CHAPTER 3: Analyzing bilateral trade using the gravity equation

TABLE OF CONTENTS

CHAPTER 3

LIST OF TABLES

LIST OF BOXES

A. Overview and learning objectives

This chapter will introduce the gravity model, a work-horse of international trade analysis. After a brief overview of the theoretical foundation of gravity models, we will guide you through possible alternative estimation methods of the "mother" gravity equation. We will then turn to the discussion of advanced issues on gravity modelling such as how to handle zero-trade flows and how to calculate tariff equivalents of non-tariff barriers. Data sources for gravity analysis and instructions on how to build a gravity database will be provided at the end of this section.

Equipped with these analytical tools, we will review the debate over the impact of the WTO membership on trade. This application will allow us to show the importance of a correct specification of the estimation equation for unbiased results, to explain to the reader how to interpret the results of a regression analysis (including inferences on trade diversion and trade creation) and to discuss a number of potential estimation problems such as endogeneity and heteroskedasticity.

Exercises are provided at the end of the chapter. Data and do files for the solution of these exercises can be downloaded from the web/DVD.

In this chapter, you will learn:

- what are the logical foundations of the gravity equation
- where to find the data needed to estimate it
- what are the main measurement issues you should be aware of
- what are the main econometric estimation issues you should be aware of
- how to present and interpret the results
- how to build the database and run a regression to estimate a standard gravity model
- how to calculate the tariff equivalent (*ad valorem* or specific) of a quantitative restriction (QR) using gravity models.

After reading this chapter, with some knowledge of econometrics and familiarity with STATA, you will be able to run a properly formulated gravity equation using STATA and to interpret it with the key caveats in mind.

B. Analytical tools

1. The gravity equation: theoretical equation

It has been known since the seminal work of Jan Tinbergen (1962) that the size of bilateral trade flows between any two countries can be approximated by a law called the "gravity equation" by analogy with the Newtonian theory of gravitation. Just as planets are mutually attracted in proportion to their sizes and proximity, countries trade in proportion to their respective GDPs and proximity.[1] Initially the gravity equation was thought of merely as a representation of an empirically stable relationship between the size of economies, their distance and the amount of their trade. Prominent models of international trade at that time included the Ricardian model, which relies on differences in technology across countries to explain trade patterns, and the Heckscher-Ohlin

(HO) model that relies on differences in factor endowments among countries as the basis for trade. It was assumed then that standard Ricardian and HO models were incapable of providing a foundation for the gravity model. In the HO model, for example, country size has little to do with the structure of trade flows.

The extraordinary stability of the gravity equation and its power to explain bilateral trade flows prompted the search for a theoretical explanation for it. Whereas empirical analysis predated theory, we know now that most trade models require gravity in order to work. The first important attempt to provide a theoretical basis for gravity models was the work of Anderson (1979). He did so in the context of a model where goods were differentiated by country of origin (the so-called Armington assumption) and where consumers have preferences defined over all the differentiated products. This structure would imply that, whatever the price, a country will consume at least some of every good from every country. All goods are traded, all countries trade and, in equilibrium, national income is the sum of home and foreign demand for the unique good that each country produces. For this reason, larger countries import and export more. Trade costs are modelled as "iceberg" costs, that is, only a fraction of the good shipped arrives to destination, the rest having melted in transit. Clearly, if imports are measured at the CIF value, transport costs reduce trade flows.

Subsequent elaborations have shown that, far from being a purely econometric tool without a theoretical basis (an early criticism against the gravity model), gravity models can arise out of a range of trade theories.[2] In particular, Bergstrand (1985 and 1989) shows that a gravity model is a direct implication of a model of trade based on monopolistic competition developed by Paul Krugman (1980). In this model, identical countries trade differentiated goods because consumers have a preference for variety. Models with monopolistic competition overcome the undesirable feature of Armington models whereby goods are differentiated by location of production by assumption. Firm location is endogenously determined and countries are specialized in the production of different sets of goods. Deardorff (1998) shows that a gravity model can arise from a traditional factor-proportions explanation of trade. Eaton and Kortum (2002) derive a gravity-type equation from a Ricardian type of model, and Helpman et al. (2008) and Chaney (2008) obtained it from a theoretical model of international trade in differentiated goods with firm heterogeneity.[3]

In its general formulation, the gravity equation has the following multiplicative form:

$$X_{ij} = G S_i M_j \phi_{ij} \tag{3.1}$$

where X_{ij} is the monetary value of exports from i to j, M_j denotes all importer-specific factors that make up the total importer's demand (such as the importing country's GDP) and S_i comprises exporter-specific factors (such as the exporter's GDP) that represent the total amount exporters are willing to supply. G is a variable that does not depend on i or j such as the level of world liberalization. Finally, ϕ_{ij} represents the ease of exporter i to access of market j (that is, the inverse of bilateral trade costs).

The contribution of recent research concerning the theoretical foundation of the gravity equation is to have highlighted the importance of deriving the specifications and variables used in the gravity model from economic theory in order to draw the proper inferences from estimations using the

gravity equation. Particularly important has been in this respect the contribution of Anderson and van Wincoop's (2003) paper, where they show that controlling for relative trade costs is crucial for a well-specified gravity model. Their theoretical results show that bilateral trade is determined by *relative* trade costs, i.e. the propensity of country *j* to import from country *i* is determined by country *j*'s trade cost toward *i* relative to its overall "resistance" to imports (weighted average trade costs) and to the average "resistance" facing exporters in country *I*; not simply by the absolute trade costs between countries *i* and *j* (Anderson and van Wincoop, 2003). The rationale for including these "multilateral trade-resistance" (MTR) terms, as they are called, is that, ceteris paribus, two countries surrounded by other large trading economies, say Belgium and the Netherlands bordered by France and Germany respectively as well as by each other, will trade less between themselves than if they were surrounded by oceans (such as Australia and New Zealand) or by vast stretches of deserts and mountains (such as the Kyrgyz Republic and Kazakhstan).

In particular, Anderson and van Wincoop show that in a context of world of N countries and a variety of goods differentiated by the country of origin a well-specified theoretically funded gravity equation takes the form:

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

(3.2)

where Y denotes world GDP, Y_i and Y_j the GDP of countries i and j respectively, t_{ij} (one plus the tariff equivalent of overall trade costs) is the cost in j of importing a good from i, $\sigma > 1$ is the elasticity of substitution and Π_i and P_j represent exporter and importer ease of market access or country i's outward and country j's inward *multilateral resistance* terms. They are low if a country is remote from world markets, remoteness being determined by physical factors such as physical distance from large markets as well as policy factors such as high tariff barriers or other trade costs. This result highlighted the severe mistake made in estimating gravity models by those studies that proxy S_i and M_j in equation (3.1) with exporting and importing country GDPs without controlling for multilateral resistance terms.

For simplicity we have omitted time indexes from equation (3.1). However, all variables in equation (3.1) can vary over time. Furthermore, we only considered aggregate data, but gravity models can also be run using sectoral data as well (see Box 3.2).

2. Estimation methods

Given the multiplicative nature of the gravity equation, the standard procedure for estimating a gravity equation (3.1) is simply to take the natural logarithms of all variables and obtain a log-linear equation that can be estimated by ordinary least squares regression (clearly easier than non-linear estimation methods). This yields the estimation equation:

$$\ln X_{ij} = \ln G + \ln S_i + \ln M_j + \ln \phi_{ij}$$

(3.3)

or more specifically in the case of the Anderson and van Wincoop model:

$$\ln X_{ij} = a_0 + a_1 \ln Y_i + a_2 \ln Y_j + a_3 \ln t_{ij} + a_4 \ln \Pi_i + a_5 \ln P_j + \varepsilon_{ij} \tag{3.4}$$

where a_0 is a constant, $a_3 = 1 - \sigma$ and ε is the error term.

In practice, the gravity equation relates the natural logarithm of the monetary value of trade between two countries to the log of their respective GDPs, a composite term measuring barriers and incentives to trade between them, and terms measuring barriers to trade between each of them and the rest of the world. This specification allows in addition an easy interpretation of the estimated parameters: the parameters of an equation estimated in logarithms are elasticities. For example, the estimated parameter for the GDP in a gravity equation estimated in logarithms is the elasticity of trade to GDP, indicating the percentage variation in trade following a 1 per cent increase in GDP.

A number of variables are generally used to capture trade costs ϕ_{ij}. Typically, empirical studies proxy trade costs with bilateral distance. However, a number of additional variables are also customarily used. These include dummies[4] for islands, landlocked countries and common borders. They are used to reflect the hypotheses that transport costs increase with distance and that they are higher for landlocked countries and islands but are lower for neighbouring countries. Dummies for common language, adjacency or other relevant cultural features such as colonial history are used to capture information costs. Search costs are probably lower for trade between countries whose business practices, competitiveness and delivery reliability are well known to one another. Firms in adjacent countries, countries with a common language or other relevant cultural features are likely to know more about each other and to understand each other's business practices better than firms operating in less-similar environments. For this reason, firms are more likely to search for suppliers or customers in countries where the business environment is familiar to them. Tariff barriers are generally included in the form of dummies for the existence of regional trade agreements. Very few studies use information on bilateral tariffs, one reason being the lack of data over time.

The problem with estimating equation (3.4) is that the so-called multilateral resistance terms (MRTs) are not directly observable. Several alternative ways of proxying for MTRs are possible. One is to use iterative methods to construct estimates of the price-raising effects of barriers to multilateral trade (Anderson and van Wincoop, 2003). However, this procedure is not frequently used since it requires a non-linear least square (NLS) program to obtain an estimate. A simpler alternative, often used, is to use a proxy for these indexes called a "remoteness" variable. An even simpler — and widely used — method consists of using country fixed effects for importers and exporters (Rose and van Wincoop, 2001; Feenstra, 2004; Baldwin and Taglioni, 2006). The next subsection will focus on these two simple approaches.[5]

a. Controlling for the multilateral trade resistance (MTR)

An important factor in the choice of the estimation method for the multilateral resistance terms is the specific research interest.

i.   Case a: Interest of the research focuses on the coefficient of a bilateral variable

Unbiased estimates of the impact of distance and other bilateral variables on bilateral trade flows can be obtained by replacing the multilateral resistance indexes in equation (3.4) with importer and exporter dummies (Anderson and van Wincoop, 2004), or country effects.[6] These country dummies are binary (0,1) variables that will capture all country-specific characteristics and will control for a country's overall level of imports/exports. In a gravity equation, one such variable will be set to one whenever the *exporting* country is, say, Kazakhstan and zero otherwise. Another one will be set to one whenever the *importing* country is Kazakhstan and zero otherwise, and so on for each country.

For a cross section, that is, when information on variables of interest is available only for one specific year, the empirical gravity equation estimated using fixed effects in its baseline form is:

$$\ln X_{ij} = a_0 + a_1 I_i + a_2 I_j + a_3 \ln t_{ij} + \varepsilon_{ij} \tag{3.5}$$

where $\ln t_{ij}$ is the logarithm of trade costs between two countries i and j, I_i is a dummy variable equal to one when the country is i and zero otherwise. In a cross section with n countries, if one-way trade flows are not combined, there are $2n^2$ country pairs (the unit of observation) but only $2n$ such fixed effects, so estimation is still possible.

In the gravity literature it is in general assumed that trade costs take the form:

$$t_{ij} = d_{ij}^{\delta_1} \cdot \exp\left(\delta_2 \text{cont}_{ij} + \delta_3 \text{lang}_{ij} + \delta_4 \text{ccol}_{ij} + \delta_5 \text{col}_{ij} + \delta_6 \text{landlock}_{ij} + \delta_7 \text{RTA}_{ij}\right) \tag{3.6}$$

where d_{ij} is bilateral distance, and cont_{ij}, lang_{ij}, ccol_{ij}, col_{ij}, landlock_{ij} and RTA_{ij} are dummy variables denoting respectively whether the two countries have a common border, common language and common colonizer, whether one was a colony of the other at some point in time, whether one of the two is a landlocked country (including when both countries are landlocked), or whether the two countries are members of a regional trade agreement (Box 3.1 discusses more extensively the issue of estimating the impact of trade agreements on trade). All these variables have been found to be significant determinants of bilateral trade.

In STATA
* generate importer and exporter dummies
tab (importer), gen(importer_)
tab (exporter), gen(exporter_)
reg lnexports lndist cont lang ccol col landlock RTA importer_* exporter_*, robust

alternatively

xi: reg lnexports lndist cont lang ccol col landlock RTA i.importer i.exporter, robust

Note that a gravity equation deals with observations that may be heterogeneous in a variety of ways. The assumption of homoskedasticity of the error term — under which all disturbances affecting individual observations are drawn from a common distribution — being likely to be violated, robust standard errors should be used systematically.[7]

When information is also available over time, the gravity equation in its baseline form is:

$$\ln X_{ijt} = a_0 + a_1 I_{it} + a_2 I_{jt} + a_3 \ln t_{ijt} + a_4 I_t + u_{ijt} \tag{3.7}$$

where I_t designates a dummy variable for a specific year one per year. Suppose, for instance, that our sample period spans 2001–2006. We would define a variable $I_1 = 1$ if year $= 2001$ and 0 otherwise and similarly for 2002, … , 2006, giving us six such dummy variables that could be nonzero only one at a time.[8] I_{it} and I_{jt} are importer and exporter time-varying individual effects. For a sample of T periods there are T of these variables. Note that I_{it} are importer time-varying effects; they therefore allow us to take into account the fact that MRT may change over time. There is a set of *2nT* of these variables.

The use of panel data (over time bilateral trade data) also has the advantage of mitigating the bias generated by heterogeneity across countries. While in a single cross section, country-pair propensity to trade can only be controlled for by observed country-pair characteristics (such as common language, common currency), in a panel country-pair heterogeneity can be controlled for by using country-pair fixed effects. Note, however, that if the interest of the research focuses on estimating the coefficient of a bilateral time-invariant coefficient, the fixed effect estimation is not a viable option because of perfect collinearity. The researcher may want to control with random effects in this case. The Hausman test can be used to test whether the random effect model is a suitable option.

```
In STATA
tab (year), gen (year_)
gen impyear = group(importer year)
gen expyear = group(exporter year)
tab (impyear), gen (impyear__)
tab (expyear), gen (expyear__)
xtreg lnexports lndist cont lang ccol col landlock RTA impyear_* expyear_* year_*, robust
```

or with random effects

```
xtreg lnexports lndist cont lang ccol col landlock RTA impyear_* expyear_* year_*, re robust
```

If the bilateral variable of interest is time-varying as in the case of a dummy denoting whether countries belong to the same regional trade agreement (RTA; also see Box 3.1), you can control for fixed effects (country-pair effects)

```
xtreg lnexports RTA impyear_* expyear_* year_*, fe robust
```

or with random effects

```
xtreg lnexports lndist cont lang ccol col landlock RTA impyear_* expyear_* year_*, re robust
```

Note: for relatively short time periods you may want to use non-time-varying exporter and importer country effects and control for country-specific factors such as importer and exporter GDP. See section below.

Box 3.1   Estimating trade creation and trade diversion with gravity models

The gravity equation provides a way of looking for evidence of trade diversion through the ex-post analysis of trade flows. Suppose that countries i and j belong to a common RTA, whereas country k does not. If, after the RTA's formation, i imports more from j and less from k, trade diversion is likely. If, in contrast, country i imports more from j and k, trade creation is likely. We will see here how to make this conjecture empirically testable.

Suppose we are interested in finding out whether MERCOSUR is trade diverting or trade creating. Then, letting M denote MERCOSUR, we will construct two dummy variables:

$BothinM = 1$ if i and j are both members of MERCOSUR at time t and 0 otherwise

$OneinM = 1$ if the importer (i) belongs to MERCOSUR but the exporter (j) does not.

Then we estimate the augmented gravity equation:

$$\ln X_{ijt} = \beta_0 + \beta_1 I_{it} + \beta_2 I_{jt} + \beta_3 \ln\left(\text{dist}_{ij}\right) + \beta_4 \text{cont}_{ij} + \beta_5 \text{lang}_{ij} + \beta_6 \text{ccol}_{ij} + \beta_7 \text{col}_{ij} + \beta_8 \text{landlock}_{ij} +$$
$$+ \beta_9 \text{OneinM}_{ijt} + \beta_{10} \text{BothinM}_{ijt} + \varepsilon_{ijt} \tag{3.8}$$

A positive (and significant) coefficient on both β_9 and β_{10} is suggestive of trade creation; a positive on the first but negative on the second is suggestive of trade diversion.

Two important limitations related to using gravity models for estimating the impact of a RTA should be mentioned. First, RTAs may be endogenous variables. That is, the causal link between the formation of a RTA and trade flows may proceed from the latter to the former; thus RTAs are determined by rather than determine trade flows. This affects traditional gravity-based estimates and the extent of the bias may be quite large (see subsection 3.d). Second, recent literature is replete with models in which regional integration agreements are formed in the pursuit of other, non-trade goals (see, for instance, Limao, 2006) or in which they have "non-traditional" gains (see Ethier, 1998). Indeed, South–South agreements have been rather more successful in non-trade dimensions like the management of common resources than in the dimension of pure trade-liberalization. Thus a complete analysis of RTAs should avoid limiting itself to measuring trade diversion and creation, although these are important issues for the welfare of member countries.

ii.   *Case b: When the interest of the research relies on the country-specific variable*

The country effects approach discussed above provides unbiased estimates of the coefficients of the gravity model but it has a notable drawback: it precludes direct estimation of the partial effects of country-specific explanatory variables. For instance, numerous gravity studies attempt to estimate the impact on trade of quality of infrastructure, quality of institutions or of the regulation system. These variables would be perfectly collinear with country-specific dummies.[9] In this section

we will discuss two options for addressing this problem: using time invariant exporter and importer dummies in a short sample period and calculating the remoteness variable.

Exporter and importer dummies in a short sample period

Let's assume, for example, that we are interested in testing the relevance of a country GDP in determining trade flows.

$$\ln X_{ijt} = a_0 + a_1 \ln\left(\text{GDP}_{it}\right) + a_2 \ln\left(\text{GDP}_{jt}\right) + a_3 \ln\left(t_{ij}\right) + a_4 I_i + a_5 I_j + a_6 I_t + u_{ijt} \tag{3.9}$$

This is not perfect, however, if MRTs vary over time, which is likely – say because the geographical composition of a country's trade varies. Nevertheless, they may not vary tremendously over a reasonably short sample period (see the discussion in Baldwin and Taglioni, 2006). Note that one can add a whole lot of control variables and other variables of interest to this basic equation such as the quality of institutions and the quality of infrastructure.

```
STATA
gen lnGDPexp= ln (GDPexp)
gen lnGDPimp= ln (GDPimp)
xtreg lnexports lnGDPexp lnGDPimp lndist importer_* exporter_* year_*, robust
```

Measuring remoteness

A method frequently used to control for the multilateral resistance terms for exporting and importing countries is to include a proxy for these indexes called "remoteness". This is often calculated as:

$$\text{Re} m_i = \sum_j \frac{dist_{ij}}{\text{GDP}_j / \text{GDP}_W} \tag{3.10}$$

a formula that measures a country's average weighted distance from its trading partners (Head, 2003), where weights are the partner countries' shares of world GDP (denoted by GDP_W).

There are two criticisms usually made of using this procedure: one is that it is not theoretically correct, since the only type of trade barrier that it captures is distance (Anderson and van Wincoop, 2003). The other one relates to the appropriate measure of internal distance, as the summation requires us to specify also a country's distance from itself (Head and Mayer, 2000 suggest using the square root of the country's area multiplied by about 0.4).

Recently, Baier and Bergstrand (2009) suggest estimating a linear approximation (by means of a first order Taylor series expansion) of the multilateral resistance terms, thus avoiding the non-linear procedure used in Anderson and van Wincoop (2003). Following this approach, the OLS reduced-form gravity equation is:

$$\ln X_{ij} = \beta_0 + \ln \text{GDP}_i + \ln \text{GDP}_j - (\sigma - 1)\ln t_{ij} + (\sigma - 1)\left[\sum_j \theta_j \ln t_{ij} - \frac{1}{2}\sum_i \sum_j \theta_i \theta_j \ln t_{ij}\right] +$$
$$+ (\sigma - 1)\left[\sum_i \theta_i \ln t_{ij} - \frac{1}{2}\sum_i \sum_j \theta_i \theta_j \ln t_{ij}\right] \tag{3.11}$$

where time indexes have been omitted for simplicity and θ denotes GDP shares and *t* trade costs. The terms in square brackets are the linear approximation of the MRTs. Intuitively, the first term in the bracket is a form of remoteness term (rather than only geographical distance the term reflects overall trade costs); the second term is a measure of world trade costs. Most importantly, this linearization shows that bilateral trade between i and j depends on the level of bilateral relative to multilateral trade costs and multilateral relative to world trade costs. Note that Bergstrand and Baier estimate equation (3.11) proxying trade costs with distance and border and θ with 1/N (where N is the number of countries).

How to calculate remoteness in STATA?

- Compute the share of GDP in world GDP

  ```
  bys exporter year: egen gdptotal = sum(gdp)
  gen gdpshare = gdp / gdptotal
  ```

- Compute the spatially weighted GDP share

  ```
  bys exporter year: egen remoteness = total(dist*gdpshare)
  ```

- Compute the spatially weighted GDP share according to Head (2003)

  ```
  bys exporter year: egen Remoteness_head=total(dist/gdpshare)
  ```

Note: the same procedure applies to the importer.

---

b. Lessons learnt

The recent effort of economic theory to derive a theoretically founded gravity equation has identified three typical mistakes in the traditional approach. Baldwin and Taglioni (2006) call these errors the gold, silver and bronze mistakes, respectively.

Gold medal mistake: Traditionally, the gravity equation uses log GDPs (and possibly other variables) as proxies for the $\ln S_i$ and $\ln M_j$ and omits what Anderson and van Wincoop call the multilateral resistance terms or Head (2003) and Baier and Bergstrand (2007) call "remoteness". These omitted terms are correlated with trade costs. Hence, estimations are biased.

Silver medal mistake: Averaging the reciprocal trade flows. The theoretically founded gravity model suggests that trade should preferably be treated separately each way (exports from *i* to *j* at time *t* being one observation, and exports from *j* to *i* at time *t* another).

Bronze medal mistake: Inappropriate deflation of trade flows, typically by the US aggregate price index. Gravity is an expenditure function allocating nominal GDP into nominal imports; therefore inappropriate deflation probably creates biases via spurious correlations. Note, however, that time dummies or country effects take this into account. Therefore, if you take care of the gold medal mistake, you cannot get bronze.

3. Advanced gravity modelling issues

a. The issue of zero-trade flows

A much-discussed issue is how to handle zero trade in a given year between two given countries. This is as much an estimation issue as a measurement one and affects all gravity estimation exercises. The problem stems from the fact that the standard way of estimating a gravity model is to take logarithms and estimate its log-linear version. Therefore, zero trade flows will be dropped out of the estimation as the log of zero is not defined.

Traditionally, three alternative approaches have been used to handle zero trade: (i) truncating the sample by dropping the observations with zero trade; (ii) adding a small constant (1 dollar, say) to the value of trade before taking logarithms; or (iii) estimating the model in levels.

The first methodology is correct if the zeros are randomly distributed, e.g. when zeros are random missing data or random rounding errors. The intuition for this is that these zeros are not informative, therefore they can be dropped.

However, if zero trade reported in the data is really zero trade or if it reflects systematic rounding errors associated with very small trade flows, throwing zero trade flows out of the sample will result in a loss of useful information and will yield inconsistent results. For example, if zero trade reflects, say, prohibitive transportation costs due to distance or landlockedness or the smallness of the economies involved, a density mass of observations at zero is, in this case, informative and should be treated as such.

Retaining zero trade flows in the sample requires using appropriate estimation techniques. Strategies (ii) and (iii) are incorrect if an OLS estimation method is used. First, the substitution of small values to prevent the omission of observations from the model is *ad hoc* and there is no guarantee that it reflects the underlying expected values, thus yielding inconsistent estimates. Second, the use of OLS estimation on levels is not supported by theoretically founded gravity equations that present a multiplicative form.

What estimator should be used once we have established the true zeros? The answer to this question depends partially on the reason we believe to be at the origin of the zero trade flows: zeros may also be the result of rounding errors, zeros can simply be missing observations that are wrongly recorded as zeros or zeros may be the result of firms' decision not to export.

Empirical literature on trade has adopted various approaches. One approach often used is to employ a Tobit estimator with left-censoring at zero on the log of trade plus a constant. However, the appropriateness of this approach to solving the zero trade issue has been questioned. The Tobit model reflects a situation where some observations are censored (unobservable) and recorded as zero. The model applies to situations where small values of trade are rounded to zero or actual zero trade might reflect "desired" negative trade. The censoring of trade flows below some positive value is a plausible assumption for some countries but it is hard to believe for other countries for which trade data are reported at a very high degree of accuracy. Therefore, from this perspective the use of a Tobit estimation can only be partially justified. With regard to the second hypothesis, Linders and de Groot (2006: p. 5) noted that "it is unclear which optimizing framework would justify negative desired trade, even if caused by randomly distributed factors not explicitly identified in the

model. As a consequence, the Tobit model is not the appropriate model to explain why some trade flows are missing."

An alternative approach is to use the (Pseudo) Poisson maximum likelihood (ML) estimator. This method can be applied to the levels of trade, thus estimating directly the non-linear form of the gravity model and avoiding dropping zero trade. An influential paper by Santos Silva and Tenreyro (2006) highlights that, in the presence of heteroskedasticity (as is usual in trade data), the PPML is a robust approach. This approach has been used in a number of estimations of gravity equations, such as for example Westerlund and Wilhelmsson (2006).

> In a cross section, the STATA commands to implement these techniques are:
>
> gen lnexports1=ln(exports+1)
> tobit ln(exports1) lndist cont lang ccol col landlock RTA exporter_* importer_*, ll(o) robust
>
> Or
>
> poisson exports lndist cont lang ccol col landlock RTA exporter_* importer_*, robust
>
> In the case of panel data, the commands are *xttobit* and *xtpoisson.* In this case, the fixed effect option (fe) allows us to account for country-pair fixed effects.[10]

Most importantly, it is also possible — indeed, likely — that the probability of having positive (non-zero) trade between two countries is correlated with unobserved characteristics of that country pair. If this is the case, a selection model à la Heckman is called for. In this context, zero trade flows result from the firms' decisions not to export to a certain market. The appropriate estimation procedure is therefore to model these decisions and correct the estimation on the volume of trade for this selection bias. As we will discuss further in subsection 3b, an important limitation in the use of the Heckman approach to solving sample selection bias is that we can only have confidence in the results if we identify a variable that explains firms' decisions to export or not to a certain market but does not affect the volume of trade.

Before turning to a more in-depth discussion of the sample selection problem, it is worth stressing the importance of distinguishing whether zero reported trade between two countries in year *t* is really zero trade and not simply a reporting error. Unlike true zero trade, reporting errors should be treated either by leaving them out of the sample or by interpolating trade values before and after the missing observation (when they are sufficiently stable around a trend). The problem is that, typically, the researcher does not know what he or she is really looking at. Again, there is no perfect fix and it is a matter of judgment how zero-trade observations should be treated. In cross sections, it is virtually impossible to identify real zero trade from missing observations. In panels, a good rule of thumb is to plot the whole time series for the country pair in question. A zero-trade observation sandwiched between ones that are regularly positive should be regarded as suspect. Note that the number of reporting errors is inversely correlated with income per capita (for two reasons: because rich countries have better statistics and because their trade is also large and large numbers are less likely to go unreported). The problem also occurs more frequently when estimating gravity equations at the sectoral level (see Box 3.2) than it does at the aggregate level because it is more common to have no trade in one particular commodity than it is for the entirety of bilateral trade between two countries.

Box 3.2   A cautionary note on gravity models with disaggregated data

Applying the logic of the gravity equation to sectoral trade flows (trade in one particular good) is not entirely straightforward. In the monopolistic-competition model, larger countries produce more varieties of goods and that contributes to increasing their trade. That is, they do not necessarily trade more of each good but they trade more goods. Thus, the idea that trade flows between i and j in a certain sector k are increasing in i's GDP is not necessarily warranted. Recent empirical research (e.g. Hummels and Klenow, 2005) suggests that as economies grow, trade expands both at the extensive margin (more products) and at the intensive one (more volume for each product). Thus, at the risk of finding an insignificant coefficient on exporter GDP, we are by and large justified in using the gravity framework to predict trade in one particular commodity.

When looking at sectoral trade flows, trade barriers are of particular importance. They also matter of course at the aggregate level, but aggregation of trade barriers into overall indices takes a lot of useful information out, which justifies their usual absence from aggregate gravity equations (trade barriers being bunched with country fixed effects or in the error term). When analyzing sectoral trade flows, the bad pretext of aggregation is no longer there, so trade barriers ought to be explicitly in the equation. Indeed, at the sectoral level the gravity equation becomes a good vehicle for analyzing how trade barriers affect trade flows. This is the focus of this case study, in which we will use the simultaneous presence of tariff and non-tariff barriers in the equation to derive a tariff equivalent of the latter on the basis of observed effects on trade flows.

By and large, database construction and estimation issues for sectoral trade flows are as in the previous case study, which used aggregate trade flows. It is worth keeping in mind, however, that GDPs are not always good proxies for demand and supply and that, when country fixed effects are included, they should be interacted with sectoral dummies. Two issues deserve special mention: zero trade and heterogeneity.

The issue of zero trade is of course likely to arise more frequently in sectoral trade flows than in aggregate ones. Again, how we should treat those is a matter of judgment. In some cases — say, bulky goods — frequent occurrence of no-trade country pairs will reflect prohibitive transport costs due to excessive distances or the smallness of the economies involved. A density mass of observations at zero is, in this case, informative and should be treated as such (e.g. using a Poisson or Tobit). In other cases — e.g. agricultural goods whose production requires particular local conditions, like bananas — no-trade country pairs may simply reflect the fact that neither of the countries is suitable for production of the good. For instance, there is no information in the fact that Norway and Sweden do not trade bananas. Such country pairs can be dropped altogether from the sample without information loss. Finally, zero trade flows may result from the self-selection of firms out of the export sector because of high fixed costs of export to certain destinations. In this case a model à la Heckman or — if firm-level information is available — a Tobit model with firm-specific censoring levels (as suggested in Crozet et al., 2009). Use the cnr (censored normal regression) command in STATA to allow for censoring values to change from observation to observation.

Finally, sectoral trade flows are likely to be more heterogeneous than aggregate ones — in which sectoral idiosyncrasies are averaged out — so outliers and heteroskedasticity should be handled with special care.

b. Zero trade and heterogeneity

Trade theory in general assumes that firms are identical and that their behaviour can be characterized by a representative firm. These models can only explain zero trade flows as a measurement error, missing information or as the consequence of prohibitive trade costs, but these factors cannot explain the significant presence of zero trade flows in the data. Helpman, Melitz and Rubinstein (HMR hereafter, 2008) explain zero trade flows across countries in a model with heterogeneous firms, where firms differ in terms of their productivity and there are fixed costs of exporting. In this set up, variable trade costs reduce the amount that exporting firms export, while fixed entry costs reduce the probability that a firm has decided to export. Zero trade costs are associated with high bilateral fixed costs of trade. Another interesting feature of the model is that it can explain asymmetric trade flows between country pairs.

Building on the monopolistic-competition model of heterogeneous firms developed by Melitz (2003), HMR have specified a model where consistent estimates of the value of trade can be obtained following a two-stage procedure. In the first stage, a Probit equation is used to estimate the extent of firms' entry into an export market, which is an unobserved variable in the gravity equation. The first-stage Probit is estimated as:

$$\rho_{ij} = \Pr(T_{ij} = 1) = \Theta(\gamma_0 + \xi_j + \zeta_i - \gamma d_{ij} - \kappa \phi_{ij})$$ (3.12)

where the probability of positive trade flows between i and j, ρ, depends on the importer and exporter dummies (ξ and ζ) and bilateral trade costs, where d denotes variable trade costs and ϕ denotes bilateral fixed costs of entry.

HMR's second stage equation is a gravity model of positive trade values where the results of the first stage are used to correct for the sample selection bias introduced by omitting zero trade flows (the standard Heckman correction term, the inverse of Mill's ratio) and to estimate the (unobserved) share of firms selecting into the export market. The gravity equation augmented by these terms is:

$$x_{ij} = \beta_0 + I_j + I_i - \gamma d_{ij} + \ln\left\{\exp\left[\delta\left(z_{ij} + \eta_{ij}\right)\right] - 1\right\} + \beta_\eta \eta_{ij} + e_{ij}$$ (3.13)

where I_j and I_i denote the exporter and importer individual effects and the term in the curly brackets is the estimator for the share of firms that export to i, z is the fitted variable for the latent variable from the first-stage Probit and η is the inverse Mill's ratio.

Since equation (3.13) is nonlinear in δ, it is estimated using non-linear least squares.

Note that there are several estimation issues related to the use of HMR methodology. First, the use of the Probit with fixed effects in the first stage estimation may induce the so-called "incidental parameters problems", which lead to inconsistent estimation of all parameters of the model, especially in short panels. One possible solution is to use random effects (see Cameron and Trivedi, 2005, p. 786). Second, the estimation of the HMR model requires exclusion restrictions (i.e. the cost variables that enter into the first stage, but not into the second stage equation) to help identification, as regressors are allowed to have different effects on the extensive and the intensive

margin of trade. However, HMR suggest using regulation costs of firm entry or a variable denoting the degree of communality of religion; recent studies suggest that the HMR Probit model is wrongly specified (see Santos Silva and Tenreyro, 2009).

Third, the estimation of the second stage regression may have problematic results in a large sample because of the large number of exporter and importer dummies required. To solve the problem, it may be necessary to use other software.

Notwithstanding these limitations, we provide below the main commands to run the HMR model in STATA. Furthermore, an illustrative example of the HMR estimation method is provided in the application.

```
STATA commands
/* first stage, probit */
probit rho ldist contig colony comlang_off religion xi_* zeta*, robust
* Compute the inverse Mills ratio
predict z_hat, xb
predict pr, pr
gen pdf_z_hat = normalden(z_hat)
gen cdf_z_hat = normprob(z_hat)
gen eta_hat = pdf_z_hat / cdf_z_hat
/* Second stage, non-linear estimation*/
nl (limport = {constant} + {xb: ldist contig colony comlang_off xi1-xiN zeta1-zetaN} +
{etastar}*z_hat + ln(exp(exp({delta=1})*( z_hat + eta_hat )) - 1) ), vce(robust)
*notice the variable religion, the identifier, is not in the second stage
```

c. Measuring overall trade costs and calculating the tariff equivalent of non-tariff barriers

Although the gravity equation is typically used to measure the impact of trade costs on bilateral trade flows, it can also be used in reverse to measure bilateral trade costs and to decompose trade costs into a tariff and non-tariff component (Head and Ries, 2001; Jacks et al., 2008; Novy, 2009). The idea is to solve a theoretical gravity equation for the trade costs term instead of trade flows and to express these costs as a function of the observable trade data. The advantage of this approach compared with the alternative approaches based either on price differences across border[11] or on direct measures of certain trade costs is the lighter data requirement. It is indeed hard to get reliable price data on comparable goods in different countries as well as ascertaining many trade cost components.

Algebraically, the expression for trade costs is readily obtained. The first step is to use the gravity equation (3.2) to find an expression for country i's intra-national trade:

$$X_{ii} = \frac{Y_i Y_i}{Y} \left(\frac{t_{ii}}{\Pi_i P_i} \right)^{1-\sigma} \tag{3.14}$$

where t_{ii} represents intra-national trade costs such as domestic transportation costs. In particular, equation (3.14) implies that intra-national trade does not only depend on a country's economic size but also on multilateral resistances. In particular, the closer is an economy the higher the intra-national trade.

The second step is to multiply the gravity equation (3.2) for X_{ij} by the corresponding gravity equation for trade flows in the opposite direction X_{ji}, $X_{ji} = Y_j Y_i / Y \left(t_{ji} / \Pi_j P_i \right)^{1-\sigma}$, replace (3.14) into this expression and rearrange for trade costs. The resulting expression for trade costs is:

$$\frac{t_{ij} t_{ji}}{t_{ii} t_{jj}} = \left(\frac{X_{ii} X_{jj}}{X_{ij} X_{ji}} \right)^{1/(\sigma-1)}$$

(3.15)

The tariff equivalent of bilateral trade costs relative to domestic trade costs can be expressed as the geometric average of trade barriers in both directions:

$$\tau = \left(\frac{t_{ij} t_{ji}}{t_{ii} t_{jj}} \right)^{1/2} - 1$$

(3.16)

denoting the extent to which international trade is more costly than domestic trade, i.e. bilateral trade costs relative to domestic trade costs. Following this approach, overall trade costs are derived from the gravity equation without imposing a cost function. Furthermore, it is neither assumed that domestic trade costs are zero nor that they are the same across countries (t_{ii} may differ from t_{jj}) nor that bilateral trade costs are symmetric (t_{ij} may differ from t_{ji}).

Novy (2009) shows that similar trade costs measures can be obtained from a variety of models. The difference is in the sensitivity of the implied trade costs to trade flows. This depends on the degree of product differentiation in the Anderson and van Wincoop (2003) model, whereas it depends on the heterogeneity of countries' relative productivities in a Ricardian model and on the degree of firm heterogeneity in a heterogeneous firms model.

Using trade flows data, equation (3.15) can be used to estimate the overall trade costs, i.e. including both tariff and non-tariff barriers. Note, however, that the precise estimate for the *level* of trade cost depends on the parameter of the elasticity of substitution σ (or the parameter denoting the degree of productivity heterogeneity across firms or across countries), while *changes* over time do not. Since there is no consensus in the literature concerning the exact value of these parameters (in general, assumed to fall in the range from 5 to 10 – see Anderson and van Wincoop, 2004), it may be less controversial to look at *changes* of trade costs over time that – changes in elasticity apart – are not affected by the level of the elasticity.

The difficulty in calculating equation (3.15) is to get figures for intra-national trade. One approach is to estimate these figures as the difference between production and exports (see Wei, 2006 and Novy, 2009). The use of GDP instead of production data tends to overstate intra-national trade, and therefore trade costs, because a growing share of GDP is services – largely non-tradable.

Under the specific assumption that domestic trade costs are zero and bilateral trade costs are symmetric (as implied by taking the geometric average as measure of bilateral trade costs), it is also possible to decompose overall trade costs (as calculated from equation (3.15)) into their various cost components by assuming an arbitrary trade cost function, such as for example the log-linear

version of equation (3.6). For instance, it is possible to decompose overall trade costs into their tariff and non-tariff components simply by estimating:[12]

$$\ln \tau_{ij} = \delta_1 \ln \left(\text{distance}_{ij} \right) + \delta_2 \ln \left(1 + \text{tariff}_{ij} \right) + \delta_3 N T B_{ij} + \varepsilon_{ij} \qquad (3.17)$$

where NTB_{ij} is a dummy that denotes the existence of an NTB. Head and Ries estimate the effect of non-tariff barriers (defined to comprise all barriers other than tariffs including transportation costs, home bias, TBTs, etc.) with year dummies. Jacks et al. (2008) estimate a similar equation with random effects as well as year dummies.

```
STATA commands
gen internal_tradeii=productionii-totalexportsi
gen lntotal_trade_costs = ln(exportsij/internal_tradeii *exportsji/internal_tradejj)
regress lntotal_trade_costs lndist lntariff lnNTB
```

In order to calculate the tariff equivalent of a quota, we only need to calculate the tariff that has the same impact on trade costs as a quota.

$$\text{tarrif equivalent}^{13} \; = \; \exp \left(\delta_3 / \delta_2 \right) - 1 \qquad (3.18)$$

Refer to application 2 to further explore this issue.

d. Endogeneity

A problem of endogeneity often arises in gravity models when estimating the impact of trade policies. The typical example is that regional trade agreements (RTAs) are unlikely to be purely exogenous: Countries are likely to form RTAs with partners with which they already trade a lot (following the "natural trading partners" hypothesis). If this is the case, the RTA dummy on the right-hand side of the gravity equation is correlated with the error term because unobserved characteristics of some pairs of countries explain why they trade a lot and at the same time make it more likely that they would form a RTA. Reverse causality apart, endogeneity issues may arise because of omitted variable bias. That is, it may be that RTAs are signed by countries that have other characteristics omitted in the regression (peaceful relationship, common legal origin, etc.) that facilitate trade.

There is not an easy fix to the problem of endogeneity of RTA. In a panel the use of (country-pair) fixed effects can help to overcome part of the endogeneity problem due to the omitted variable bias, although time-varying omitted variables remain a problem.

In general, one would like to use an instrumental variable (IV) approach. The usual problem with IV techniques is finding instruments that are correlated with the PTA dummy but not with trade. There is, unfortunately, no perfect solution to this problem. An instrumental-variable technique developed by Hausman and Taylor (1981) can be used; alternatively, one can use the Generalized Method of Moments (GMM) estimation, in particular system GMM where lagged levels are used as instruments for current differences and vice versa. GMM estimates are, however, typically sensitive to the number of lags used.

Alternatively, endogeneity may be addressed by trying to identify a natural experiment. For example, Frankel (2010) uses the case of the 14 CFA countries as a natural experiment to study the impact

of a currency union on trade. Using firm level data, one may want to remove large firms and study the impact of a RTA on the trade of small firms, as entry in an RTA is likely to be exogenous to small firms' trade. Another way would be to look at the impact of RTAs on the extensive margin of trade.

Recent analysis of the effects of FTA on trade has been using nonparametric (matching) econometric techniques (Baier and Bergstrand, 2006). Other studies have used Probit models to estimate the likelihood that an RTA is formed between two countries; these models use economic and political economy determinants of RTAs (e.g. Mansfield and Reinhardt, 2003; Baier and Bergstrand, 2004; Mansfield et al., 2008). However, these approaches mainly solve for selection bias rather than endogeneity.

4. Data sources

Estimation of a gravity equation requires data on bilateral trade, GDPs, distances, tariffs and possibly other determinants of bilateral trade including contiguity (common border), common language, colonial ties, exchange rates and so on. There is a wealth of databases from which the researcher can draw on for these variables.

As discussed in Chapter 1, bilateral trade flows can be found in the IMF's DOTS, in COMTRADE, in BACI or in the World Bank's Trade, Production and Protection database by Nicita and Olarreaga.[14] In a gravity model, trade flows are typically expressed in current international prices (dollars) and import statistics are generally preferred to export statistics. Sources for tariff data include TRAINS, IDB and CTS.[15] Data on NTBs can be found in TRAINS and Kee, Nicita and Olarreaga.[16]

GDPs in current dollars, converted at current exchange rates, can be found in the IMF's International Financial Statistics (IFS) and in the World Bank's World Development Indicators (WDI; available online) together with a wealth of other indicators. Additional relevant data can also be found in the Penn World Tables (PWT),[17] which currently include time series such as:

- population
- real GDP per capita, in constant dollars and relative to the US at PPP, and growth rates
- exchange rates
- decompositions of national expenditure (consumption, investment and government expenditure, with price indices for each)[18]
- openness (in current and constant prices)
- ratio of GNP to GDP
- current savings.

The PWT in their original form are described in Summers and Heston (1991). The current version, PWT 6.3, covers 189 countries over the years 1950–2007. Not all series are available in PWT 6.3 so earlier versions may be used, but the user should be aware of changes in the base year (2005 in version 6.3, 2002 in version 6.2, 2000 in version 6.1, 1996 in version 6.0 and 1985 in version 5.6).

A complete gravity database is available on the website of Thierry Mayer.[19] A good source of macro data already in STATA is the MACRODAT for STATA website.[20] Some gravity covariates datasets

are available at Andrew Rose's and CEPII's websites.[21] These include bilateral distances — typically the "great-circle" distances between capitals, i.e. the shortest distances measured on the earth's surface irrespective of actual highways or sailing routes — common language, common border, whether either country is an island, whether either country is landlocked, etc. See Haveman's webpage.[22] Information on existing RTAs and the year of their entry into force as well as some of the contents of RTAs can be found at the WTO website.[23]

C. Applications

1. Building a database and estimating a gravity model

Estimation of a gravity equation requires a substantial upfront investment in data collection and organization of the data. One reason is that the estimation of a gravity model typically involves a large database. This has an advantage and a drawback. The advantage is that with a large sample, estimation is typically precise and stable. The drawback is that large samples are cumbersome to work with and use up computing power.

The large size of the typical database of a gravity model is due to the basic principle that even if the researcher is interested only in the factors influencing a particular trading relationship — say, whether the creation of a certain free trade area has diverted trade from a certain country — the effect should be measured on the basis of a gravity equation estimated for *all* countries, not just the countries involved. In addition, gravity equations can be estimated for either cross sections or panels of countries. In the first case, the unit of observation is a pair of countries; so with n countries there are $n(n-1)$ observations. In the second case the unit of observation is a pair of countries in a year, so there are $Tn(n-1)$ observations with T being the number of time periods covered by the panel. Thus sample size in a gravity equation is typically very large (in a panel of 100 countries over 10 time periods, there are 100,000 observations).[24] However, whenever possible panels should be preferred.

Another difficulty related to building a database for a gravity model is that data from a variety of different data sources (see subsection B.4) have to be merged in a single database. Since data may be available in different formats or classifications, the researcher needs to invest some time in organizing this information. This section helps to address some of the problems that may emerge in this process.

As an example of the typical steps required to build a database for the estimation of a gravity model, we describe here the simple case of a standard gravity model estimated on data aggregated at the country level and augmented with a dummy denoting WTO membership (see Rose, 2004). For this purpose we extract raw data for bilateral trade for the period 1990–2005 from the UN COMTRADE, GDP data for the same time period from the World Bank World Development Indicators, the set of bilateral covariates (distance, common language, border and so on) from the CEPII website and get information on the GATT/WTO accession from the WTO website. The do files we will refer to in this subsection are provided in the folder "Chapter3\Applications\1_Building a database and estimating a gravity model".

For illustrative purposes we will distinguish nine steps:

Step 1: Import data into STATA

While data from CEPII are provided in STATA format, data on GDP, trade flows and WTO accession are in .txt, .csv or .xls format (data can be found in the folder *Datasets/Original* or *Datasets/Stata*). Obviously, we will need to import the data in STATA to be able to manipulate them. Provided that all data are in STATA format, we can proceed with the data manipulation. Note that since data on GDP are provided for the Benelux region rather than Belgium and Luxembourg separately, we will need to adjust all covariate variables for which data are available for the two countries separately in a way that matches GDP data. There are different methodologies to match covariate data to GDP data: (1) to pick one of the two countries as representative of the region, (2) to take the average of the variable (e.g. distance) between the two countries. In the do file we will show the commands for this matching.

To import data in STATA use the "insheet" command.

```
insheet tradeflows.cvs, clear
save tradeflos.dta
insheet joinwto.cvs, clear
save joinwto.dta
```

Step 2: Create all possible country-pairs–year combinations

Let's first check whether the data on bilateral trade flows are correct and complete. It is always good first to check randomly that the data have been properly put in. For this you should randomly select some observations and go back to the original dataset to verify whether they are correct.

Then, you may want to get a complete database of bilateral trade including zero trade flows. The command *fillin* in STATA will create all possible country-pair–year combinations. You must end up with n*n*T observations, where n is the number of countries and T the time period. Then, if you accept the assumption that all non-reported information is zero (Helpman et al., 2007 adopt this approach), you may want to replace all missing observations with zeros.

```
use tradeflows.dta, clear
fillin importer exporter year
replace imports=0 if imports==.
```

Step 3: Reshape and merge country-specific data with bilateral trade flows

In our database each observation is identified by three indicators: importer, exporter and the year. Trade data are in "long" format, i.e. information for each variable is displayed in the form of a vector.

Data on GDP extracted from WDI are instead in "wide" format, i.e. they are in the form of a matrix n×T. Before merging GDP country-specific data into our database we need to transform the matrix into a vector data form. The command *reshape* in STATA allows this transformation (note in the reshape command the letters i, j are used to indicate the cross section and the time dimension of the observations). In particular, given a file where the GDP figures of countries A, B, C for various years, say 1990–1992, are reported in columns named as gdpyear. That is, given a file of the form:

```
(wide form)
i ....... variable_ij ........
country gdp1990 gdp1991 gdp1992
-------------------------------
A 5000 5500 6000
B 2000 2200 3300
C 3000 2000 1000
```

we can reshape it as:

```
(long form)
i j variable_ij
country year gdp
----------------------
A 1990 5000
A 1991 5500
A 1992 6000
B 1990 2000
B 1991 2200
B 1992 3300
C 1990 3000
C 1991 2000
C 1992 1000
```

by simply using the command

```
reshape long gdp, i(country) j(year)
```

Country-specific information (e.g. GDP data in long format) can be merged into the bilateral database with trade flows information over time. However, you need to specify whether the data refer to the exporting or the importing country before merging. A simple way to do this is to save the data on GDP twice: the first time you will refer to the variable country as the exporter and save the file as gdp_exporter, the second time you will refer to the country as the importer and save the file as gdp_importer. In STATA, simply use the commands:

```
use gdp.dta, clear
rename country exporter
rename gdp gdp_exporter
sort exporter year
save gdp_exporter.dta
use gdp.dta, clear
rename country importer
rename gdp gdp_importer
sort importer year
save gdp_importer.dta
```

Now you can merge the two GDP data files with trade flows. Recall that you need to sort data by the variable you intend to use to match observations in the two databases. For example, to merge the file with data for GDP for the exporting country use the STATA commands:

```
use tradeflows, clear
sort exporter year
merge exporter year using gdp_exporter /*gdp_exporter must be sorted by exporter year*/
save gravity
```

The same procedure can be used to merge the database with the information about a country's year of accession to the WTO.

Step 4: Merge with pair-specific data

Use the command *merge* in STATA to merge bilateral covariates from CEPII. As an alternative it would be possible to use *joinby*. The advantage of using *merge* is that STATA in this case automatically creates an additional variable called "merge_" that allows you to check that all observations across databases match. Note that since the information does not vary over time, the importer and exporter identifiers are the only matching variables.

```
use gravity.dta, clear
sort exporter importer
merge exporter importer using cepii.dta /*cepii.dta must be sorted by exporter and importer*/
```

Step 5: Generate new country-pair variables

Bilateral variables can be generated using country-specific information. Suppose, for example, that like Rose (2004) we want to study the impact on trade of WTO accession. Provided that we have information on the year in which a country joins the WTO, we can easily generate a set of variables that denote whether both the exporter and the importer are WTO members (Bothin) or only one of the two countries is a WTO member (Onein).

```
/*Generate dummies for WTO membership status */
gen onein=0
gen bothin=0
replace onein=1 if (join_exporter<=year & join_importer>year) | (join_importer<=year & join_
exporter>year) /* the symbol | corresponds to the operator "or" */
replace bothin=1 if (join_exporter<=year & join_importer<=year)
```

Step 6: Generate dummies

As discussed above a simple way to obtain unbiased estimations of a gravity equation is to use country fixed effects. In a cross section (e.g. when data are only available for a certain year), it is possible to use country dummies on the exporter and importer side — often called "country fixed effects" — because the unit of observation in a gravity model is a pair of countries, not a country. Hence, while there are n^2 observations in a cross section, there will be n country dummies on the exporter side and n on the importer side using a total of $2n$ (clearly less than n^2) degrees of freedom. Note, however, that using exporter and importer fixed effects in a cross section, you will not be able to estimate the coefficients on country-specific variables (e.g. GDP) due to perfect collinearity.

```
/* Country dummies */
tab exporter, gen(exporter_)
tab importer, gen(importer_)
```

In a panel, time effects (a set of dummy variables, one for each year) should be included in order to control for global economic effects (booms or slowdowns in the global economy). In addition, a set of time-varying fixed effects both for the exporter and the importer should be included to control for the time-varying multilateral resistance terms. Again, note that time-varying fixed effects are perfectly collinear with time-varying country-specific variables. Therefore, they render estimation of the impact of country-specific variables such as GDP impossible.

```
/*Time dummies*/
tab year, gen(year_)
```

```
/*Country-time dummies*/
tab gdp_exporter, gen(exporteryear_)
tab gdp_importer, gen(importeryear_)
```

There may be circumstances when estimates may be improved by using country-pair effects. Assume, for example, that there is a monopoly over shipping lines between two countries. This would reduce trade between those two but would not affect the trade of each of the two with third countries. It would therefore not be controlled for by country effects. Country-pair dummies can be easily generated in STATA:

```
/*Pair dummies*/
egen pairid = group(importer exporter)
tab pairid, gen(pair_)
```

Note that in a panel it is possible to model country-pair factors as random rather than fixed effects (see Brun et al., 2005 or Carrère, 2006).[25] This approach preserves the possibility of estimating separately the effect of bilateral factors such as distance, common borders etc. that would otherwise be confounded with the fixed effects. However, as with country effects, random-effect estimation bunches unobserved country-pair characteristics in the error term. If they happen to be correlated with some of the regressors (like GDPs), estimates are inconsistent. This can be tested for by using a Hausman test.

Table 3.1 shows a snapshot of the STATA editor with gravity database.

Step 7: Data transformation

Typically, the estimated gravity equation takes a log-linear form. We therefore need to take the natural logarithm of the continuous variables in the equation (i.e. except dummies). Note that in so doing the zero trade flows will drop out of the sample.[26]

```
gen limports=ln(imports)
/* idem for gdp exporter, gdp importer and distance*/
```

Table 3.1 Excerpt from an illustrative gravity database

	importer	exporter	year	imports	gdp_exporter	gdp_importer	dist	onein	bothin	nonein
15657	IDN	EMU	1999	9.659	3.324e+09	1.400e+11	16992.25	0	1	0
15658	IDN	BOL	1999	219.247	8.285e+09	1.400e+11	17453.36	0	1	0
15659	IDN	BRA	1999	316304.1	5.870e+11	1.400e+11	15644.48	0	1	0
15660	IDN	BRB	1999	.85	2.468e+09	1.400e+11	18349.78	0	1	0
15661	IDN	BRN	1999	37654.51	4.600e+09	1.400e+11	1533.213	1	0	0
15662	IDN	BTN	1999	0	3.967e+08	1.400e+11	4169.54	1	0	0
15663	IDN	BWA	1999	0	5.623e+09	1.400e+11	8807.14	0	1	0
15664	IDN	CAF	1999	3763.993	1.051e+09	1.400e+11	9863.057	0	1	0
15665	IDN	CAN	1999	421145.6	6.610e+11	1.400e+11	15815.39	0	1	0
15666	IDN	CHE	1999	131234	2.680e+11	1.400e+11	11222.69	0	1	0
15667	IDN	CHL	1999	88932.25	7.300e+10	1.400e+11	15614.28	0	1	0
15668	IDN	CHN	1999	1242141	1.080e+12	1.400e+11	5220.879	1	0	0
15669	IDN	CIV	1999	40985.74	1.256e+10	1.400e+11	12371.63	0	1	0
15670	IDN	CMR	1999	2768.479	1.049e+10	1.400e+11	10640.69	0	1	0
15671	IDN	COG	1999	16.813	2.354e+09	1.400e+11	10134	0	1	0
15672	IDN	COL	1999	1768.534	8.630e+10	1.400e+11	19772.34	0	1	0
15673	IDN	COM	1999	31.135	2.229e+08	1.400e+11	6995.042	0	1	0
15674	IDN	CPV	1999	46.621	5.834e+08	1.400e+11	14514.75	1	0	0
15675	IDN	CRI	1999	142.145	1.580e+10	1.400e+11	18767.43	0	1	0
15676	IDN	CUB	1999	137.164	.	1.400e+11	17903.42	1	0	0
15677	IDN	CYM	1999	0	.	1.400e+11	18324.11	1	0	0
15678	IDN	CYP	1999	69.055	9.780e+09	1.400e+11	8922.735	0	1	0
15679	IDN	CZE	1999	10698.96	6.019e+10	1.400e+11	10705.36	0	1	0

Note: The variables are, in the order in which they appear on the table:

importer	The exporter's 3-letter ISO code (here Indonesia)
exporter	The importer's 3-letter ISO code
year	Year
imports	Imports value (one way), in dollars
gdp_exporter	The exporter's GDP (from WDI or PWT) in current dollars
gdp_importer	The importer's GDP (from WDI or PWT) in current dollars
dist	Distance between the importer and exporter countries (from CEPII database)
onein	Only one of the two countries in the pair is a WTO member
bothin	Both countries in the pair are WTO members
nonein	None of the two countries in the pair are WTO members

It is important to recall that if you need to average flows over a certain period, theoretical foundations suggest that when you are averaging, the log transformation should come first (this is a geometric mean).

```
/*period (5 year) averages */
gen period=1
replace period=2 if year<2000 & year>=1995
replace period=3 if year>=2000
gen limports=ln(imports)
collapse (mean) limports lgdp_exporter ldist, by(period) /* non-time varying covariates can also be added
to the variable list*/
save av_gravity
```

Step 8: Identify the panel and run regressions

Panel data refer to the situation when information on bilateral trade flows is available over time. Once the individual and time dimensions of the panel have been identified, a number of easy commands in STATA allow you to generate lag variables or growth rates. The command *xtdes* allows you to describe the pattern of panel data, e.g. whether the panel is balanced or not.

```
use gravity.dta, clear
/* to identify the dimension of the panel*/
xtset pairid year
/*generate the first-lag variable of imports*/
gen Llimports=L1.limports
/* generate the growth rate of import*/
gen Gimports=limports-Llimports
```

Now the database is ready for you to run regressions. Look at the discussion above on estimation methods and the STATA commands reg, tobit, ivreg, heckman, poisson, nl along with their time series versions (e.g. xtreg, xttobit, xtpoisson). For example, the simplest gravity model

```
*for a cross section
xi: reg lnimports lndist other covariates i.exporter i.importer if year=2000, robust
```

```
* for a panel
xtreg lnimports lndist other covariates exporter_* importer_* year_*, robust fe /* where the option fe
provides for country-pair fixed effects*/
```

Step 9: Presenting and interpreting regression results

Professional presentations of regression results should always include the following information:

- dependent variable
- unit of observation (country pair, country pair/time)
- number of observations
- estimation method
- for each estimation method, a list of coefficient estimates in column with t-statistics (preferably to standard errors for ease of interpretation) in parentheses below, and variable names that the reader can comprehend
- indications of what sets of dummy variables were used (fixed effects, time effects, etc.) without typically reporting estimates on those
- R-squared and standard tests (depending on the case, over-identification, properties of residuals like homoskedasticity or absence of serial correlation, etc.).

A special STATA package called "outreg2" that makes it fast and easy to present results in the standard way (parameter estimates with t-stats in parentheses below, stars for levels of significance, and variable labels instead of names, which improves readability) can be freely downloaded from the web using the "findit outreg2" command in STATA.

As gravity estimates of the influence of distance and GDPs on trade have proved remarkably stable across studies, they provide checks to determine how reasonable the estimates are from a new study. Most of the variables are expressed in natural logarithms, so coefficients obtained from linear estimation can be read directly as elasticities. The elasticity of trade to distance, for instance, is usually between −0.7 and −1.5, so a 10 per cent increase in distance between two countries cuts their trade, on average, by 7 to 15 per cent. Elasticities with respect to importing-country GDPs are also typically unitary, suggesting unitary income elasticities of imports at the aggregate level. Estimates departing significantly from these orders of magnitude would signal that something is probably wrong in either estimation or measurement.

Note, however, that while the coefficients for the natural logarithm of continuous variables (e.g. GDP, distance) are elasticities, the coefficients for the dummies (such as a dummy denoting whether two countries belong to the same trade agreement) are not. They need to be transformed as follows in order to be interpreted as elasticities: elasticity = exp(a)−1 where a is the estimated coefficient of the dummy variable.[27] As stressed by Anderson and van Wincoop (2003) and Feenstra (2004), however, the multilateral resistance terms are essential in order to estimate the general equilibrium, for instance comparative static effects on bilateral trade of the existence of a regional trade agreement between two countries. Once estimated with a gravity model, one would need to retrieve estimates for the multilateral resistance terms Π_i and P_j of equation (3.2) with and without a regional trade agreement and derive the effect of a regional trade agreement on trade using the equation:

$$\text{elasticity} = e^a \left(\Pi_i^{*1-\sigma} / \Pi_i^{1-\sigma} \right) \left(P_j^{*1-\sigma} / P_{ji}^{1-\sigma} \right) - 1 \qquad (3.19)$$

where * denotes the estimate of the multilateral resistance term for the two countries without a regional trade agreement.

2. Measuring the effect of NTBs

As shown in subsection B.3, gravity models can be used to retrieve the tariff equivalent of non-tariff barriers. In this application we will study the case of the European Union (EU) bananas regime as an example of a special regime combining tariffs and quotas. While the theoretical arguments have been presented for a general case, in practice economic literature generally assumes symmetric trade costs, zero domestic trade costs and that trade costs are of the form specified in equation (3.6) for estimating tariff equivalents. This application will retain these assumptions. Data and do files can be found in the folder "Chapter3\Applications\2_Measuring the effect of NTBs".

Let τ_{ijt} and Q_{ijt} be respectively any tariff and quota imposed by j on i in the sector in question (here bananas) at time t, and other variables be as before. In its general formulation, a gravity equation estimated at the sectoral level looks like this:

$$\ln X_{ijt} = \beta_0 + \beta_1 \ln\left(1 + \tau_{ijt}\right) + \beta_2 \ln Q_{ijt} + \beta_3 \ln GDP_{it}$$
$$+ \beta_4 \ln GDP_{jt} + \beta_5 \ln dist_{ij} + \beta_6 I_i + \beta_7 I_j + \sum_{\ell=1}^{T} \beta_{7+\ell} I_\ell + u_{ijt} \qquad (3.20)$$

127

Note that 1 has been added to the tariff because, the equation being in logs, zero tariffs would send the log to minus infinity whereas $\ln(1) = 0$.

Table 3.2 gives regression results for a gravity equation estimated for worldwide banana trade over 1989–2004. In particular, regression results are reported for the case in which exporter and importer exchange rates. The dummy for the Framework Agreement and ACP countries are included as control variables to account for these countries' preferential regime. Observe first that, as usual, row headings indicate the name of the explanatory variable. Thus, for example, the number -1.150 that appears in column 1 and in the row market "$\ln(1 + \tau)$, applied" is the estimated value for the coefficient β_1 in the equation above and gives an approximation to the price elasticity of importing countries' demand for bananas, estimated "on average" for all years and countries.[28] With sufficiently many countries and no domestic production in the importing countries, it is also the elasticity of substitution.

The elasticity of banana trade flows relative to distance is close to unity, as in most gravity equations (however, the present exercise cannot be directly compared in this regard with standard gravity estimates as we are dealing with a single commodity with special transport arrangements). The elasticity with respect to the importing country's GDP can be taken as a rough approximation of the income elasticity of banana consumption (since there is no domestic production in most importing countries) and is less than unity (between 0.72 and 0.84 depending on the estimation method). This is probably a slight underestimate, since the share of bananas in household budgets tends to rise with income (bananas are a "luxury").

Exchange rates are significant and with the expected sign for exporter countries (banana exports go up when the exporting country's currency depreciates *vis-à-vis* the dollar) but not for importing countries. With the exception of the Framework Agreement dummy, special-regime variables are highly significant and with the expected sign.

We can now use these estimates to retrieve the tariff equivalent of the quota. Let us use hats over variables for estimated coefficients and predicted trade values. Labelling Z_{nt}^{ij} the explanatory variables other than the quota and bunching together the other coefficients from β_0 to β_5, we have

$$\ln \hat{X}_{ijt} = \sum_{n \neq 2} \hat{\beta}_n Z_{ijnt} + \hat{\beta}_2 Q_{ijt} \tag{3.21}$$

Note that Q_{ijt} is equal to one if a quota applies and zero otherwise. Thus, the predicted difference in trade between a country pair *with* a quota and the same country pair *without* the quota would be

$$\ln \hat{X}_{ijt,quota} - \ln \hat{X}_{ijt,noquota} = \sum_{n \neq 2} \hat{\beta}_n Z_{ijnt} + \hat{\beta}_2 (1) - \left[\sum_{n \neq 2} \hat{\beta}_n Z_{ijnt} + \hat{\beta}_2 (0) \right]$$
$$= \hat{\beta}_2 \tag{3.22}$$

A similar calculation can be performed for the effect of a tariff at rate τ_t^{ij} compared to no tariff at all:

$$\ln \hat{X}_{t,\text{tariff}}^{ij} - \ln \hat{X}_{t,\text{no tariff}}^{ij} = \sum_{\ell \neq 4} \hat{\beta}_\ell Z_{\ell t}^{ij} + \hat{\beta}_4 \left(\ln\left(1 + \tau_t^{ij}\right) \right) - \left[\sum_{\ell \neq 5} \hat{\beta}_\ell Z_{\ell t}^{ij} + \hat{\beta}_4 (1) \right]$$
$$= \hat{\beta}_4 \ln\left(1 + \tau_t^{ij}\right) \tag{3.23}$$

Table 3.2   Gravity estimates, banana market

Dep. var.: trade value	OLS	OLS robust	OLS robust	iterative	iterative
1n(1+τ), applied	−1.150**	−1.150**		−1.261***	
	−0.488	−0.497		−0.469	
1n(1+τ), unconstrained			−1.195**		−1.136**
			−0.582		−0.537
1n(1+τ), constrained			−1.065		−1.486**
			−0.675		−0.69
MFN quota dummy	−0.671***	−0.671***	−0.691***	−0.515***	−0.459**
	−0.163	−0.165	−0.205	−0.157	−0.2
Framework Agr. dummy	0.426	0.426	0.428	0.395	0.387
	−0.29	−0.321	−0.322	−0.278	−0.279
ACP dummy	1.046***	1.046***	1.044***	0.993***	0.997***
	−0.218	−0.24	−0.24	−0.21	−0.21
Ivory Coast * time trend	0.0871	0.0871	0.0877	0.157**	0.156**
	−0.0652	−0.0756	−0.0754	−0.0627	−0.0628
Cameroon * time trend	0.211***	0.211***	0.212***	0.260***	0.258***
	−0.0743	−0.0787	−0.0788	−0.0714	−0.0716
log distance	−1.119***	−1.119***	−1.119***	−1.269***	−1.271***
	−0.0631	−0.0753	−0.0753	−0.0607	−0.0607
log importer's GDP	0.852***	0.852***	0.853***	0.729***	0.723***
	−0.28	−0.303	−0.304	−0.269	−0.269
log exporter's GDP	0.178	0.178	0.177	0.251	0.252
	−0.219	−0.222	−0.222	−0.21	−0.21
log importer's exch. rate	−0.0514	−0.0514	−0.0511	−0.0335	−0.0346
	−0.0884	−0.0793	−0.0794	−0.0851	−0.0851
log exporter's exch. rate	0.0701**	0.0701**	0.0703**	0.0859***	0.0854***
	−0.032	−0.0274	−0.0274	−0.0308	−0.0308
Constant	−14.44*			−321.5**	−11.74
	−8.063			−125.8	−7.754
Observations	6,983	6,983	6,983	6,969	6,983
R-squared	0.58	0.81	0.81	0.62	0.62
Cook-Weisberg chi-sq.	4.03				
Implied tariff equiv.					
(€/ton)	**346**	**346**	**343**	**221**	**158**

Source: Author calculations

Notes: Iterative estimates (fourth and fifth columns) were obtained using STATA's rreg procedure, which uses iteratively reweighted least squares with Huber and biweight functions, where the more extreme an outlier is, the less heavily it gets weighted in the regression calculations. Very extreme cases get dropped altogether. An alternative is qreg (quantile regression), whose most common form is median regression. Median regression, also called Least Absolute Value regression, minimizes the sum of the absolute value of the residuals rather than their squares, thus giving less weight to outliers. The Cook-Weisberg test statistic shows that the homoskedasticity assumption is rejected at the 10% level. The variable labelled "1n(1+τ), applied" uses the tariff applied on bananas by importing countries, irrespective of other measures they may have. The variable labelled "1n(1+τ), constrained" uses the tariff applied by the EU on dollar bananas post-1994; the variable labelled "1n(1+τ), unconstrained" uses the tariff applied in all other cases.[29]

A tariff equivalent of quota Q_t^{ij} is a tariff that has the same effect on trade flows. This is equivalent to equating the left-hand sides of (3.22) and (3.23). But if their left-hand sides are equal, so are their right-hand sides; thus, the tariff equivalent $\tilde{\tau}$ of quota Q_t^{ij} satisfies

$$\hat{\beta}_1 \ln\left(1 + \tau_t^{ij}\right) = \hat{\beta}_2 \qquad (3.24)$$

or

$$\tilde{\tau} = \exp\left(\hat{\beta}_2 / \hat{\beta}_1\right) - 1 \tag{3.25}$$

This simple calculation can be programmed after the estimation of the gravity equation, yielding an *ad valorem* tariff equivalent to the EU's quota. In a final step, observed unit values can be used to translate the *ad valorem* equivalent into a specific form.

Tariff equivalents calculated in accordance with this method and converted into specific rates on the basis of a CIF unit value of €438/ton are reported at the bottom of Table 3.2. The last estimate (€158/ton) is obtained using as "β_1" the coefficient on the tariff on QR-constrained markets. This choice has advantages and disadvantages. On the one hand, what is measured is the effect of in-quota tariffs which might be expected not to be binding (although the data suggests that they are binding). On the other hand, because the QR-constrained market is the EU's, the coefficient measures the elasticity of substitution on the EU market which, if the assumption of constant elasticity of substitution is not taken at face value, is better than using the elasticity of substitution applying to other markets. Note that the estimate of €158/ton does not include the €75/ton in-quota tariff, so it is equivalent to a price gap of 158 + 75 = 233 euros per ton. An interesting feature of the €158/ton estimate is that it is roughly in line with anecdotal estimates of the market value of import licences.

```
reg lnvalue lnApptariff quotaregime frameworkregime ACPregime CIVtime CMRtime lndistance lnmGDP
lnxGDP lnmrate lnxrate Y2-Y15 M2-M96 X2-X118
```

```
* Compute the quota's specific tariff equivalent, unit value= 438 euros/ton
gen t_advalorem = exp(_b[quotaregime] / _b[lnApptariff]) - 1
gen t_specific = t_advalorem * 438
```

Note that, following the approach described in this application, non-tariff barriers other than quotas such as norms can be converted into tariff equivalents using the same methodology. This is potentially important to measure the market-access restrictions implied by sanitary and phyto-sanitary measures in foodstuffs, for instance.[30]

D. Exercises

1. Estimating the impact of a regional trade agreement

The aim of this exercise is to assess the impact on trade of the formation of a regional trade agreement. The exercise focuses on NAFTA and refers to the 1985–2004 period, but a similar exercise can be applied to other PTAs provided the sample period covers sufficient years before and after the agreement. The do file "AnalyzingBilateralTradeUsingGravity.do" including solutions can be found at "Chapter3\Exercises\1_Estimating the impact of a Regional Trade Agreement". (use STATA 10 for this exercise).

1. Preliminaries
    Import in STATA the Trade, Protection and Production data base constructed by the World Bank (Nicita and Olarreaga 2006) available in the folder the folder "Chapter3\Datasets" and merge it with trade data in "Chapter3\Datasets\BilateralTrade.dta" for the period 1985–2004. Then aggregate data at the country level. Note that this process may take time since you need to load a large database. For this reason you may also need to set memory at a high level (e.g. 800m). The do file to construct the data can be found at "Chapter3\Exercises\Preliminary\TPPGravity.do".
    Hints: insheet, forvalues, foreach, append, merge, collapse
    a. Define a dummy variable equal to one for intra-NAFTA trade (i.e. between Canada, Mexico and the US since 1994, the year of entry into force of NAFTA), and another one for each NAFTA's country import from the rest of the world (i.e. build the dummies to measure trade creation and trade diversion).
        Hints: generate
    b. Plot the temporal evolution of intra-NAFTA imports and of NAFTA's imports from the rest of the world. Comment.
        Hints: twoway tsline, xtline, graph save
2. Average trade creation and diversion
    a. Introduce the two above dummies in a gravity equation (fixed effects model with country-pair fixed effects and year dummies) estimated on all countries, not just the countries involved.
        Hints: tsset, gen, replace, tabulate, xtreg, outreg2
    b. Present the results and interpret the coefficients associated with the dummies (i.e. quantify the trade creation/diversion).
        Hints: outreg2
3. Evolution of trade creation and diversion
    a. Interact the two NAFTA dummies (denoting intra and extra NAFTA country pairs) with year effects and introduce them into the gravity equation (with still a fixed effects model with country-pair fixed effects and year dummies).
        Hints: levelsof, foreach, gen, replace, xtreg, outreg2
    b. Interpret the evolution of the coefficients in terms of trade creation and trade diversion.
4. Export diversion
    a. Add one dummy capturing the exports of NAFTA towards the rest of the world and add this third RTA dummy in the gravity equation.
        Hints: gen, replace, xtreg, outreg2
    b. Repeat Questions 1b, 2b, 3a and 3b using this new variable in addition to the others. Comment on all your results.
        Hints: collapse, gen, replace, twoway tsline, graph save, levelsof, foreach, tsset, xtreg, outreg2.

2. Calculating tariff equivalent

The aim of this exercise is to use the gravity model to measure the tariff equivalent of non-tariff trade barriers. The exercise follows the approach developed by Jacks et al. (2008). The do files "i_NTB.do" and "ii_Tariff_Equiv.do" including the solutions can be found in "Chapter3\Exercises\2_ Calculating tariff equivalent".

1. Preliminaries

    Import in STATA the Trade, Protection and Production data base constructed by the World Bank (Nicita and Olarreaga 2006) available in the folder "Chapter3\Datasets" and merge it with trade data in "Chapter3\Datasets\BilateralTrade.dta" and the standard gravity variables in "Chapter3\Datasets\GravityData.dta". Then aggregate data at the country level. Note that this process may take time as you need to load a large database. For this reason you also may need to set memory at a high level (e.g. 800m). The do file to construct the data can be found at "Chapter3\Exercises\Preliminary\TPPGravity.do".

    Hints: insheet, forvalues, foreach, append, merge, collapse

    a. Build the reporter and partner country tariff variables: $tariff_i$ and $tariff_j$. Then compute domestic consumption variables for the importer and the exporter by taking the difference between output and total exports.

        Hints: generate, joinby

    b. Bilateralize tariffs by multiplying importer and exporter tariffs and then taking logs, i.e. $t_{ij} = \ln\left(tariffs_i \cdot tariffs_j\right)$. Calculate the tariff equivalent of total trade costs as in equation (3.16), $\tau_{ij} = \left(X_{ii}X_{jj} / X_{ij}X_{ji}\right)^{1/[2(\sigma-1)]} - 1$, using an elasticity of substitution set to $\sigma = 11$.

2. Gravity estimation

    a. Generate a dummy variable equal to one for the existence of core non-tariff barriers implemented.

    b. Estimate the following gravity equation with country-pair fixed effects.

    $$\tau_{ijt} = \beta_0 + \beta_1 \ln\left(distance_{ij}\right) + \beta_2 \ln\left(t_{ijt}\right) + \beta_3 Quota_{ijt} + TimeDummies + u_{ijt}$$

    Then, compare the results when the country-pair effects are assumed to be random and when the country-pair fixed effects are replaced by exporter and importer dummies.

    Hints: reg, xtreg

3. Tariff equivalent

    a. For each gravity estimation result, determine the tariff equivalent of the non-tariff barriers.

        Hints: estimates store, estimate restore

    b. Check for the sensitivity of the trade cost measure, by re-estimating the gravity equation and computing the tariff equivalent, assuming the elasticity of substitution is set to $\sigma = 5$ or $\sigma = 15$. Comment.

Endnotes

1. A clear and concise introduction to the gravity equation can be found in Head (2003). A thorough treatment for the advanced reader is in Chapter 5 of Feenstra (2004).
2. See Feenstra et al. (2001), Evenett and Keller (2002) and Feenstra (2004).
3. This is a model built along the lines of Melitz (2003), where firms face fixed and variable costs of exporting. Firms vary by productivity and only the more productive firms will find it profitable to export.
4. Dummies are variables that only assume the value zero or one. For example, a dummy denoting whether the importing country is an island takes the value one for all observations when the importing country is indeed an island and zero otherwise.
5. Alternative methodologies to estimate the gravity models without directly controlling for MRT rely on manipulations of the basic gravity equation aiming at getting rid of one or both MRTs. One methodology consists in normalizing bilateral trade flows with respect to a certain country-pair (Martin et al., 2008 and Romalis, 2008). Another method is the so-called "Tetrads method" (Head and Mayer's website provides the do file for estimations conducted with this methodology) that consists in taking ratio of ratios so as to remove both MRTs.
6. Recall that the unit of observation (the "individual") in a gravity equation is a pair of countries, not a single country. So when we speak of "country effects", those are not identical with "fixed effects" in the standard meaning, which refers to dummy variables marking each individual pair.
7. This can be done most easily in STATA using the "robust" option after the estimation command "regress". In cases where outliers create particular problems, alternative estimation methods, like Huber's, can be preferable to OLS. Outliers can be detected and dropped from the estimation. To detect outliers use the Hadi test (STATA command: hadimvo). Another problem associated with using country-pair observations is that the errors may be correlated across country-pairs (that is, country-pair observations are not independent). To address this problem, one should correct for cluster errors. In STATA, use the option "cluster" or the command "cgmreg" for the multi-way clustering.
8. Note that "time effects" should not be confused with a "time trend". The former are dummy variables, one per year, as explained in section 1. The latter is a variable that starts from 1 and goes up one unit each period: 1 for the first, 2 for the second, and so on. The former is more general because it does not impose the assumption of a linear time trend. For instance, if production is depressed in year t because of a hurricane, this temporary drop will be picked up by the variable I_t, which will have a negative coefficient. It won't be picked up, in contrast, by a time trend.
9. The latter in fact can be obtained as a linear combination of the set of fixed effects.
10. Note that the standard Poisson model is vulnerable to problems of overdispersion and excess zero flows. To overcome these problems, Brueger et al. (2009) suggest using modified Poisson fixed-effects estimations (negative binomial, zero-inflated). STATA commands to run these estimations are: "nb" for the negative binomial, and "zip" and "zinb" for the zero-inflated Poisson and negative binomial, respectively.
11. The motivation for using price gaps as a measure of trade costs is that in absence of such costs arbitrage will equalize prices.
12. See Head and Ries (2001) and Jacks et al. (2008).
13. For small changes this formula is equivalent to that calculated in Kee et al. (2009), where tarrif equivalent $= (\exp(\delta 3) - 1) / \delta 2$.
14. See http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESEARCH/0,,contentMDK: 2108 5384~pagePK:64214825~piPK:64214943~theSitePK:469382,00.html for the latter.
15. See Chapter 2.
16. Found at http://siteresources.worldbank.org/INTRES/Resources/469232-1107449512766/OTRI_INDICES_2008.xlsx as of April 2011.
17. PWT Mark 6.3 is freely available on the web at http://pwt.econ.upenn.edu/php_site/pwt_index.php. It can also be found with a different data-extraction interface at the CHASS center of the University of Toronto at http://datacentre2.chass.utoronto.ca/pwt/. Country codes are not identical across databases but concordance tables can be found on Jon Haveman's page at http://www.macalester.edu/research/economics/PAGE/HAVEMAN/Trade.Resources/Concordances/OthMap/country.txt as of April 2011.

18. Note that as country price indices combine a PPP adjustment with an exchange rate, they can change because of either inflation without exchange rate adjustment or exchange-rate change without inflation. This can cause year-to-year jumps.

19. Found at http://econ.sciences-po.fr/thierry-mayer/publications as of April 2011.

20. Found at http://www.graduateinstitute.ch/md4stata as of April 2011.

21. Found at http://www.cepii.fr/anglaisgraph/bdd/gravity.htm as of April 2011.

22. Found at http://www.macalester.edu/research/economics/page/haveman/trade.resources/tradedata.html as of April 2011.

23. Found at http://rtais.wto.org/UI/PublicMaintainRTAHome.aspx as of April 2011.

24. In practice, however, there are always many country/year pairs with missing data, either for trade or, more frequently, for other variables the researcher may be interested in (e.g. transportation infrastructure levels).

25. The commands to use in STATA in the case of a panel for fixed and random effect estimations are xtreg …, fe and xtreg …, re, respectively.

26. See discussion above on estimations with zero trade flows.

27. To derive this formula, consider that $\ln X_{ij}(1)$ is the predicted value of trade when the dummy is equal to 1 (e.g. two countries are in a regional trade agreement, share a common language, a border, etc.) while $\ln X_{ij}(0)$ is the value of trade when the dummy takes the value 0. It follows that the difference $\ln X_{ij}(1) - \ln X_{ij}(0) = a$, where a is the estimated coefficient for the dummy variable. It follows that $X_{ij}(1)/X_{ij}(0) = \exp(a)$, which in turn implies that the percentage change in trade value due to the dummy switching from 0 to 1 is: $X_{ij}(1) - X_{ij}(0)/X_{ij}(0) = \exp(a) - 1$.

28. It is not exactly the price elasticity of import demand because trade is measured in dollars, so what we are really measuring is $d\ln(p^*q)/d\ln(1+\tau)$ instead of $d\ln(q)/d\ln(1+\tau)$. However, the price that is implicit in the value of trade flows, p^*, is the world price (which is why we put a star on it); if the tariffs affect a sufficiently small share of world trade, the world price can be considered as fixed and the approximation is acceptable.

29. For country-pairs with no QRs, the variable "$\ln(1+\tau)$, constrained" is zero, but then the variable "$\ln(1+\tau)$, unconstrained" is positive. They cannot both be zero.

30. For a survey of methods for the quantification of non-tariff barriers, see e.g. Maskus et al. (2001) or Beghin and Bureau (2001).

References

Anderson, J. E. and van Wincoop, E. (2003), "Gravity with gravitas: a solution to the border puzzle", *American Economic Review* 93: 170–92.

Anderson, J. E. and van Wincoop, E. (2004), "Trade costs", *Journal of Economic Literature* 42: 691–751.

Baier, S. L. and Bergstrand, J. H. (2004), "Economic determinants of free trade agreements", *Journal of International Economics* 64(1): 29–63.

Baier, S. L. and Bergstrand, J. H. (2006), "Estimating the effects of free trade agreements on trade flows using matching econometrics", mimeo, published in *Journal of International Economics* (2009) 77(1): 63–76.

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

Baier, S. L. and Bergstrand, J. H. (2009), "Bonus vetus OLS: a simple method for approximating international trade-cost effects using the gravity equation", *Journal of International Economics* 77(1): 77–85.

Baldwin, R. and Taglioni, D. (2006), "Gravity for dummies and dummies for gravity equations", National Bureau of Economic Research Working Paper 12516, NBER.

Beghin, J. and Bureau, J.-C. (2001), "Quantification of sanitary, phytosanitary, and technical barriers to trade for trade policy analysis", Iowa State University, Center for Agricultural and Rural Development Working Paper 01-WP-291.

Bergstrand, J. H. (1985), "The gravity equation in international trade: some microeconomic foundations and empirical evidence", *The Review of Economics and Statistics* 67(3): 474–81.

Bergstrand, J. H. (1989), "The generalized gravity equation, monopolistic competition and the factor-proportions theory in international trade", *The Review of Economics and Statistics* 71(1): 143–53.

Brun, J.-F., Carrère, C., Guillaumont, P. and de Melo, J. (2005), "Has distance died? Evidence from a panel gravity model", *World Bank Economic Review* 19: 99–120.

Cameron, A.C. and Trivedi, P. K. (2005), *Microeconometrics: Methods and Applications*, Cambridge University Press.

Carrère, C. (2006), "Revisiting the effects of regional trade agreements on trade flows with proper specification of the gravity model", *European Economic Review* 50: 223–47.

Chaney, T. (2008), "Distorted gravity: the intensive and extensive margins of international trade", *American Economic Review* 98: 1707–21.

Crozet, M., Head, K. and Mayer, T. (2009), "Quality sorting and trade: firm-level evidence for French wine", Centre for Economic Policy Research Discussion Paper 7295, CEPR.

Dixit, A. and Norman, V. (1980), *Theory of International Trade*, Cambridge University Press.

Eaton, J. and Kortum, S. (2002), "Technology, geography and trade", *Econometrica* 70: 1741–79.

Ethier, W. (1998), "Regionalism in a multilateral world", *Journal of Political Economy* 106: 1214–45.

Evenett, S. and Keller, W. (2002), "On theories explaining the gravity equation", *Journal of Political Economy* 110: 281–316.

Feenstra, R. (2004), *Advanced International Trade*, MIT Press.

Feenstra, R., Markusen, J. and Rose, A. (2001), "Using the gravity equation to differentiate across alternative theories of trade", *Canadian Journal of Economics* 34: 430–47.

Frankel, J. (2010), "The estimated effects of the Euro on trade: why are they below historical evidence on effects of monetary unions among smaller countries?", in Alesina, A. and Giavazzi, F. (eds.), *Europe and the Euro*, University of Chicago Press.

Hausman, J. A. and Taylor, E. (1981), "Panel data and unobservable individual effects", *Econometrica* 49: 1377–98.

Head, K. (2003), "Gravity for beginners", mimeo, University of British Columbia.

Head, K. and Ries, J. (2001), "Increasing returns versus national product differentiation as an explanation for the pattern of US–Canada trade", *The American Economic Review* 91(4): 858–76.

Helpman, E., Melitz, M. and Rubinstein, Y. (2008), "Trading partners and trade volumes", *Quarterly Journal of Economics* 123: 441–87.

Hummels, D. and Klenow, P. J. (2005), "The variety and quality of a nation's exports", *American Economic Review* 95: 704–23.

Jacks, D. S., Meissner, C. M. and Novy, D. (2008), "Trade costs, 1870–2000", *American Economic Review: Papers and Proceedings* 98(2): 529–34.

Kee, H.L., Nicita, A. and Olarreaga, M. (2009), "Estimating trade restrictiveness indices", *The Economic Journal* 119: 172–199.

Krugman, P. (1980), "Scale economies, product differentiation and the pattern of trade", *American Economic Review* 70: 950–9.

Limao, N. (2006), "Preferential trade agreements as stumbling blocks for multilateral trade liberalization: evidence for the United States", *American Economic Review* 96: 896–914.

Linders, G. M. and de Groot, H. L. (2006), "Estimation of the gravity equation in the presence of zero flows", Tinbergen Institute Discussion Paper 2006–072/3.

Magee, C. (2003), "Endogenous preferential trade agreements: an empirical analysis", *Contributions to Economic Analysis and Policy* 2, Article 15.

Maskus, K., Wilson, J. and Otsuki, T. (2001), "Quantifying the impact of technical barriers to trade: a framework for analysis", in Maskus, K. and Wilson, J. (eds.), *Quantifying the Impact of Technical Barriers to Trade: Can It Be Done?*, University of Michigan Press.

Melitz, M. and Ottaviano, G. (2008), "Market size, trade and productivity", *Review of Economic Studies* 75: 295–316.

Novy, D. (2009), "Gravity redux: measuring international trade costs with panel data", mimeo, University of Warwick.

Rose, A. and van Wincoop, E. (2001), "National money as a barrier to international trade: the real case for currency union", *American Economic Review* 91(2): 386–90.

Santos Silva, J. and Tenreyro, S. (2006), "The log of gravity", *The Review of Economics and Statistics* 88: 641–58.

Santos Silva, J. and Tenreyro, S. (2009), "Trading partners and trading volumes: implementing the Helpman-Melitz-Rubinstein model empirically", Center of Economic Performance (CEP) Discussion Paper 935.

Summers, R. and Heston, A. (1991), "The Penn World Table (Mark 5): an expanded set of international comparisons, 1950–1988", *Quarterly Journal of Economics* 101: 327–68.

Tinbergen, J. (1962), *Shaping the World Economy*, Twentieth Century Fund.