of the PPML estimator, the estimates of the gravity fixed effects correspond exactly to the structural gravity terms (Arvis and Shepherd, 2013; Fally, 2015). Second, from a practical perspective, this implies that consistent aggregation of bilateral trade costs at any level can be obtained in three simple steps:

*Step 1:* *Estimate the gravity model with the PPML estimator.*

*Step 2:* *Construct bilateral trade costs* $t_{ij,t}$ *for each pair.*

*Step 3:* *Aggregate bilateral trade costs at the desired level with the estimates of:*

- *importer fixed effects used as weights for the supply-side analysis* $(CU_{(i)})$.
- *exporter fixed effects used as weights for the demand-side analysis* $(CU_{(j)})$.

## 4. Gravity data: sources and limitations

Gravity equations have been estimated using a variety of country-specific and bilateral variables as determinants of bilateral trade flows. The goal of this section is to review the main data sources and the data limitations that researchers have faced when using these sources. Following the recommendation that gravity should be estimated with *exporter*(-*time*) and *importer*(-*time*) fixed effects, and also for brevity purposes, this section will mainly focus on data for the dependent gravity variable, i.e. bilateral trade flows, and on data that can be used to construct proxies for bilateral trade distortions. All web links for the data sources discussed in this section are provided as active links in Appendix C.

### (a) Bilateral trade flows data

Traditionally, gravity estimations have mostly been performed with aggregate data. However, mainly due to availability of more and more reliable disaggregated data, an increasing number of studies present sectoral and even product gravity analysis.

#### *Aggregate trade flows data*

The primary source of information for aggregated (country-level) bilateral trade flows is the International Monetary Fund (IMF)'s Direction of Trade Statistics (DOTS). The database covers 184 countries. Annual data are available from 1947, while monthly and quarterly data start from 1960. Data are reported in US dollars. Relying on DOTS and other national sources of data, Barbieri and Keshk have created a database (Correlates of War Project) that tracks total national trade and bilateral trade flows (imports and exports) between states from 1870-2009 in current US dollars.

### *Merchandise trade flows data*

Availability of trade flows data at the disaggregated level depends on the sector in question. Data on merchandise trade flows are available at disaggregated level and for a long period of time for several data sources. The UN Commodity Trade Statistics Database (COMTRADE) is the most common source of data of disaggregated trade by commodity. It reports annual bilateral trade flow data expressed in gross value and volume from 1962 for more than 160 countries on average. Monthly data are also available since 2010. Trade values are in current US dollars converted from national currencies. Data are available online through the UN website or through the World Bank's World Integrated Trade Solution (WITS) portal. The data are accessible in different nomenclatures and in different levels of disaggregation. Trade data classified according to the Harmonised System (HS) are available up to the 6-digit level (that is, at a level of detail that distinguishes about 5,000 separate goods items), which is the most disaggregated classification that is consistent across countries at the international level. Annual trade data are also classified using the Standard International Trade Classification (SITC). This classification focuses more on the economic functions of products at various stages of processing rather than the physical characteristics of a product. In its Rev. 4 version this classification reaches 5 digit (2,970 lines). Concordance tables exist to match data in HS and SITC classifications.

Measurement error is a standard problem with trade data. Import data have been traditionally more reliable because imports are monitored much more closely than exports by customs administrations, since the former are often subject to an import duty. Therefore, it is often advisable to use import data to construct the main dependent variable in gravity regressions. It is also recommended to use "mirror data", that is to use imports data from destination countries as a measure of exports from origin countries. It should be noted, however, that mirroring may not be a good idea in cases when the importing country applies very high tariffs and has weak monitoring capability at customs. In these cases, the incentive to avoid tariffs and border controls may lead to largely underestimated import data. For this reason, it is not uncommon to have declared imports of country *j* from exporter *i* that are lower than the declared exports of *i* to destination *j*, even though imports are reported at *cost, insurance and freight* (*c.i.f) prices* and exports are reported at *free on board (f.o.b) prices*, which do not include any costs associated with transportation.

In an attempt to reconcile declarations of importers and exporters in COMTRADE, the Centre d'Etudes Prospectives et d'Informations Internationales (CEPII) has created the Base Analytique du Commerce International (BACI). In addition, the BACI data are also cleaned to exclude re-exports. The BACI database provides trade data at the 6-digit HS level for more than 200 countries from 1995. Because the construction and processing of the BACI database requires time and it is based on original data from other primary sources, such as COMTRADE, the BACI data are available to the public with a time lag of one or two years as compared to COMTRADE. Finally, to tackle the problem of measurement errors, the World Trade Flows (WTF) database developed by Feenstra and Romalis (2014) omits observation where the ratio between *c.i.f* and *f.o.b.* is either less than 0.1 or larger than 10 and where the *c.i.f.* value is smaller than 50,000 US dollars. The WTF database contains bilateral trade data for 185 countries covering, on average, the period 1984-2014.

### *Services trade flows data*

Although there has been a significant effort and advances to offer data on trade in services, the availability of data on bilateral trade flows in services remains relatively limited. The Organisation for Economic Co-operation and Development (OECD) Trade in Services database offers data on bilateral services trade for 12 main services sectors and several sub-sectors according to the Extended Balance of Payments (EBOPS) 2010 classification (146 categories in total). The OECD database covers 35 countries including 32 OECD member countries as well as the Russian Federation, Colombia and Latvia from 1999 onwards. The UN Service Trade Database covers 46 economies from 2000 onwards and follows the EBOPS 2002 classification (114 categories: 86 standard items (11 main items), 24 memorandum items and 4 supplementary items). The WTO, UNCTAD and International Trade Centre (ITC) also jointly develop a database which contains bilateral annual service flows data for 36 countries at the same level of disaggregation as the OECD data from 2005 onwards according to the EBOPS 2010 classification. These bilateral data can be retrieved from the ITC TradeMap. An older version of this database, following the previous services classification (EBOPS 2002), covers 49 countries for the period 1980-2013.[14] Finally, based on adjusted data from the OECD, Eurostat, UN and IMF, the Trade in Service database, developed by Francois and Pindyuk (2013), reports bilateral service flows data classified according to the EBOPS 2002 classification and covering 248 countries, on average, for the period 1981-2010. Data comes from the OECD, Eurostat, UN and IMF. Adjustments have been made using mirroring, reconciliation of aggregated with underlying flows and consolidation. Services are classified according to the EBOPS 2002 classification.

### *Agriculture and resource sectors data*

The UN Food and Agriculture Organization Corporate Statistical Database (FAOSTAT) collects information on an annual basis in more than 100 countries. The Detailed Trade Matrix reports information on agricultural bilateral trade flows for over 600 food and commodities per year. It provides data for both quantities (in tons) and values (in thousands of US dollars) of agricultural imports and exports. Data are available for the period 1986-2013 and are gathered from national sources.

### *Intra-national trade flows data*

As argued above, the use of *intra-national* trade flows data is desirable and consistent with gravity theory. However, such data are not readily available and their use requires caution. Some countries, such as Canada, have devoted significant resources and special attention to carefully construct intra-national trade flows.[15] However, constructing an international database of intra-national trade flows is challenging for at least two reasons. First, traditionally researchers have constructed intra-national trade flows as apparent consumption, defined as the difference between production and total exports. However, aggregate production data are usually measured and reported as *value added* (e.g. GDP), while total exports are reported as *gross* value. That is why the production databases described below are based on sectoral data, usually covering goods only, for which value added and gross values are available and reported. Typically production data are classified using the Standard Industrial Classification (SIC). This nomenclature classifies products at the 4-digit level at the highest level of detail. Although concordance tables between various nomenclatures exist, matches are not perfect and one may need to move to higher levels of aggregation to guarantee a better match.

Despite these (and other) limitations, there have been efforts to merge bilateral trade and production data in order to construct consistent databases of international and intra-national trade flows, in particular for the manufacturing sectors. The World Bank's Trade, Production and Protection (TPP) database covers approximately 100 countries for the period 1976-2004 with information classification according to the Industrial Standard Industrial Classification (ISIC) Rev. 3 at the 3-digit level. The CEPII's Trade, Production and Bilateral Protection (TradeProd) provides data for over 150 countries for the period 1980-2006 expressed in ISIC Re. 2 at the 3-digits level. The UN Industrial Development Organization (UNIDO) Industrial Statistics (INDSTAT) reports data from 1962 onwards at the 2-digit level of ISIC Rev. 3 (INDSTAT2) or from 1990 onwards at the 4-digit level (INDSTAT4) for 166 countries.[16]

## (b) Bilateral trade costs data

As discussed earlier, one of the estimation challenges with gravity equations is to proxy for the unobservable bilateral trade costs $t_{ij,t}$ formulated in the structural gravity model. Traditionally, the bilateral trade costs in gravity equations are proxied by a series of observable variables that determine trade costs. From a broad practical perspective, trade costs can be divided into their time-varying and time-invariant components. Although, by nature, trade policy variables are time-varying, most of the standard gravity variables that are routinely included in gravity estimations include time-invariant (or very slowly time varying) covariates such as physical distance, contiguous borders, common language, and common history and colonial ties. The CEPII's GeoDist database reports data on time-invariant gravity variables for 225 countries (Mayer and Zignago, 2011).

## (c) Trade policy data

Trade policies are typically divided in tariffs and non-tariff measures (NTMs). Data on various trade policies are available through three main portals: the WTO's Tariff and NTM portals; the World Bank's World Integrated Trade Solution (WITS); and the ITC's various web-based "Map" tools. The main data sources on specific trade policy measures are presented below.

### *Tariff data*

Tariffs can be classified into three groups:

(i) MFN bound tariffs are the tariff ceiling above which countries have committed not to raise their applied tariff.[17]

(ii) MFN applied tariffs are the MFN tariffs imposed by a WTO member country on imports from other WTO members.

(iii) Preferential tariff rates are the tariffs countries have bilaterally negotiated under RTAs.

The WTO provides facilities to download tariff data for each of these three groups of tariffs. Applied tariffs data notified to the WTO can be found in the WTO's Integrated Data Base (IDB). The IDB contains data on MFN duties for applied and preferential duties for WTO member countries on an annual basis from 1996 onwards. Data are available at the tariff line level as reported by the country imposing these tariffs, i.e. more detailed than HS 6-digit level. The Consolidated Tariffs Schedules (CTS) database contains bound tariffs, tariff quotas and export subsidies bound commitments at the tariff line level, as well as domestic support commitments. Access to the IDB and CTS databases is possible through the WTO's Tariff Analysis Online (TAO) interface and for HS 6-digit pre-aggregated data through the Tariff Download Facility. The World Bank's WITS, developed in collaboration with UNCTAD, gathers together the WTO's IDB and CTS database, UNCTAD's Trade Analysis Information System (TRAINS) data along with trade flows data from the UN COMTRADE, data from Market Access Maps MacMap and the OECD's Agriculture Market Access Database in a unique interface to facilitate data extraction.

In practice, countries set tariffs at the tariff line level, which can be at the 6-, 8-, 9-, 10-, or 12-digit HS level depending on the country. Yet, researchers may need to aggregate tariffs in order to perform cross country comparisons, work with a dataset of a manageable size and/or match the information with information available for other variable, such as bilateral trade flows. Two simple approaches to aggregate tariffs have been proposed in the literature: (i) a simple average aggregation procedure, and (ii) an import-weighted average method. While simple and easy to implement, each of these procedures is subject to caveats. For example, when import-weighted averages are used to estimate the average degree of protection in a certain country, tariff lines with very high tariffs will have a low weight, because imports subject to high protection rates are likely to be small. At the extreme, paradoxically, for a given level of total imports, the contribution to the import-weighted average tariff of goods subject to prohibitive tariffs is the same as the contribution of goods subject to zero tariffs. In fact, in both cases the product between the tariff and the level of import will be zero. Similarly, using the simple average method may also be misleading, because the tariff rate associated with a good that represents an important share of the total trade of a sector has the same impact on the calculated average tariff as that of a good that represents a minimal share of trade. In order to tackle these tariff aggregation problems, the ITC MAcMap database includes weighted tariffs at the HS 6 digit level that are calculated on the basis of a reference group weighting scheme. Five groups of reference countries have been identified according to the PPP GDP per capita and trade openness. Total imports by a given group are normalized to account for its size. Then, the measure obtained is used as weight to aggregate data across partners and products (Bouët et al., 2005). MAcMap includes MFN and preferential tariff data for the years 2005-2014 up to the national tariff line level for 190 countries.

### *Main non-tariff measures (NTMs) databases*

Non-tariff measures (NTMs) have gained a more prominent role as trade and consumer protection tools in the current world economy. As a result, a number of databases have been developed. Six main NTM databases are readily available:

*UNCTAD's TRAINS* database was the first comprehensive database on NTMs. The database covers import (technical and non-technical measures) and export measures as well as information on "procedural obstacles" (e.g. administrative burdens, transparency issues or infrastructural challenges).

Information in TRAINS is coded in a binary form at the tariff line level, which bears the limitation that the data does not allow to distinguish between mild and stiff non-tariff measures.

NTMs information can also be retrieved from the *World Bank's TPP database*. This database provides information on a set of "core non-tariff trade barriers (NTBs)" including price-control, finance control, and quantity control measures. Variables included in the database consist of frequency measures, coverage ratios, and simple and import-weighted *ad-valorem* equivalents of NTMs at the HS 3-digit level.

*The CEPII's NTM-MAP database* is also based on the UNCTAD's TRAINS data and provides frequency measures, coverage ratios, and prevalence score ratios for technical barriers to trade (TBT), sanitary and phytosanitary (SPS) measures, pre-shipment inspections, contingent trade protective measures and non-automatic licensing, quotas, prohibitions, and quantity-control measures.

*The WTO Integrated Trade Intelligence Portal* (*I-TIP*) includes information on antidumping, countervailing measures, quantitative restrictions, safeguard measures, tariff rate quotas, export subsidies, TBT and SPS measures. In addition, information on specific trade concerns (STCs) raised in the WTO TBT and SPS committees are provided. All information available through I-TIP refers to countries' notifications to the WTO. As a result the availability of information for a given WTO country member depends on its compliance with the WTO's notification obligations.

*The WTO's Trade Monitoring database* gathers information about trade-related measures, such as trade remedies, export duties, and quantitative restrictions, implemented by WTO member countries following the 2008 global financial crisis.

Finally, the *Global Trade Alert database* reports policies that may affect trading partners' commercial interests, such as import tariffs, export incentives, export taxes, as well as other NTMs.

### *Specific non-tariff measures (NTMs) data*

Besides the NTMs databases presented above, several other databases focusing only on specific types of NTMs are available.

*Subsidies and government support measures*. The *OECD's Agricultural Policy database* accounts for different measures of agricultural support, such as the total support estimate, producer support estimate, consumer support estimate and general services support estimate (GSSE). Data are available from 1986 onwards.

The *WTO's Agriculture Information Management System* includes a series of measures notified by WTO member countries to the WTO Agricultural Committee, including export subsidies. These data are available from 1995 onwards. In addition, the WTO Consolidated Tariff Schedules (WTO-CTS) provides information about agricultural non-tariff commitments, which include tariff quotas and subsidies.

The *International Energy Agency* (*IEA*) website contains data about fossil fuel subsidies. The dataset covers oil, electricity, natural gas, coal and total fossil fuels in billions of real US dollars over the period 2012-2014. The data also include estimates for the average subsidisation rate (percent), subsidies per capita, and total subsidies as a share of GDP (percent).

Finally, the *World Bank* reports data about aggregated subsidies and other transfers in current local currency unit (LCU) by country from 1981 to 2015. Subsidies, grants, and other social benefits reported include all unrequited, non-repayable transfers on current account to private and public enterprises; grants to foreign governments, international organizations, and other government units; and social security, social assistance benefits, and employer social benefits in cash and in kind.

*Export restriction.* The OECD has developed and maintains data aggregated at the 6-digit level of HS 2007 on export restrictions for primary agricultural products as well as raw materials (minerals, metals, and wood). Various kinds of export restrictions are reported, such as export duties, export prohibitions, and licensing requirements. Information on primary agriculture products covers the period 1996-2012, while data on raw materials restrictions are only available for the years 2009-2014.

*Safeguards and antidumping/countervailing measures.* A series of useful databases, developed by Chad Bown and hosted by the World Bank's Temporary Trade Barriers Database (TTBD), provide information on safeguard, and antidumping and countervailing measures. The Global Antidumping (GAD) and the Global Countervailing Duties (GCVD) databases gather data for the period 1980-2015. The China-Specific Safeguards (CSFG) includes information for the period 2002-2015. The Global Safeguards (GSFG) and the WTO Dispute Settlement Understanding (WTO-DSU) Cases related to Antidumping, Safeguards or Countervailing Duties cover the period 1995-2015.

*Technical regulations and sanitary and phytosanitary measures.* The WTO's TBT Information Management System (TBT-IMS) and SPS Information Management System (SPS-IMS) provide access to all the TBT and SPS measures notified by WTO member countries to the WTO as well as any documents submitted to and released in the respective WTO committee. In addition, both TBT- and SPS-IMS report various STCs raised by WTO country members in their respective committees. Other relevant sources of information on TBT, include Perinorm, which is a bibliographic database, developed by the British Standards Institution, the Association Française de Normalisation, and the Deutsches Institut für Normung, with information on national, European, and international standards in 23 countries.

*Services trade restrictiveness indices.* The OECD Services Trade Restrictiveness Index (STRI) identifies policies restricting foreign entry and movement of people, and imposing barriers to competition and transparency as well as other measures. The World Bank's Services Trade Restrictions Database collects also information for different services trade policies in 103 countries and five main sectors (covering telecommunications, finance, transportation, retail, and professional services) and key modes of service supply for the period 2008-2010 (Borchert et al., 2012).

The I-TIP Services database, developed jointly by the WTO and World Bank, provides information on WTO members' commitments under the General Agreement on Trade in Services (GATS), RTAs applied measures in services and service statistics for 12 groups and 160 sub-groups according to the Services Sectoral Classification List, developed during the Uruguay Round of the General Agreement on Tariffs and Trade (GATT).

*Trade facilitation restrictions.* Information on trade facilitation is available in various databases covering different types of information related to trade facilitation measures (WTO, 2015).

The World Bank's Doing Business database reports various "trading across borders" indicators relevant to trade facilitation. In particular, information on the respective time and costs to import and export due to documentary requirement, border compliance, and domestic transport is available for 189 countries for the period 2004 onwards.

The OECD Trade Facilitation Indicators (TFIs) covers 16 indicators strongly linked to the provisions of the WTO Trade Facilitation Agreement (TFA) provisions, such as advance ruling, appeal procedures, fees and charges, formalities (documents, automation, procedures), (internal and external) cooperation as well as transit (fees and charges, formalities, guarantees and agreements and cooperation). The TFIs database tracks the trade facilitation performance of 152 countries for the years 2009 and 2015.

The World Bank's Logistic Performance Index (LPI) focuses on the logistic friendliness of a country and ranks countries along six dimensions: customs; infrastructure; ease of arranging shipments; quality of logistics services; tracking and tracing; and timeliness. The database covers 160 countries for the following years: 2007, 2010, 2012 and 2014.

The World Economic Forum (WEF) Enabling Trade Index (ETI) assesses the extent to which economies have in place institutions, policies, infrastructure and services facilitating the flow of goods over borders and their destinations. The index includes 79 indicators grouped into 4 areas: market access; border administration; infrastructure; and operating environment.

### *Regional trade agreements*

When countries form a RTA not only do they apply lower tariffs, but they also cooperate on a number of other policy areas that reduce overall bilateral trade costs among member countries beyond the removal of explicit trade barriers. One way to take this information into account is by including as a covariate in a gravity equation a dummy indicating whether or not there is a trade agreement in place between a specific pair of countries.

The WTO Regional Trade Agreements Information System (RTA-IS) reports detailed information on RTAs notified to the WTO, including the agreement's nature (customs unions, free trade agreements, or partial scope agreements); scope (goods, services or goods and services); signature date; and signatory countries as well as links to the official documents. The WTO's preferential trade arrangements (PTA) database also provides information on unilateral trade agreements, namely agreements of non-reciprocal nature, such as the General System of Preferences (GSP) and sub-schemes for least-developed countries (LDCs). Data on preference utilization rates are also available through the WTO's Tariff Analysis Online interface.

Building on the information provided in the WTO RTA-IS, other databases on RTAs have been developed. For instance, the RTAs database, developed by Mario Larch and readily available in STATA format, covers 468 RTAs from 1950 onwards. Similarly, the database on Economic Integration Agreements (EIAs), developed by Jeffrey H. Bergstrand, categorizes the bilateral relationship for the pairings of 195 countries during the period 1950-2005 by applying a multi-faceted index that distinguishes between unilateral, bilateral agreements, RTA, customs unions and common markets.

While informative, an indicator variable of the existence of RTAs cannot capture the fact that RTAs also differ in terms of scope and types of specific provisions covered. In order to address this issue, indexes of the depth of RTAs can be built starting from basic information on the coverage of the agreement. The 2013 World Trade Report codified provisions for a set of 100 RTAs signed between 1958 and 2011 by extending the data developed by Horn et al. (2010). The different RTAs' provisions are classified into one of the 52 policy areas identified by the authors. Some of these policy areas are defined as "WTO+" provisions when the RTA's provisions fall under the WTO's current mandate, reconfirm existing commitments and specify additional related obligations. Conversely, other policy areas are denoted as "WTO-X" provisions when the RTA's provisions establish obligations that are outside the WTO's current mandate. The codification also ascertains the legal enforceability of the RTA's obligations by assuming that the clearer, more specific and imperative the legal language used to express a commitment or undertaking, the more successfully it can be invoked by a complainant in a dispute settlement proceeding, and thus the greater likelihood of it being enforced. Following this methodology, the World Bank's Global Preferential Trade Agreements (GPTA) extends the coverage of the RTAs to include 330 agreements. Data on RTAs' depth of integration can also be retrieved from the Design of Trade Agreements (DESTA) database, which covers 587 trade agreements for the period 1947-2010.

## C. Applications

This section highlights the usefulness, validity and applicability of the recommendations suggested in section B.2 by presenting a series of empirical applications estimating the impact of trade policies on trade, such as RTAs and MFN tariffs, within the structural gravity model. The purpose of these applications is primarily instructional. Therefore, the model specifications considered in each of the applications focus on the effects of specific covariates instead of specifying comprehensive sets of trade policy variables.

In order to emphasize the importance of the various considerations that should be taken into account when estimating the effects of trade policy, each application is presented as a sequence of estimating equations and corresponding results. For instructional purposes, examples of the main STATA commands used to implement each application are presented. Consistent with the recommendations formulated in section B.2, all estimation results are obtained with panel data with specific year intervals. In addition, standard errors in all estimations are clustered by trading pair in order to account for any intra-cluster correlations at the trading pair level.[18]

In all the applications presented in this chapter, the results are obtained from the same balanced panel data covering the aggregate manufacturing sector of 69 countries over the period 1986-2006.[19] The sample combines data from several sources. Most importantly, it includes consistently constructed international and intra-national trade flows data, which were assembled and provided by Thomas Zylkin. The original sources for the international trade data are the UN COMTRADE database and the CEPII TradeProd database. COMTRADE is the primary data source and TradeProd is used for instances when it includes positive flows for observations when no trade flows are reported in COMTRADE. Intra-national trade for each country is constructed as the difference between total manufacturing production and total manufacturing exports. Importantly, both of these variables are reported on a *gross basis*, which ensures consistency between intra-national and international trade. Three sources are used to construct the production data: the UN UNIDO INDSTAT database, the CEPII TradeProd database, and the World Bank's TPP database.[20] The data on RTAs were taken from Mario Larch's Regional Trade Agreements Database. Finally, all standard gravity variables including distance, contiguous borders, common language, and colonial ties are from the CEPII GeoDist database. An important advantage of the GeoDist database is that the weighted-average methods used to construct distance ensure consistency between the measures of intra-national and international distance, because each method uses population-weighted distances across the major economic centres within or across countries, respectively.

## 1. Traditional gravity estimates

This first application discusses the estimates of the effects of traditional gravity variables by applying different methods to account for these multilateral resistance terms and different estimators (OLS and PPML estimators).

### (a) OLS estimation ignoring multilateral resistance terms

The analysis begins with an OLS estimation of the empirical specification that includes standard gravity variables with panel data with 4-year intervals:

$$\begin{aligned}\ln X_{ij,t} = \beta_0 + \beta_1 \ln DIST_{ij} + \beta_2 CNTG_{ij} + \beta_3 LANG_{ij} + \beta_4 CLNY_{ij} + \\ \beta_5 \ln Y_{i,t} + \beta_6 \ln E_{j,t} + \varepsilon_{ij,t}\end{aligned} \quad (1\text{-}23)$$

As is standard in the literature, the variable $\ln X_{ij,t}$ corresponds to the logarithm of nominal bilateral international trade flows from exporter $i$ to importer $j$ at time $t$. $\beta_0$ is a constant term, whose structural interpretation is as world output. As defined in section B.2, $\ln DIST_{ij}$ represents the logarithm of bilateral distance between trading partners $i$ and $j$, $CNTG_{ij}$ is an indicator variable capturing the presence of contiguous borders between trading partners $i$ and $j$, $LANG_{ij}$ denotes a dummy variable for the existence of a common official language between partners $i$ and $j$, and $CLNY_{ij}$ is an indicator for the presence of colonial ties between countries $i$ and $j$. Finally, the covariates $\ln Y_{i,t}$ and $\ln E_{j,t}$ are the logarithms of the values of exporter output and importer expenditure, respectively.

As reported in column (1) of Table 1, the estimation results from specification (1-23) are overall as expected. With a $R^2 = 0.76$, the econometric specification delivers the standard strong fit that is commonly found in many empirical gravity models in the literature. The estimates on all covariates in equation (1-23) are statistically significant and have the expected signs. The estimate of the effect of distance is statistically significant at any conventional level and virtually equal to the benchmark estimate of –1, as documented by Disdier and Head (2008) and Head and Mayer (2014), confirming that distance is a significant impediment to bilateral trade. The impact of sharing a common border, speaking the same official language, and sharing colonial ties on international trade are positive and statistically significant, in line with the literature.[21] Overall, the gravity estimates obtained here are widely accepted in the literature and, therefore, establish the representativeness of the sample.

The estimates on output and expenditure are, as expected, positive and statistically significant. Although the estimates of both variables are very close to one, as predicted by the structural gravity model, both of them are statistically different from one. A possible explanation for this result is that both output and expenditure covariates may account for dynamic forces in the panel specification (Olivero and Yotov, 2012). Finally, in terms of magnitude, each of the estimates reported in column (1) of Table 1 is readily comparable to the corresponding summary indexes developed by Head and Mayer (2014).

**Table 1 Traditional gravity estimates**

| | (1) *OLS* | (2) *OLS Remoteness* | (3) *OLS Fixed Effects* | (4) *PPML Fixed Effects* |
|---|---|---|---|---|
| Log distance | –1.002 | –1.185 | –1.216 | –0.841 |
| | (0.027)** | (0.031)** | (0.038)** | (0.032)** |
| Contiguity | 0.574 | 0.247 | 0.223 | 0.437 |
| | (0.185)** | (0.177) | (0.203) | (0.083)** |
| Common language | 0.802 | 0.739 | 0.661 | 0.247 |
| | (0.082)** | (0.078)** | (0.082)** | (0.077)** |
| Colony | 0.735 | 0.842 | 0.670 | –0.222 |
| | (0.144)** | (0.150)** | (0.149)** | (0.116)+ |
| Log output | 1.190 | 1.164 | | |
| | (0.009)** | (0.010)** | | |
| Log expenditure | 0.908 | 0.903 | | |
| | (0.010)** | (0.010)** | | |
| Exporter remoteness index | | 0.972 | | |
| | | (0.068)** | | |
| Importer remoteness index | | 0.274 | | |
| | | (0.060)** | | |
| Constant | –11.283 | –35.219 | 1.719 | 15.867 |
| | (0.296)** | (1.986)** | (0.715)* | (0.214)** |
| Observations | 25689 | 25689 | 25689 | 28152 |
| $R^2$ | 0.759 | 0.765 | 0.843 | 0.614 |
| *Exporter-time* fixed effects | No | No | Yes | Yes |
| *Importer-time* fixed effects | No | No | Yes | Yes |
| RESET test (p-value) | 0.000 | 0.000 | 0.000 | 0.642 |

*Source:* Authors' calculations

*Notes:* All estimates are obtained with data for the years 1986, 1990, 1994, 1998, 2002, and 2006. Columns (1)-(3) use the OLS estimator. Column (1) does not control for the multilateral resistances. Column (2) uses "remoteness indexes" to control for multilateral resistances. Column (3) uses *importer-time* and *exporter-time* fixed effects, whose estimates are omitted for brevity, to control for multilateral resistances. Finally, column (4) employs the PPML estimator. Standard errors are clustered by country pair and are reported in parentheses. The *p*-values read as follows: $^{+}$ $p < 0.10$; $^{*}$ $p < 0.05$; and $^{**}$ $p < 0.01$.

```
* STATA commands to estimate standard gravity model with the OLS estimator and
* without intra-national trade flows:
    generate ln_trade = ln(trade)
    generate ln_DIST = ln(DIST)
    generate ln_Y = ln(Y)
    generate ln_E = ln(E)
    regress  ln_trade ln_DIST CNTG LANG CLNY ln_Y ln_E ///
             if exporter != importer, cluster(pair_id)
```

## (b) OLS estimation controlling for multilateral resistance terms with remoteness indexes

As famously demonstrated by Anderson and van Wincoop (2003), failure to account for the multilateral resistance terms may lead to severe biases in the estimates of the gravity variables. The following specification attempts to account for the multilateral resistances by considering the "remoteness indexes" mentioned in section B.2:[22]

$$\ln X_{ij,t} = \beta_0 + \beta_1 \ln DIST_{ij} + \beta_2 CNTG_{ij} + \beta_3 LANG_{ij} + \beta_4 CLNY_{ij} + \beta_5 \ln Y_{i,t} + \beta_6 \ln E_{j,t} + \beta_7 \ln REM\_EXP_{i,t} + \beta_8 \ln REM\_IMP_{j,t} + \varepsilon_{ij,t} \quad (1\text{-}24)$$

where the new covariates on the exporter side, $\ln REM\_EXP_{i,t}$, and on the importer side, $\ln REM\_IMP_{j,t}$, are constructed, respectively, as the logarithms of output- and expenditure-weighted averages of bilateral distance (Head, 2003):

$$\ln REM\_EXP_{i,t} = \ln\left(\sum_j DIST_{ij} \Big/ \frac{E_{j,t}}{Y_t}\right) \quad (1\text{-}25)$$

$$\ln REM\_IMP_{j,t} = \ln\left(\sum_i DIST_{ij} \Big/ \frac{Y_{i,t}}{Y_t}\right) \quad (1\text{-}26)$$

Estimates from specification (1-24) are reported in column (2) of Table 1. Three main findings stand out. First, the estimates of the effects of the standard gravity variables and the activity covariates are qualitatively identical to those from column (1). The only notable difference is that the estimate on contiguity is no longer statistically significant in the new specification. Second, the estimates of the effects of distance are stronger in column (2) than in column (1), while the estimates of the effects of contiguity and common official language are smaller. These results suggest that the estimates from column (1), which did not account for the multilateral resistances, were indeed biased as suggested by Anderson and van Wincoop (2003). Finally, in accordance with the literature, the estimates of the remoteness indexes are positive, large and highly significant, confirming that, all else equal, regions that are more isolated/remote from the rest of the world tend to trade more with each other.

```
* STATA commands to estimate gravity model with the OLS estimator and remoteness indexes:
    bysort exporter year: egen TotEj = total(E)
    bysort year: egen TotE = max(TotE)
    bysort exporter year: egen REM_EXP = total(DIST / (E / TotE))
         generate ln_REM_EXP = ln(REM_EXP)
    bysort importer year: egen TotYi  = total(Y)
    bysort year: egen TotOUT = max((TotYi)
    bysort importer year: egen REM_IMP = total(DIST / (Y / TotY))
         generate REM_IMP = ln(REM_IMP)
    regress  ln_trade ln_DIST CNTG LANG CLNY ln_Y ln_E ln_REM_EXP ///
             ln_REM_IMP if exporter != importer, cluster(pair_id)
```

## (c) OLS estimation controlling for multilateral resistance terms with fixed effects

Consistent with the recommendation formulated in section B.2, the gravity specification (1-23) is modified to account for the multilateral resistances with an appropriate set of *exporter-time* and *importer-time* fixed effects:

$$\ln X_{ij,t} = \pi_{i,t} + \chi_{j,t} + \beta_1 \ln DIST_{ij} + \beta_2 CNTG_{ij} + \beta_3 LANG_{ij} + \beta_4 CLNY_{ij} + \varepsilon_{ij,t} \qquad (1\text{-}27)$$

The term $\pi_{i,t}$ denotes the vector of *exporter-time* fixed effects, which will account for the outward multilateral resistances. Similarly, the vector $\chi_{j,t}$ denotes the set of *importer-time* fixed effects to capture the inward multilateral resistances. No constant term is included in the presence of the fixed effects. By definition, both *exporter-time* and *importer-time* fixed effects will absorb, respectively, the exporter value of output and importer expenditure, as well as all other observable and unobservable exporter- and importer-specific characteristics that may influence bilateral trade.

```
* STATA commands to estimate gravity model with the OLS estimator and exporter- and
* importer-time effects:
    egen exp_time = group(exporter year)
         tabulate exp_time, generate(EXPORTER_TIME_FE)
    egen imp_time = group(importer year)
         tabulate imp_time, generate(IMPORTER_TIME_FE)
    regress  ln_trade EXPORTER_TIME_FE* IMPORTER_TIME_FE* ln_DIST CNTG LANG CLNY ///
             if exporter != importer, cluster(pair_id)
```

The estimates from specification (1-27), reported in column (3) of Table 1, reinforce the message from the results in column (2), which only partially controlled for multilateral resistances. The estimate of the negative impact of distance on trade flows from column (3) is larger than the corresponding numbers from columns (1) and (2), while the estimates of the effects of contiguous borders and common official language decrease further relative to the results from columns (1) and (2). Overall, these results confirm the importance of accounting properly for multilateral resistances in order to obtain consistent gravity estimates.

### (d) PPML estimation controlling for multilateral resistance terms with fixed effects

Following the last recommendation suggested in section B.2, the gravity specification (1-27), which accounts for the full set of *exporter-time* and *importer-time* fixed effects, is reformulated in multiplicative form and re-estimated by applying the PPML estimator instead of the OLS estimator:

$$X_{ij,t} = \exp\left[\pi_{i,t} + \chi_{j,t} + \beta_1 \ln DIST_{ij} + \beta_2 CNTG_{ij} + \beta_3 LANG_{ij} + \beta_4 CLNY_{ij}\right] \times \varepsilon_{ij,t} \tag{1-28}$$

The PPML estimates from specification (1-28) listed in column (4) of Table 1 point to two important findings. First, comparison between the OLS estimates in column (3) and the PPML estimates in column (4) reveals significant differences in terms of magnitudes, significance, and even signs. Overall, and despite the different samples used, the results presented here are very similar to those reported in Table 5 of Santos Silva and Tenreyro (2006). Specifically, as compared to the estimated coefficients associated with the OLS estimation, the PPML estimate of the effect of distance is significantly smaller in absolute value. Similarly, the estimate of contiguous borders becomes statistically significant, and the estimate of common language decreases in magnitude but remains significant. Although the estimate of the effects of colony decrease in magnitude in both studies, it becomes negative and marginally significant here compared to Santos Silva and Tenreyro (2006). Second, and more important, the *p*-values of the Ramsey RESET test, reported at the bottom of Table 1, reveal that the PPML specification is the only one to pass the misspecification test.[23] Overall, these estimates favour the PPML estimator over the OLS estimator.

```
* STATA commands to estimate gravity model with the PPML estimator and exporter- and
* importer-time effects:
    ppml trade EXPORTER_TIME_FE* IMPORTER_TIME_FE* ln_DIST CNTG LANG CLNY ///
         if exporter != importer, cluster(pair_id)
```

## 2. The “distance puzzle” resolved

Despite its popularity and great predictive power, the gravity model has been subject to significant criticism on the ground that gravity estimates fail to capture the effects of globalization on international trade. Based on a meta-analysis of a rich data set of 1,467 distance estimates from gravity equations from 103 papers, Disdier and Head (2008) conclude that the estimated negative impact of distance on trade has remained persistently high, even after controlling for many important differences in samples and methods.[24] This finding, known as the “distance puzzle” in international trade, is in direct contradiction with the empirical evidence of declining trade-related costs (Coe et al., 2002).

This application applies the methods developed by Yotov (2012) in order to solve the “distance puzzle” of trade. In particular, capitalizing on the properties of the structural gravity model, Yotov (2012) recognizes that the structural gravity system can only ever identify relative trade costs. Therefore, studies that only use international trade data cannot resolve the distance puzzle, because the effects of distance on international trade are measured relative to other international trade costs. Yotov (2012) proposes to measure the effects of distance and globalization relative to internal trade costs and demonstrates that the "distance puzzle" disappears when, consistent with gravity theory, *internal trade* and *internal distance* are explicitly accounted for in the standard gravity specification. In fact, an empirical model allowing for a decrease in international trade costs relative to *internal* trade costs is more likely to capture the effects of globalization than a model analysing the impact of trade costs relative to a reference group that has been affected similarly (equally) by globalization.

For expositional clarity and instructional purposes, the analysis is presented sequentially. The first set of results capture the "distance puzzle" as described in the literature. The following results address the "distance puzzle" and reproduce some of the estimates found in Borchert and Yotov (2016).

## (a) Uncovering the "distance puzzle"

The analysis starts with an OLS estimation of the gravity model with 4-year interval data. The empirical specification includes traditional gravity covariates, including *exporter-time* and *importer-time* fixed effects, and considers only international trade flows (i.e. for $i \neq j$), as is standard in the literature:

$$\ln X_{ij,t} = \pi_{i,t} + \chi_{j,t} + \sum_{T=1986}^{2006} \beta_T \ln DIST\_(T)_{ij} + \beta_2 CNTG_{ij} + \beta_3 LANG_{ij} + \beta_4 CLNY_{ij} + \varepsilon_{ij,t} \quad (1\text{-}29)$$

In order to determine the change in the impact of the distance variable on trade, the model specification allows for different effects of distance in each of the six years, considered in the analysis $T \in \{1986, 1990, 1994, 1998, 2002, 2006\}$. As highlighted in column (1) of Table 2, the negative impact of distance on bilateral trade has actually increased by 7.95 percent between 1986 and 2006, confirming the presence of the "distance puzzle" in the sample.

In accordance with the last recommendation suggested in section B.2, the gravity specification (1-29) is re-estimated in multiplicative form by applying the PPML estimator to the same sample (with international trade only):

$$X_{ij,t} = \exp\left[\pi_{i,t} + \chi_{j,t} + \sum_{T=1986}^{2006} \beta_T \ln DIST\_(T)_{ij} + \beta_2 CNTG_{ij} + \beta_3 LANG_{ij} + \beta_4 CLNY_{ij}\right] \times \varepsilon_{ij,t} \quad (1\text{-}30)$$

As reported in column (2) of Table 2, the estimate of the effect of distance in 2006 is only marginally smaller in absolute value than the one associated with the distance variable for 1986. Yet, the –2.75 percent change in the distance estimate between 1986 and 2006 is statistically not different from zero, confirming once again that the data is subject to the "distance puzzle".

## (b) Solving the "distance puzzle"

Following the recommendations proposed in Section B.2, the gravity specification (1-30) is modified to consider *international* and *intra-national* trade data, and to include a measure of *intra-national* distance, $\ln DIST\_INTRA_{ii}$, taking the value of zero for international trade flows (for $i \neq j$):

$$X_{ij,t} = \exp\left[\pi_{i,t} + \chi_{j,t} + \sum_{T=1986}^{2006} \beta_T \ln DIST\_(T)_{ij} + \beta_2 CNTG_{ij} + \beta_3 LANG_{ij}\right] \times \exp\left[\beta_4 CLNY_{ij} + \beta_5 \ln DIST\_INTRA_{ii}\right] \times \varepsilon_{ij,t} \quad (1\text{-}31)$$

**Table 2 A simple solution to the "distance puzzle" in trade**

| | (1) *OLS* | (2) *PPML* | (3) *INTRA* | (4) *BRDR* | (5) *FEs* |
|---|---|---|---|---|---|
| Log distance 1986 | –1.168<br>(0.044)** | –0.859<br>(0.037)** | –0.980<br>(0.072)** | –0.857<br>(0.063)** | –0.910<br>(0.032)** |
| Log distance 1990 | –1.155<br>(0.042)** | –0.834<br>(0.038)** | –0.940<br>(0.073)** | –0.819<br>(0.063)** | –0.879<br>(0.032)** |
| Log distance 1994 | –1.211<br>(0.046)** | –0.835<br>(0.035)** | –0.915<br>(0.072)** | –0.796<br>(0.063)** | –0.860<br>(0.032)** |
| Log distance 1998 | –1.248<br>(0.043)** | –0.847<br>(0.035)** | –0.887<br>(0.071)** | –0.770<br>(0.063)** | –0.833<br>(0.032)** |
| Log distance 2002 | –1.241<br>(0.044)** | –0.848<br>(0.032)** | –0.884<br>(0.071)** | –0.767<br>(0.063)** | –0.829<br>(0.032)** |
| Log distance 2006 | –1.261<br>(0.044)** | –0.836<br>(0.031)** | –0.872<br>(0.071)** | –0.754<br>(0.062)** | –0.811<br>(0.032)** |
| Contiguity | 0.223<br>(0.203) | 0.437<br>(0.083)** | 0.371<br>(0.140)** | 0.574<br>(0.155)** | 0.442<br>(0.082)** |
| Common language | 0.661<br>(0.082)** | 0.248<br>(0.077)** | 0.337<br>(0.168)* | 0.352<br>(0.137)* | 0.241<br>(0.076)** |
| Colony | 0.670<br>(0.149)** | –0.222<br>(0.116)+ | 0.019<br>(0.156) | 0.027<br>(0.125) | –0.220<br>(0.117)+ |
| Log intra-national distance | | | –0.488<br>(0.101)** | –0.602<br>(0.109)** | |
| Intra-national trade dummy | | | | 1.689<br>(0.574)** | |
| Observations | 25689 | 28152 | 28566 | 28566 | 28566 |
| Percent change in log distance between 1986 and 2006 | 7.950<br>(3.759)* | –2.750<br>(3.004) | –10.965<br>(1.058)** | –11.969<br>(1.173)** | –10.931<br>(0.769)** |
| *Intra-national* trade | No | No | Yes | Yes | Yes |
| Country-specific *intra-national* fixed effects | No | No | No | No | Yes |

*Source:* Authors' calculations

*Notes:* All estimates are obtained with data for the years 1986, 1990, 1994, 1998, 2002, and 2006, and use *exporter-time* and *importer-time* fixed effects. The estimates of the fixed effects are omitted for brevity. Columns (1) and (2) use data on international trade flows only. Column (1) employs the OLS estimator and column (2) uses the PPML estimator. Column (3) adds *internal* trade observations and uses *intra-national* distance as an additional covariate. Column (4) adds an indicator covariate for international trade. Finally, column (5) uses country-specific dummies for *intra-national* trade. Standard errors are clustered by country pair and are reported in parentheses. The bottom panel of the table reports the percentage change in the estimates of the effects of bilateral distance between 1986 and 2006. Standard errors for the percentage changes are obtained with the delta method. The *p*-values read as follows: $^{+}$ $p < 0.10$; $^{*}$ $p < 0.05$; and $^{**}$ $p < 0.01$.

Two main results stand out from the PPML estimates of the gravity specification (1-31) reported in column (3) of Table 2. First, the impact of internal distance on domestic sales is much smaller as compared to the distance effects on international trade (where by definition the log of distance for intra-national trade is zero). This is consistent with the estimates reported in Anderson et al. (2016c) for the effects of intra-provincial vs. international distance in the case of Canada, confirming Head and Mayer's (2013) argument that international distance accounts for a host of obstacles to trade. Second, and more important for the current purposes, the results show a statistically significant decrease of -10.965 percent of the effects of distance on trade between 1986 and 2006, solving thus the "distance puzzle".

```
* STATA commands to estimate gravity model with international and intra-national trade:
    * Create variable for the log of internal distance
    generate SMCTRY = 1 if exporter == importer
        replace SMCTRY = 0 if SMCTRY == .
    generate ln_DIST_INTRA = ln_DIST*SMCTRY
    forvalues Year = 1986(1)2006 {
        replace ln_DIST_`Year' = 0 if SMCTRY == 1
    }
    * Estimate the gravity model
    ppml trade EXPORTER_TIME_FE* IMPORTER_TIME_FE* ln_DIST_1986 ln_DIST_1990 ///
        ln_DIST_1994 ln_DIST_1998 ln_DIST_2002 ln_DIST_2006 CNTG LANG CLNY ///
        ln_DIST_INTRA, cluster(pair_id)
```

Next, the gravity specification (1-31) is modified to better account for potential forces affecting international relative to internal trade in addition to distance (Borchert and Yotov, 2016):[25]

$$X_{ij,t} = \exp\left[\pi_{i,t} + \chi_{j,t} + \sum_{T=1986}^{2006} \beta_T \ln DIST\_(T)_{ij} + \beta_2 CNTG_{ij} + \beta_3 LANG_{ij} + \beta_4 CLNY_{ij}\right] \times \exp\left[\beta_5 \ln DIST\_INTRA_{ii} + \beta_6 SMCTRY_{ii}\right] \times \varepsilon_{ij,t} \quad (1\text{-}32)$$

The additional dummy variable $SMCTRY_{ii}$, taking the value of one for *intra-national* trade and zero for *international* trade, is motivated by three reasons. First, the covariate $SMCTRY_{ii}$ enables to distinguish home bias effects and the fact that domestic trade tends to be much larger than international trade. Second, the variable $SMCTRY_{ii}$ can potentially capture any other effects affecting international trade differentially that have not been covered by the other covariates of the model. Finally, as noted by Anderson and Yotov (2010a), the dummy variable $SMCTRY_{ii}$ has the advantage of being an exogenous variable that controls for all the relevant forces that discriminate between *intra-national* and *international* trade. Consistent with the standard treatment of internal trade costs in the trade literature, a common coefficient is assigned to the variable $SMCTRY_{ii}$, which can be interpreted as setting the elasticity of *intra-national* trade costs to be equal across countries in the sample.

Three findings stand out from the PPML estimates of the gravity specification (1-32) listed in column (4) of Table 2. First, as expected, the impact of the variable $SMCTRY_{ii}$ is large, positive, and statistically significant, suggesting a significant "home bias" with *intra-national* trade about exp(1.689) = 5.5 times larger than international trade. This estimate is significantly smaller compared to the famous border estimate of 22 for inter-provincial trade within Canada relative to international trade between Canadian provinces and US states reported in McCallum (1995). The proper econometric specification of the structural gravity model (i.e. controlling for the multilateral resistance terms as suggested by Anderson and van Wincoop (2003)) may explain this result. Second, although the impact of international distance and internal distance, respectively, falls and increases, in absolute magnitude, they both converge toward each other in terms of magnitude. The intuition for this result is twofold. First, international distance has indeed been capturing more than just the effects of transportation costs. Second, the estimate on internal distance reported in column (3) has also been capturing "home bias" effects. Finally, and most important, the effects of distance have decreased by 11.969 percent between 1986 and 2006, implying that the "distance puzzle" has once again disappeared.

Following the recommendations formulated in Section B.2, the last and most comprehensive gravity specification considered for the purpose of this application is modified to include country-specific fixed effects for *intra-national* trade ($\mu_{ii}$):

$$X_{ij,t} = \exp\left[\mu_{ii} + \pi_{i,t} + \chi_{j,t} + \sum_{T=1986}^{2006} \beta_T \ln DIST\_(T)_{ij} + \beta_2 CNTG_{ij} + \beta_3 LANG_{ij} + \beta_4 CLNY_{ij}\right] \times \varepsilon_{ij,t} \qquad (1\text{-}33)$$

The country-specific fixed effects $\mu_{ii}$ are defined as dummy variables taking the value of one for *intra-national* trade and zero otherwise. As such, due to perfect collinearity, they will absorb the *intra-national* distance and trade variables ($\ln DIST\_INTRA_{ii}$ and $SMCTRY_{ii}$). Therefore, the fixed effects $\mu_{ii}$ control for country-specific intra-national trade costs and "home-bias" effects, as well as any other country-specific time-invariant characteristics that may drive a wedge between internal and international trade.

As reported in column (5) of Table 2, the PPML estimation results of specification (1-33) reveal that the effects of distance are smaller compared to the estimates from the previous specification that accounts for internal distance. This finding confirms the fact that the estimates of the effects of distance in standard gravity regressions reflect more than just the effects of transportation costs. The results also confirm the absence of the "distance puzzle" with a statistically significant reduction of 10.931 percent of the effects of distance between 1986 and 2006.

```
* STATA commands to estimate gravity model with intra-national trade fixed effects:
     egen intra_pair = group(exporter) if exporter == importer
          replace intra_pair = 0 if intra_pair == .
     tabulate intra_pair, generate(INTRA_FE)
     ppml trade INTRA_FE* EXPORTER_TIME_FE* IMPORTER_TIME_FE* ln_DIST_1986 ///
          ln_DIST_1990 ln_DIST_1994 ln_DIST_1998 ln_DIST_2002 ln_DIST_2006 CNTG ///
          LANG CLNY, cluster(pair_id)
```

## 3. Regional trade agreements effects

In the last 25 years, the number of RTAs has increased more than four-fold, to more than 450 agreements notified to the WTO. In this context, the objective of this application is to obtain estimates of the effects of RTAs on trade.[26] The analysis starts with a basic OLS specification. Then, capitalizing on various contributions from the literature, each additional step introduces a new feature to the initial specification.

### (a) Traditional estimates of RTAs

The analysis starts with an OLS estimation of the gravity model with 4-year interval data. The empirical specification includes traditional gravity covariates, including *exporter-time* and *importer-time* fixed effects, and considers only international trade flows (i.e. for $i \neq j$), as is standard in the literature:

$$\ln X_{ij,t} = \pi_{i,t} + \chi_{j,t} + \beta_1 \ln DIST_{ij} + \beta_2 CNTG_{ij} + \beta_3 LANG_{ij} + \beta_4 CLNY_{ij} + \beta_5 RTA_{ij,t} + \varepsilon_{ij,t} \qquad (1\text{-}34)$$

The covariate $RTA_{ij,t}$ is a dummy variable that takes a value of one if countries *i* and *j* are partners in a RTA at time *t*, and zero otherwise. Two main findings of the OLS estimates reported in Table 3 stand out. First, the estimates of the effects of the standard gravity variables are in accordance with theory and consistent with the findings reported in the previous applications. Second, interestingly, the results suggest that RTAs play no statistically significant role in promoting international trade. One possible explanation for the small (in fact negative) and not statistically significant coefficient of the variable *RTA* could be that specification (1-34) does not account properly for the potential endogeneity of RTAs. This issue is addressed in subsequent specifications.

But first, following the last recommendation suggested in section B.2, the gravity specification (1-34) is re-formulated in multiplicative form and re-estimated by applying the PPML estimator to the same sample (with international trade only):

$$\ln X_{ij,t} = \exp\left[\pi_{i,t} + \chi_{j,t} + \beta_1 \ln DIST_{ij} + \beta_2 CNTG_{ij} + \beta_3 LANG_{ij} + \beta_4 CLNY_{ij} + \beta_5 RTA_{ij,t}\right] \times \varepsilon_{ij,t} \qquad (1\text{-}35)$$

As reported in column (2) of Table 3, the PPML estimates of the standard gravity variables are virtually identical to the corresponding numbers from specification (1-28), which does not control for the covariate *RTA* and is listed in column (4) of Table 1. This finding suggests that the omission of the variable *RTA* has not heavily biased the estimates of the model specifications considered in the first application. In addition, the positive and significant estimate of the effects of RTAs ($\hat{\beta}_5 = 0.191$) suggests that, all else equal, RTAs increase trade between member countries by about $\left[\exp\left(\hat{\beta}_5\right) - 1\right] \times 100 = \left[\exp(0.191) - 1\right] \times 100 = 21$ percent. Although the estimated coefficient of the variable *RTA* is positive and statistically significant, it is significantly smaller than corresponding numbers found in the literature (Baier and Bergstrand, 2007; Anderson and Yotov, 2016).

### (b) Allowing for trade-diversion from domestic sales

Following Dai et al. (2014) and Anderson and Yotov (2016), the gravity specification (1-35) is re-estimated by expanding the sample to include *intra-national* trade flows data in addition to *international* trade flows. The idea is that RTAs may be diverting trade from domestic to international sales and, therefore, the estimates of the variable *RTA* that are based on *international* trade only may be biased downward. As reported in column (3) of Table 3, the estimates of the standard gravity variables based on the sample with *international* and *intra-national* trade are statistically not different from the corresponding estimated parameters based on the sample with *international* trade only and listed in column (2) of Table 3. However, and more importantly, the results in column (3) show that extending the sample to include *intra-national* trade increases the estimated effect of RTAs, which has more than doubled in magnitude (from $\hat{\beta}_5 = 0.191$ to $\hat{\beta}_5 = 0.409$). This finding supports the hypothesis that RTAs enhance trade between members at the expense of domestic sales.

**Table 3** Estimating the Effects of Regional Trade Agreements

| | (1) *OLS* | (2) *PPML* | (3) *INTRA* | (4) *ENDG* | (5) *LEAD* | (6) *PHSNG* | (7) *GLBZN* |
|---|---|---|---|---|---|---|---|
| Log distance | –1.216 | –0.822 | –0.800 | | | | |
| | (0.039)** | (0.031)** | (0.030)** | | | | |
| Contiguity | 0.223 | 0.416 | 0.393 | | | | |
| | (0.203) | (0.083)** | (0.079)** | | | | |
| Common language | 0.661 | 0.250 | 0.244 | | | | |
| | (0.082)** | (0.077)** | (0.077)** | | | | |
| Colony | 0.670 | –0.205 | –0.182 | | | | |
| | (0.149)** | (0.114)+ | (0.113) | | | | |
| RTA | –0.004 | 0.191 | 0.409 | 0.557 | 0.520 | 0.291 | 0.116 |
| | (0.054) | (0.066)** | (0.069)** | (0.102)** | (0.086)** | (0.089)** | (0.087) |
| RTA($t+4$) | | | | | 0.077 | | |
| | | | | | (0.092) | | |
| RTA($t-4$) | | | | | | 0.414 | 0.288 |
| | | | | | | (0.067)** | (0.062)** |
| RTA($t-8$) | | | | | | 0.169 | 0.069 |
| | | | | | | (0.043)** | (0.048) |
| RTA($t-12$) | | | | | | 0.119 | 0.002 |
| | | | | | | (0.030)** | (0.029) |
| International border 1986 | | | | | | | –0.706 |
| | | | | | | | (0.048)** |
| International border 1990 | | | | | | | –0.480 |
| | | | | | | | (0.043)** |
| International border 1994 | | | | | | | –0.367 |
| | | | | | | | (0.033)** |
| International border 1998 | | | | | | | –0.158 |
| | | | | | | | (0.023)** |
| International border 2002 | | | | | | | –0.141 |
| | | | | | | | (0.017)** |
| Observations | 25689 | 28152 | 28566 | 28482 | 28482 | 28482 | 28482 |
| Total RTA effect | | | | | | 0.992 | 0.475 |
| | | | | | | (0.094)** | (0.109)** |
| Intra-national trade | No | No | Yes | Yes | Yes | Yes | Yes |

*Source:* Authors' calculations

*Notes:* All estimates are obtained with data for the years 1986, 1990, 1994, 1998, 2002, and 2006, and use exporter-time and importer-time fixed effects. The estimates of the fixed effects are omitted for brevity. Columns (1) and (2) use data on *international* trade flows only. Column (1) applies the OLS estimator and column (2) uses the PPML estimator. Column (3) adds *intra-national* trade observations and uses country-specific dummies for internal trade. Column (4) adds pair fixed effects. The estimates of the pair fixed effects are omitted for brevity. Column (5) introduces RTA lead. Column (6) allows for phasing-in effects of RTAs. Finally, column (7) accounts for the effects of globalization. Standard errors are clustered by country pair and are reported in parentheses. The *p*-values read as follows: $^{+}$ $p<0.10$; $^{*}$ $p<0.05$; and $^{**}$ $p<0.01$.

### (c) Addressing potential endogeneity of RTAs

As noted previously, failure to address the potential endogeneity of RTAs may bias the gravity estimates. Following Baier and Bergstrand (2007), the gravity specification (1-35) is modified to include pair fixed effects ($\mu_{ij}$) in addition to the theoretically-motivated *importer-time* and *exporter-time* fixed effects:

$$X_{ij,t} = \exp\left[\pi_{i,t} + \chi_{j,t} + \mu_{ij} + \beta_5 RTA_{ij,t}\right] \times \varepsilon_{ij,t} \tag{1-36}$$

Because of perfect collinearity, using pair fixed effects ($\mu_{ij}$) does not allow to include in the model, and therefore estimate, any of the standard gravity variables that do not vary over time (distance, contiguity, common language and colonial ties). In addition, one of the bilateral fixed effects has to be dropped from the model specification. For practical purposes, the fixed effect for *intra-national* trade $\mu_{ii}$, captured by the variable $SMCTRY_{ii}$ defined in the application 2, is removed from specification (1-36).[27] In effect, this implies that all internal trade costs are set to one and all *international* fixed effects $\mu_{ij}, j \neq i$, are estimated relative to the *intra-national* fixed effect $\mu_{ii}$.

```
* STATA commands to estimate gravity model with the country-pair fixed effects:
    egen pair_id = group(exporter importer)
        tabulate pair_id, generate(PAIR_FE)
    ppml trade PAIR_FE* EXPORTER_TIME_FE* IMPORTER_TIME_FE* RTA, cluster(pair_id)
```

The PPML estimation results, which are obtained with pair fixed effects, are reported in column (4) of Table 3. Although not presented in column (4), the estimates of all pair fixed effects are negative and smaller than -1, reflecting the fact that the pair fixed effects absorb all trade costs and that *international* trade costs are larger than *intra-national* trade costs. More importantly for the purpose of the application, the coefficient of the variable *RTA* is statistically significant and positive, and much larger ($\hat{\beta}_5 = 0.557$) than the estimated coefficient obtained with the previous specifications. The positive and highly significant estimate of the effects of RTAs is in accordance with Baier and Bergstrand's (2007) predictions that the estimates of the RTAs impact on trade obtained without proper account for endogeneity are biased downward.[28] The estimated coefficient of the variable *RTA* reported in column (4) suggests that, all else equal, the formation of RTAs leads to an average increase of about $[\exp(0.577) - 1] \times 100 = 75$ percent in international trade between members, which is much closer to existing estimates from the literature.

### (d) Testing for potential "reverse causality" between trade and RTAs

In order to test whether specification (1-36) has properly accounted for possible "reverse causality" between trade and RTAs through the pair fixed effects, an easy test can be implemented to assess the "strict exogeneity" of RTAs by adding a new variable capturing the future level of RTAs, $RTA_{ij,\,t+4}$, to specification (1-36) (Wooldridge, 2010; Baier and Bergstrand, 2007):

$$X_{ij,t} = \exp\left[\pi_{i,t} + \chi_{j,t} + \mu_{ij} + \beta_5 RTA_{ij,t} + \beta_6 RTA_{ij,t+4}\right] \times \varepsilon_{ij,t} \tag{1-37}$$

If RTAs are exogenous to trade flows, the parameter $\beta_6$ associated with the variable $RTA_{ij,t+4}$ should be statistically not different from zero. As shown in column (5) of Table 3, the PPML estimate of the "future lead" of RTAs is neither economically nor statistically different from zero, confirming the absence of "reverse causality" in the results associated with specification (1-36).

```
* STATA commands to estimate gravity model with lead-in RTA variable:
    tsset pair_id year
    generate RTA_LEAD4 = f4.RTA
         replace RTA_LEAD4 = 0 if RTA_LEAD4 == .
    ppml trade PAIR_FE* EXPORTER_TIME_FE* IMPORTER_TIME_FE* RTA ///
         RTA_LEAD4, cluster(pair_id)
```

## (e) Allowing for potential non-linear and phasing-in effects of RTAs

In order to allow for non-linear effects of RTAs and/or to capture the possibility that the effects of RTAs change over time, specification (1-36) is further modified to include various lags (up to 12 years) of the variable *RTA*:

$$X_{ij,t} = \exp\left[\pi_{i,t} + \chi_{j,t} + \mu_{ij} + \beta_5 RTA_{ij,t} + \beta_6 RTA_{ij,t-4} + \beta_7 RTA_{ij,t-8} + \beta_8 RTA_{ij,t-12}\right] \times \varepsilon_{ij,t} \quad (1\text{-}38)$$

As highlighted in column (6) of Table 3, the estimated coefficients of the three lagged *RTA* variables point to strong phasing-in effects of RTAs, which is consistent with findings from existing related studies (Baier and Bergstrand, 2007; Anderson and Yotov, 2011). In particular, the results suggest a non-monotonic relationship, where the relatively small average effects of RTAs over the first four years after the RTAs' entry into force more than double in the second four-year period, and decrease almost three times as compared to their peak after twelve years. That being said, the effects of RTAs remain significant twelve years after their implementation, which explains why the overall RTA effect, reported at the bottom of column (6) of Table 3, is strong and statistically significant.

```
* STATA commands to estimate gravity model with lagged RTA variables and compute the total
* RTA effects and associated standard errors with delta method:
    tsset pair_id year
    forvalues t = 4(4)12 {
         generate RTA_LAG`t' = L`t'.RTA
         replace RTA_LAG`t' = 0 if RTA_LAG`t' == .
    }
    ppml trade PAIR_FE* EXPORTER_TIME_FE* IMPORTER_TIME_FE* RTA RTA_LAG4 RTA_LAG8 ///
         RTA_LAG12, cluster(pair_id)
    lincom _b[RTA]+_b[RTA_LAG4]+_b[RTA_LAG8]+_b[RTA_LAG12]
```

## (f) Addressing globalization effects

The final experiment applies the methods developed by Bergstrand et al. (2015) to test and account for the possibility that the estimated effects of RTAs from specification (1-38) may be biased upward because they capture globalization effects, such as technology and innovation. Specifically, specification (1-38) is adjusted to include a set of new indicator variables capturing the existence of international borders between countries *i* and *j* for each year *T*:

$$X_{ij,t} = \exp\left[\pi_{i,t} + \chi_{j,t} + \mu_{ij} + \beta_5 RTA_{ij,t} + \beta_6 RTA_{ij,t-4} + \beta_7 RTA_{ij,t-8} + \beta_8 RTA_{ij,t-12}\right] \times \exp\left[\sum_{T=1986}^{2002} \beta_T INTL\_BRDR\_(T)_{ij}\right] \times \varepsilon_{ij,t} \quad (1\text{-}39)$$

The new covariate, $INTL\_BRDR\_(T)_{ij}$, is a dummy variable taking the value of one for international trade for each year $T$, and zero otherwise. Because of perfect collinearity with the rest of the fixed effects included in specification (1-39), it is impossible to estimate these international border dummies for all the years in the sample. For practical purposes, the international border dummy for 2006, *INTL_BRDR_*2006, is dropped from specification (1-39). As a result, the estimated coefficients of the other border dummy variables $INTL\_BRDR\_(T)$ for the years $T \in \{1986, 1990, 1994, 1998, 2000, 2002\}$, should be interpreted relative to the corresponding estimate for 2006.

```
* STATA commands to estimate gravity model with international border variables:
    generate INTL_BRDR = 1 if exporter != importer
         replace INTL_BRDR = 0 if INTL_BRDR == .
    forvalues Year = 1986(1)2006 {
         generate INTL_BRDR_`Year' = 1 if INTL_BRDR == 1 & year == `Year'
             replace INTL_BRDR_`Year' = 0 if INTL_BRDR_`Year' == .
    }
    ppml trade PAIR_FE* EXPORTER_TIME_FE* IMPORTER_TIME_FE* RTA RTA_LAG4 RTA_LAG8 ///
         RTA_LAG12 INTL_BRDR_1986 INTL_BRDR_1990 INTL_BRDR_1994 INTL_BRDR_1998 ///
         INTL_BRDR_2002, cluster(pair_id)
    lincom _b[RTA]+_b[RTA_LAG4]+_b[RTA_LAG8]+_b[RTA_LAG12]
```

Two main findings stand out from the estimates of specification (1-39) reported in column (7) of Table 3. First, the estimated coefficients of the different RTAs variables remain positive, even though they all decrease in magnitude. Furthermore, only the estimate of the first lagged *RTA* variable ($RTA_{ij,t-4}$) remains statistically significant. This result suggests that, once globalization forces are accounted for, not only the impact of RTAs takes time to show up in the data but it also phases-in faster. This explains why the total estimated RTA effects, reported at the bottom of column (7), are slashed in half once globalization forces are explicitly taken into account in the model specification. This result, consistent with Bergstrand et al. (2015), suggests that the estimates of RTAs in the previous specifications may have indeed captured the effects of globalization.

Second, the estimates of the international border variables reveal that borders have fallen significantly over time. To see this point, note that the border estimates should be interpreted as deviations from the international border effect in 2006, defined as the reference group. For example, the estimated coefficient of the dummy variable *INTL_BRDR_*1986 suggests that the effects of borders on trade in 1986 were exp(0.706) = 2.03 larger than the corresponding effects in 2006. Overall, the estimates of the trend in the international border dummies are similar to those reported in Bergstrand et al. (2015) confirming a steady and strong effect of globalization on trade over time.

# D. Exercises

## 1. Estimating the effects of WTO accession

The aim of this exercise is to assess the impact on trade of the accession to the WTO. A similar exercise can be applied to any other trade agreement provided the sample period covers sufficient years before and after the agreement. The STATA do-file "WTOAccession.do" providing the solution to the exercise can be found in "Chapter1\Exercises\".

(i) Preliminaries

a. Open the data file "Chapter1Exercise1.dta".
b. Create a histogram reporting the frequency of the number of the member countries of the WTO by year of accession.
*Hints: hist*

(ii) Benchmark gravity estimation

a. Generate exporter-time and importer-time effects.
*Hints: generate*
b. Estimate the following standard gravity specification by considering only *international* trade flows (i.e. for $i \neq j$) and applying the OLS estimator:

$$\ln X_{ij,t} = \pi_{i,t} + \chi_{j,t} + \boldsymbol{GRAVITY} \times \boldsymbol{\beta} + \beta_5 RTA_{ij,t} + \beta_6 WTO_{i,t} + \varepsilon_{ij,t}$$

where the vector ***GRAVITY*** includes the log of the distance and dummy variables for contiguity, common language, and colonial ties.
*Hints: regress*
c. Re-estimate the same specification expressed in multiplicative form with the PPML estimator. Compare the results and comment.
*Hints: ppml*

(iii) Gravity estimation accounting for *intra-national* trade and potential endogeneity

a. Generate pair fixed effects.
b. Re-estimate the specification presented above with the PPML estimator but this time by considering *international* and *intra-national* trade. Compare the results and comment.
c. In order to correct for potential endogeneity of the RTAs variable, estimate the following gravity specification with the PPML estimator:

$$X_{ij,t} = \exp\left[\pi_{i,t} + \chi_{j,t} + \mu_{ij} + \beta_5 RTA_{ij,t} + \beta_6 WTO_{i,t}\right] \times \varepsilon_{ij,t}$$

(iv) Gravity estimation accounting for globalization

a. Generate international border dummies for the years 1986, 1990, 1994, 1998 and 2002.
b. Estimate with the PPML estimator the following structural gravity specification:

$$X_{ij,t} = \exp\left[\pi_{i,t} + \chi_{j,t} + \mu_{ij} + \beta_5 RTA_{ij,t} + \beta_6 WTO_{i,t} + \sum_{T=1986}^{2002} \beta_T INTL\_BRDR\_(T)_{ij,t}\right] \times \varepsilon_{ij,t}$$

Compare the results and comment.

## 2. Estimating the effects of unilateral trade policy

The aim of this exercise is to demonstrate that the gravity model can be used to estimate the effects of non-discriminatory (across trading partners) trade policies, such as MFN tariffs. The exercise follows the approach developed by Heid et al. (2015). The STATA do-file "UnilateralTradePolicy.do" providing the solution to the exercise can be found in "Chapter1\Exercises\".

(i) Preliminaries

a. Open the data file "Chapter1Exercise2.dta".

b. Determine for how many countries and years data on MFN tariffs are available.
*Hints: keep, duplicates*

(ii) Benchmark gravity estimation

a. Generate exporter-time, importer-time and pair fixed effects.
*Hints: generate*

b. Estimate with the PPML estimator the following structural gravity specification:

$$X_{ij,t} = \exp\left[\pi_{i,t} + \chi_{j,t} + \mu_{ij} + \sum_{T=1}^{4} \beta_T RTA_{ij,t-3\times(T-1)}\right] \times \varepsilon_{ij,t}$$

*Hints: ppml*

c. Compute the total effects of the RTAs and comment.
*Hints: lincom*

(iii) Gravity estimation with unilateral trade policy

a. Create the logarithm of the MFN tariffs variable (ln_*MFN*), and replace the missing values to be equal to zero.

b. Estimate with the PPML estimator the following structural gravity specification and compare the results:

$$X_{ij,t} = \exp\left[\pi_{i,t} + \chi_{j,t} + \mu_{ij} + \sum_{T=1}^{4} \beta_T RTA_{ij,t-3\times(T-1)} + \beta_5 \ln\_MFN_{j,t} \times INTL_{ij}\right] \times \varepsilon_{ij,t}$$

c. Compute the total effects of the RTAs, compare the result and comment.

d. Compute the trade elasticity of substitution based on the estimates obtained in 3.b. Discuss the result, noting that the elasticity estimates from the related trade literature usually vary between 2 and 12 (Eaton and Kortum, 2002; Anderson and van Wincoop, 2003; Broda et al., 2006).

# Appendices

## Appendix A: Structural gravity from supply side

In Section B.1, the structural gravity framework was derived from the demand side assuming an Armington (1969) setting with CES preferences. This appendix demonstrates that an isomorphic structural gravity framework can be derived from the supply side. The reader may refer to Anderson (2011) for a discussion of demand-side and supply-side gravity foundations. This derivation is based on the most influential supply-side model, the Ricardian model of international trade by Eaton and Kortum (2002).[29]

Consumer preferences are still assumed to be homothetic, globally common/identical across countries, and approximated by a Constant Elasticity of Substitution (CES) utility function:

$$U_j = \left[ \int_0^1 c(l)^{\frac{\sigma-1}{\sigma}} dl \right]^{\frac{\sigma}{\sigma-1}} \tag{1.A.1}$$

where $j$ denotes the country and $\sigma$ is the elasticity of substitution among different varieties.

Following Eaton and Kortum (2002), there is a continuum of goods $l \in [0,1]$, with consumption of individual goods denoted by $c(l)$. In addition, trade of goods from country $i$ to country $j$ imposes iceberg trade costs $t_{ij} > 1$. In contrast to the baseline framework, countries now differ in the efficiency with which they can produce goods. Let $z_i(l)$ denote country $i$'s efficiency in producing good $l \in [0,1]$. Then, with constant returns to scale the cost of producing a unit of good $l$ in country $i$ is $\varsigma_i / z_i(l)$, with $\varsigma_i$ denoting the input costs in country $i$. Taking iceberg trade costs into account, delivering a unit of good $l$ produced in country $i$ to country $j$ costs:

$$p_{ij}(l) = \left( \frac{\varsigma_i}{z_i(l)} \right) t_{ij} \tag{1.A.2}$$

With perfect competition, $p_{ij}(l)$ is the price consumers in country $j$ would pay if they decide to buy good $l$ from country $i$. In the presence of international trade, consumers are free to choose from which country to buy a good. Therefore, the actual price consumers pay for good $l$ is $p_j(l)$, the lowest price across all sources $i$:

$$p_j(l) = \min\left\{ p_{ij}(l);\ i = 1, \ldots, N \right\} \tag{1.A.3}$$

where $N$ again denotes the number of countries in the world.

Following Eaton and Kortum (2002), the country's efficiency is drawn from a Fréchet distribution: $F_i(z) = e^{-T_i z^{-\theta}}$, where $T_i$ is the location parameter for country $i$ and $\theta$ governs the variation within the distribution, which is assumed to be common to all countries. Replacing $z$ in $F_i(z)$ using Equation (1.A.2) leads to $G_{ij}(p) = Pr\left[ P_{ij} \le p \right] = 1 - e^{-\left[ T_i (\varsigma_i t_{ij})^{-\theta} \right] p^{\theta}}$. Given that the distribution of prices for which a country $j$ buys is given by $G_j(p) = Pr\left[ P_j \le p \right] = 1 - \prod_{i=1}^{N} \left[ 1 - G_{ij}(p) \right]$, it simplifies to:

$$G_j(p) = 1 - e^{-\Phi_j p^{\theta}} \tag{1.A.4}$$

where $\Phi_j = \sum_{i=1}^{N} T_i \left( \varsigma_i t_{ij} \right)^{-\theta}$.

The probability that country *i* provides good *l* at the lowest price to country *j* is given by:

$$\pi_{ij} = \frac{T_i\left(\varsigma_i t_{ij}\right)^{-\theta}}{\Phi_j} \tag{1.A.5}$$

Under the assumption of a continuum of goods between zero and one, this is also the fraction of goods that country *j* buys from country *i*. The price of a good that country *j* actually buys from any country *i* is also distributed $G_j(p)$, and the exact price index is given by $P_j = \gamma\Phi_j^{-1/\theta}$ with $\gamma = \left[\Gamma\left(\left(\theta+1-\sigma\right)/\theta\right)\right]^{1/(1-\sigma)}$ where $\Gamma$ is the Gamma function.

The fraction of goods that country *j* buys from country *i*, $\pi_{ij}$, is also the fraction of its expenditures on goods from country *i*, $X_{ij}$, due to the fact that the average expenditures per good do not vary by source, namely:

$$X_{ij} = \frac{T_i\left(\varsigma_i t_{ij}\right)^{-\theta}}{\Phi_j}E_j = \frac{T_i\left(\varsigma_i t_{ij}\right)^{-\theta}}{\sum_{j=1}^{N} T_k\left(\varsigma_k t_{kj}\right)^{-\theta}}E_j \tag{1.A.6}$$

where $E_j$ is country *j*'s total spending.

In addition, at delivered prices (because part of the shipments melt en route), the value of output in country *i*, $Y_i$, should be equal to the total expenditure on this country's variety in all countries in the world, including *i* itself:

$$Y_i = \sum_{j=1}^{N} X_{ij} = T_i\varsigma_i^{-\theta}\sum_{j=1}^{N}\frac{t_{ij}^{-\theta}}{\Phi_j}E_j \tag{1.A.7}$$

Solving for $T_i\varsigma_i^{-\theta}$ in equation (1.A.7) yields:

$$T_i\varsigma_i^{-\theta} = \frac{Y_i}{\sum_{j=1}^{N}\frac{t_{ij}^{-\theta}}{\Phi_j}E_j} \tag{1.A.8}$$

Substituting this expression for $T_i\varsigma_i^{-\theta}$ in equation (1.A.6) leads to:

$$X_{ij} = \frac{t_{ij}^{-\theta}}{\Phi_j\left(\sum_{j=1}^{N}\frac{t_{ij}^{-\theta}}{\Phi_j}E_j\right)}Y_iE_j \tag{1.A.9}$$

Replacing $\Phi_j$ using $P_j = \gamma\Phi_j^{-\frac{1}{\theta}}$ in both terms of the denominator of equation (1.A.9) yields:

$$X_{ij} = \frac{t_{ij}^{-\theta}}{\gamma^{\theta}P_j^{-\theta}\left(\sum_{j=1}^{N}\frac{t_{ij}^{-\theta}}{\gamma^{\theta}P_j^{-\theta}}E_j\right)}Y_iE_j \tag{1.A.10}$$

The term $\Pi_i$ can be defined:

$$\Pi_i = \left( \sum_{j=1}^{N} \left( \frac{t_{ij}}{P_j} \right)^{-\theta} \frac{E_j}{Y} \right)^{-\frac{1}{\theta}} \tag{1.A.11}$$

where $Y \equiv \Sigma_j Y_j$.

Similarly, $P_j$ can also be expressed as follows:

$$P_j = \left( \gamma^{-\theta} \Phi_j \right)^{-\frac{1}{\theta}} = \left( \gamma^{-\theta} \sum_{i=1}^{N} T_i \left( \varsigma_i t_{ij} \right)^{-\theta} \right)^{-\frac{1}{\theta}} = \left( \gamma^{-\theta} \sum_{i=1}^{N} \frac{t_{ij}^{-\theta} Y_i}{\sum_{l=1}^{N} \frac{t_{il}^{-\theta}}{\Phi_l} E_l} \right)^{-\frac{1}{\theta}} = \left( \sum_{i=1}^{N} \left( \frac{t_{ij}}{\Pi_i} \right)^{-\theta} \frac{Y_i}{Y} \right)^{-\frac{1}{\theta}} \tag{1.A.12}$$

Equation (1.A.10) can thus be rewritten as:

$$X_{ij} = \frac{Y_i E_j}{Y} \left( \frac{t_{ij}}{\Pi_i P_j} \right)^{-\theta} \tag{1.A.13}$$

By replacing -$\theta$ by 1-$\sigma$, the system of equations (1.A.11)-(1.A.13) corresponds to the same exact system of equations (1-8)-(1-10) derived from the demand side in Section B.1:

| | **Demand-side** | **Supply side** | |
|---|---|---|---|
| (1-8) | $X_{ij} = \frac{Y_i E_j}{Y} \left( \frac{t_{ij}}{\Pi_i P_j} \right)^{1-\sigma}$ | $X_{ij} = \frac{Y_i E_j}{Y} \left( \frac{t_{ij}}{\Pi_i P_j} \right)^{-\theta}$ | (1.A.13) |
| (1-9) | $\Pi_i^{1-\sigma} = \sum_j \left( \frac{t_{ij}}{P_j} \right)^{1-\sigma} \frac{E_j}{Y}$ | $\Pi_i^{-\theta} = \sum_j \left( \frac{t_{ij}}{P_j} \right)^{-\theta} \frac{E_j}{Y}$ | (1.A.11) |
| (1-10) | $P_j^{1-\sigma} = \sum_i \left( \frac{t_{ij}}{\Pi_i} \right)^{1-\sigma} \frac{Y_i}{Y}$ | $P_j^{-\theta} = \sum_i \left( \frac{t_{ij}}{\Pi_i} \right)^{-\theta} \frac{Y_i}{Y}$ | (1.A.12) |

Hence, the structural gravity system can be derived from the demand and the supply side. They are *isomorphic*. The only difference is that the elasticity of substitution is replaced by the Fréchet parameter governing the variation within the distribution.

## Appendix B: Structural gravity with tariffs

This appendix extends the standard gravity model to accommodate tariffs and tariff revenues. See for a similar derivation Heid and Larch (2016), and for an application of such a framework to quantify tariff evasion Egger and Larch (2012). All main assumptions are preserved. Specifically, each of the $N$ countries in the world produces a differentiated variety of goods (Armington, 1969). The supply of each variety is fixed at $Q_i$ with a corresponding factory-gate price $p_i$. Thus, the value of (income from) domestic production in country $i$ is defined as $Y_i = p_i\ Q_i$. Consumer preferences are approximated by a Constant Elasticity of Substitution (CES) utility function defined as:

$$\left\{\sum_i \alpha_i^{\frac{1-\sigma}{\sigma}} c_{ij}^{\frac{\sigma-1}{\sigma}}\right\}^{\frac{\sigma}{\sigma-1}} \tag{1.B.1}$$

where $\sigma > 1$ is the trade elasticity of substitution; $\alpha_i > 0$ is the CES preference parameter; and $c_{ij}$ denotes consumption of varieties from country $i$ in country $j$. Consumers maximize (1.B.1) subject to the following budget constraint:

$$E_j = \sum_i p_{ij} c_{ij} = Y_j + \sum_i (\tau_{ij} - 1) X_{ij} \tag{1.B.2}$$

Budget constraint (1.B.2) is adjusted to reflect the fact that tariff revenues, $\Sigma_i(\tau_{ij} - 1)X_{ij}$, are collected, and that these revenues are assumed to be fully rebated to consumers and add to their nominal income from production $Y_j$.[30] Tariffs here are defined as $\tau_{ij} = 1 + adv\_tariff_{ij}$ where $adv\_tariff_{ij}$ is the ad-valorem tariff on varieties imported in country $j$ from country $i$. Finally, equation (1.B.2) ensures that the total expenditure in country $j$ is equal to the total spending on varieties from all countries, including $j$, at delivered prices $p_{ij} = \tau_{ij} p_i t_{ij}$, which now are defined as a function of tariffs, $\tau_{ij}$, in addition to factory-gate prices in the origin, $p_i$, and the iceberg costs, $t_{ij} \geq 1$.

Solving the consumer's optimization problem yields the Marshallian consumer demand for goods shipped from origin $i$ to destination $j$, which reads as follows:

$$c_{ij} = p_{ij}^{-\sigma} \left(\frac{\alpha_i}{P_j}\right)^{1-\sigma} E_j \tag{1.B.3}$$

where $E_j$ corresponds to total expenditure in country $j$ and $P_j$ denotes the CES consumer price index defined as:

$$P_j = \left[\sum_i (\alpha_i p_{ij})^{1-\sigma}\right]^{\frac{1}{1-\sigma}} \tag{1.B.4}$$

Using equation (1.B.3) the value of exports from $i$ to $j$ at delivered prices is defined as:

$$X_{ij} = c_{ij} p_i t_{ij} = \tau_{ij}^{-\sigma} \left(\frac{\alpha_i p_i t_{ij}}{P_j}\right)^{1-\sigma} E_j \tag{1.B.5}$$

The final step in the derivation of the structural gravity model is to impose market clearance for goods from each origin, namely:

$$Y_i = \sum_j X_{ij} = \sum_j \tau_{ij}^{-\sigma} \left( \frac{\alpha_i p_i t_{ij}}{P_j} \right)^{1-\sigma} E_j \tag{1.B.6}$$

The first equality in equation (1.B.6) specifies that the pre-tariff value of total expenditure on goods from country $i$, $\Sigma_j X_{ij}$, is equal to the value of output in country $i$, $Y_i$. The second equality in equation (1.B.6) applies the definition of bilateral expenditure from equation (1.B.5). Defining the total world output, $Y \equiv \sum_i Y_i$, and dividing the left and the right side of equation (1.B.6) by $Y$ yields after rearrangement:

$$(\alpha_i p_i)^{1-\sigma} = \frac{Y_i / Y}{\sum_j \tau_{ij}^{-\sigma} \left( \frac{t_{ij}}{P_j} \right)^{1-\sigma} \frac{E_j}{Y}} \tag{1.B.7}$$

Defining the term in the denominator of equation (1.B.7) as $\Pi_i^{1-\sigma} \equiv \sum_j \tau_{ij}^{-\sigma} \left( t_{ij} / P_j \right)^{1-\sigma} E_j / Y$, and substituting this definition into equation (1.B.7) simplifies the expression as follows:

$$(\alpha_i p_i)^{1-\sigma} = \frac{Y_i / Y}{\Pi_i^{1-\sigma}} \tag{1.B.8}$$

Using equation (1.B.8) to substitute for the power transform $(\alpha_i p_i)^{1-\sigma}$ in the bilateral allocations equation (1.B.5) and in the CES price index equation (1.B.4), and combining the definition of $\Pi_i^{1-\sigma}$ with the resulting expressions that correspond to equation (1.B.5) and equation (1.B.6), the structural gravity system with tariffs is defined as:

$$X_{ij} = \frac{Y_i E_j}{Y} \left( \frac{t_{ij}}{\Pi_i P_j} \right)^{1-\sigma} \tau_{ij}^{-\sigma} \tag{1.B.9}$$

$$\Pi_i^{1-\sigma} = \sum_j \tau_{ij}^{-\sigma} \left( \frac{t_{ij}}{P_j} \right)^{1-\sigma} \frac{E_j}{Y} \tag{1.B.10}$$

$$P_j^{1-\sigma} = \sum_i \left( \frac{t_{ij} \tau_{ij}}{\Pi_i} \right)^{1-\sigma} \frac{Y_i}{Y} \tag{1.B.11}$$

The system of equations (1.B.9)-(1.B.11) resembles the system of equations (1-8)-(1-10) defined in Section B.1. However, there are two differences, which have implications for the estimation of gravity and for the welfare analysis with the gravity model. These differences are:

(i) Tariffs enter the gravity equation (1.B.9) directly and indirectly, via the multilateral resistances. The implication with respect to gravity estimations is that tariffs will appear in the estimating equation and, more importantly, the estimate of the coefficient on tariffs can be used to directly recover a value for the trade elasticity parameter. In addition, the structural gravity theory presented here can be used to calculate tariff equivalent effects for each of the gravity covariates. The interested reader may refer to Larch and Wanner (2014) for further discussion and analysis of the empirical implications of the inclusion of tariffs.

(ii) The expression for expenditure differs from the value of total production owing to tariff revenues. This difference has no implications for gravity estimations since expenditure at the country level, regardless of their functional form, will be absorbed by the *importer-time* fixed effects. However, this difference has important implications for welfare analysis. Specifically, as demonstrated by Anderson and van Wincoop (2001), the expression for real income with rents from tariff becomes:

$$\tilde{W}_j = \frac{Q_j p_j}{P_j} \frac{1}{1 - \sum_i \left(\tau_{ij} - 1\right) s_{ij}} \tag{1.B.12}$$

where $s_{ij} = \tau_{ij} X_{ij} / E_j = \left(\alpha_i p_{ij} / P_j\right)^{1-\sigma}$ is the CES expenditure share on goods from country *i* in country *j* (including tariffs). The first fraction in equation (1.B.12) is the expression for real income from the gravity model without tariffs. The second fraction is a tariff multiplier, which captures the additional welfare effects of the introduction of tariffs and rents. The expression for the tariff multiplier as an adjustment to nominal income can be obtained by using the definition of $s_{ij}$ in budget constraint (1.B.2) and then solving for total expenditure. The interested reader may refer to Anderson and van Wincoop (2001) for further comparative statics and discussion of the welfare implications of the introduction of tariffs and rents.

## Appendix C: Databases and data sources links summary

| *Data* | *Database* | *Link* |
|---|---|---|
| **Country-specific data** | IFS | http://data.imf.org/?sk=5DABAFF2-C5AD-4D27-A175-1253419C02D1 |
| | WDIs | http://databank.worldbank.org/data/reports.aspx?source=world-development-indicators |
| | PWT | http://www.rug.nl/research/ggdc/data/pwt |
| **Aggregated merchandise trade flows** | DOTS | http://data.imf.org/?sk=9D6028D4-F14A-464C-A2F2-59B2CD424B85&ss=13900303418545 |
| | Correlates of War Project | http://correlatesofwar.org/data-sets/bilateral-trade. |
| **Disaggregated merchandise trade** | Comtrade | http://comtrade.un.org |
| | BACI | http://www.cepii.fr/cepii/en/bdd_modele/presentation.asp?id=1 |
| | WTF | http://cid.econ.ucdavis.edu/Html/WTF_bilateral.html |
| | Trade Map | http://www.trademap.org/Index.aspx |
| | Global Trade Atlas | https://www.gtis.com/gta |
| **Service trade flows** | OECD's Trade in Services | https://stats.oecd.org/Index.aspx?DataSetCode=TISP |
| | UN Service Trade | http://comtrade.un.org/ |
| | Trade Map | http://www.trademap.org/Index.aspx |
| | Francois and Pindyuk's Trade in Service | https://ideas.repec.org/p/lnz/wpaper/20130101.html |
| **Agriculture trade flows** | FAOSTAT Detailed Trade Matrix | http://faostat3.fao.org/download/T/TM/E |
| **Intra-national trade flows data** | TPP | http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESEARCH/0,,contentMDK:21085384~pagePK:64214825~piPK:64214943~theSitePK:469382,00.html |
| | TradeProd | http://www.cepii.fr/CEPII/fr/bdd_modele/presentation.asp?id=5 |
| | INDSTAT | http://stat.unido.org/ |
| **Time-invariant covariates** | GeoDist | http://www.cepii.fr/cepii/en/bdd_modele/presentation.asp?id=6 |
| **Tariff data** | IDB | https://tao.wto.org/welcome.aspx?ReturnUrl=%2f%3fui%3d1&ui=1 or http://tariffdata.wto.org/. |
| | CTS | https://tao.wto.org/welcome.aspx?ReturnUrl=%2f%3fui%3d1&ui=1 or http://tariffdata.wto.org/. |
| | TRAINS | http://databank.worldbank.org/data/reports.aspx?source=UNCTAD-~-Trade-Analysis-Information-System-(TRAINS) |
| | MacMap | http://www.macmap.org/Default.aspx?ReturnUrl=%2fAdvancedSearch%2fTariffAndTrade%2fDefault.aspx |
| | I-TIP | https://www.wto.org/english/res_e/statis_e/itip_e.htm |
| | OECD's AMAD | https://www.oecd.org/site/amad. |
| **Various NTMs** | TPP | http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESEARCH/0,,contentMDK:21085384~pagePK:64214825~piPK:64214943~theSitePK:469382,00.html |
| | TRAINS | http://databank.worldbank.org/data/reports.aspx?source=UNCTAD-~-Trade-Analysis-Information-System-(TRAINS) |
| | NTM-MAP | http://www.cepii.fr/CEPII/en/bdd_modele/presentation.asp?id=28 |
| | I-TIP | https://www.wto.org/english/res_e/statis_e/itip_e.htm |
| | Trade Monitoring | http://tmdb.wto.org/searchmeasures.aspx?lang=en-US |
| | Global Trade Alert | www.globaltradealert.org/ |

*(Continued)*

*(Continued)*

| *Data* | *Database* | *Link* |
|---|---|---|
| **Export subsidies** | Agricultural Information Management System | http://agims.wto.org/ |
| **Fossil fuel subsidies** | IEA's data | http://www.worldenergyoutlook.org/resources/energysubsidies/fossilfuelsubsidydatabase/ |
| **Aggregated subsidies** | World Bank's data | http://databank.worldbank.org/data/reports.aspx?source=world-development-indicators |
| **Export restrictions** | OECD's Inventory on export restrictions | http://qdd.oecd.org/subject.aspx?Subject=ExportRestrictions_PrimaryAgriculture<br>http://qdd.oecd.org/subject.aspx?Subject=ExportRestrictions_IndustrialRawMaterials |
| **Trade remedies** | TTBD | http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESEARCH/EXTPROGRAMS/EXTTRADERESEARCH/0,,contentMDK:22561572~pagePK:64168182~piPK:64168060~theSitePK:544849,00.html |
| **Technical barriers to trade (TBT)** | TBT-IMS | http://tbtims.wto.org/ |
| **Sanitary and phytosanitary measures (SPS)** | SPS-IMS | http://spsims.wto.org/ |
| **TBT and SPS** | Perinorm | http://www.perinorm.com/home/default.aspx?ReturnUrl=%2fdefault.aspx |
| **Service trade restrictiveness index (STRI)** | OECD's STRI | http://qdd.oecd.org/subject.aspx?Subject=063bee63-475f-427c-8b50-c19bffa7392d |
| | World Bank's STRI | http://iresearch.worldbank.org/servicetrade/ |
| | Service I-TIP | http://i-tip.wto.org/services/default.aspx |
| **Trade facilitation** | Doing Business | http://databank.worldbank.org/data/reports.aspx?source=doing-business |
| | OECD's TFI | http://www.oecd.org/trade/facilitation/indicators.htm |
| | LPI | http://lpi.worldbank.org/ |
| | ETI | https://knoema.com/atlas/sources/WEF?topic=Foreign-Trade |
| **Preferential and regional trade agreements** | RTA-IS | http://rtais.wto.org/UI/PublicMaintainRTAHome.aspx |
| | Larch's RTA | http://www.ewf.uni-bayreuth.de/en/research/RTA-data/index.html |
| | Bergstrand's EIA | http://kellogg.nd.edu/faculty/fellows/bergstrand.shtml |
| | WTO's PTA | http://ptadb.wto.org/?lang=1 |
| | GTPA | http://wits.worldbank.org/gptad/database_landing.aspx |
| | DESTA | http://www.designoftradeagreements.org/www.designoftradeagreements.org/indexf908.html?page_id=884 |

# Endnotes

1 This chapter is based on the paper "Estimating Trade Policy Effects with Structural Gravity" prepared by Piermartini and Yotov (2016).

2 The CES utility function assumption is widely used in the existing gravity literature. Anderson and Neary (2005) discuss the implications of more general, non-homothetic preferences. Recent attempts to depart from CES utility function, while preserving the key properties of the structural gravity model, include Novy (2013), Behrens et al. (2014), and Arkolakis et al. (2015), who also provide an informative review of the main alternatives to CES utility functions that have been used in the trade literature.

3 See for a derivation of the CES demand equation Appendix 2.A in Baldwin et al. (2011).

4 World output, $Y$, does not appear explicitly in the general discussions of the structural gravity model as presented in some recent surveys and academic articles that adjust the definitions of equations (1-8), (1-9) and (1-10) to account for $Y$.

5 Anderson (2011) offers an insightful discussion and formal proofs of these and other, less obvious, properties based on the relationship between trade flows and country size in a frictionless world.

6 In principle, the interpretation problem can be fixed by using the inverse hyperbolic sine function (Kristjánsdóttir, 2012). However, this procedure has to be applied with caution because it is a non-linear transformation (as is the log-transformation), which means that with heteroskedastic trade data one may end up with inconsistent estimates (Santos Silva and Tenreyro, 2006).

7 For a discussion of the relative merits of the PPML estimator vs. other linear and non-linear estimators, the interested reader may refer to Santos Silva and Tenreyro (2006), Santos Silva and Tenreyro (2011), Egger and Staub (2016), and Head and Mayer (2014).

8 In principle, the heteroscedasticity issue can be addressed with other estimators, such as the Gamma estimator.

9 The reader may refer to Disdier and Head (2008) and Head and Mayer (2013) for an analysis of the use and impact of distance in gravity regressions.

10 See for example Magee (2003) and Baier and Bergstrand (2002, 2004). Egger et al. (2011) offer a recent IV treatment of the endogeneity of trade agreements in a cross-section setting.

11 Monte Carlo simulations suggest that although the PPML estimator underestimates the distance effect, the estimates of the parameter associated with the distance variable converges to the true value as the sample size increases (Head and Mayer, 2014). The interested reader may refer to the "Log of Gravity" webpage at http://personal.lse.ac.uk/tenreyro/LGW.html for a series of discussions on the benefits and potential downside of using the PPML estimator.

12 If the trade policy is a barrier, for example a dummy for a quota, the formula for calculating the ad-valorem tariff equivalent whose removal would have generated the same impact as the removal of the barrier in question would be $(e^{\beta_{QUOTA} / \beta_{TARIFF}} - 1) \times 100$.

13 The interested reader may refer to Felbermayr et al. (2015) for a discussion of the sensitivity of the results in counterfactual gravity analysis to the choice of the elasticity of substitution.

14 The OECD and WTO are currently working on building a global matrix of trade in services statistics. The dataset will include exports and imports of total services and of 11 main EBOPS 2002 items, and will cover 191 reporters and partners for the period 1995-2012. Data are obtained from OECD, Eurostat, WTO-UNCTAD-ITC and national sources. Missing observations are estimated using different techniques, such as backcasting, forecasting, interpolations, derivations, integration of EBOPS 2010 data, as well as gravity-model based estimates in order to obtain a complete, square matrix. An EBOPS 2010 version of the dataset is also envisaged for the near future.

15 Under the Project to Improve Provincial Economic Statistics (PIPES), Canada's government has created a database that includes consistent intra-national and international data for Canada's economy at the sectoral level for the period 1997-2007. See Genereux and Langen (2002) for further details.

16 The CEPII's CHELEM-International Trade database covers 94 countries from 1967 onwards with sectoral data classified according to the CHELEM nomenclature (71 sectors), GTAP (43 sectors) and ISIC classification (147 sectors). The CHELEM nomenclature has been built to allow a better correspondence between data on trade and production. However, unlike the first three databases, the CHELEM-International Trade database includes estimated observations, and therefore should not be used for gravity estimations.

17 Bound rates are in general neglected in standard specifications of gravity models, as bound rates only reflect countries commitments but are not the tariffs that importers and exporters face when trading. Recent economic literature has, however, highlighted the importance that bound rates can have in determining a firm's decision to trade insofar as they affect the certainty of trading conditions (Handley and Limão, 2013; Handley, 2014; Osnago et al., 2015). A direct measure of trade policy uncertainty is the so-called "tariff water"-the gap between the bound and applied tariff rate.

18 As the data used for the applications is a panel data set with repeated observations of pairs of countries over time, common observable and unobservable effects may naturally arise. While important bilateral time-varying effects are controlled for with explanatory variables, such as RTAs, and any bilateral time-invariant effects are taken into account with fixed effects, some correlation pattern between pairs of countries over time may still be present in the error term. This correlation pattern is captured by clustering the errors over country-pairs. Bertrand et al. (2004) provide some evidence that this general procedure does quite well.

19 The dimensions of the data were predetermined by the availability of consistently constructed *international* and *intra-national* trade flows data. The sample covers the following countries: Argentina, Australia, Austria, Belgium-Luxembourg, Bolivia, Brazil, Bulgaria, Cameroon, Canada, Chile, China, Colombia, Costa Rica, Cyprus, Denmark, Ecuador, Egypt, Finland, France, Germany, Greece, Hong Kong (China), Hungary, Iceland, India, Indonesia, Iran, Ireland, Israel, Italy, Japan, Jordan, Kenya, Kuwait, Macao (China), Malaysia, Malta, Mauritius, Malawi, Mexico, Morocco, Myanmar, the Netherlands, Nepal, Niger, Nigeria, Norway, Panama, the Philippines, Poland, Portugal, Qatar, Romania, Senegal, Singapore, South Africa, the Republic of Korea, Spain, Sri Lanka, Sweden, Switzerland, Tanzania, Thailand, Trinidad and Tobago, Tunisia, Turkey, Uruguay, the United Kingdom, and the United States.

20 The interested reader may refer to Baier et al. (2016) for further details on the construction of the trade data.

21 The interested reader may refer to Melitz (2008) and Melitz and Toubal (2014) for recent and more involved analysis of the effects of language on international trade.

22 For expositional simplicity, the notation used for the coefficients and error term is the same when moving from specification (1-23) to specification (1-24), although technically speaking it should be different.

23 The Ramsey RESET test detects whether potential variables are omitted in the model specification. The null hypothesis ($H_0$) states that the model does not suffer from misspecification errors suggesting the model is correctly specified. The null hypothesis can be rejected when the *p*-value is smaller than the critical value. Conversely, the null hypothesis cannot be rejected if the *p*-value is larger than the significance value.

24 The distance puzzle has been of significant interest to the professions. See for example Buch et al. (2004), Carrère and Schiff (2005), Brun et al. (2005), Boulhol and de Serres (2010), Lin and Sim (2012), and Larch et al. (2016).

25 Some recent studies obtain region-specific estimates of the variable SMCTRY and document wide variation of those estimates across countries (Anderson et al., 2016b) and even across provinces within Canada (Agnosteva et al., 2014). Related studies investigate the implications of intra-national trade costs for international trade and welfare (Ramondo et al., 2014).

26 While this application focuses on the effects of regional trade agreements, a different literature investigates the questions about which countries (Baier and Bergstrand, 2004; Egger and Larch, 2008) and when (Bergstrand et al., 2016) regional trade agreements are concluded.

27 When individual, country-specific fixed effects are used instead of a single *SMCTRY* indicator variable, all country-specific dummies for intra-national trade have to be dropped. This will have no effect on the estimates from column (4) of Table 3.

28 Felbermayr et al. (2015, pp. 7) explain the downward bias as follows: "If the error term in the gravity model represents unobservable policy-related barriers that reduce trade, and if those barriers make an RTA more likely, then the RTA dummy and the error term will be negatively correlated, leading to underestimation of the RTA coefficient."

29 The reader may refer to Heid and Larch (2016) for a similar derivation.

30 For the ease of notation, trade imbalances are not considered in this appendix. See for a derivation of a model with tariffs and trade imbalances for example Online Appendix A of Heid and Larch (2016).