

# 2

## **The Basis for the Gravity Model: From Intuition to Theory**

## 2 The Basis for the Gravity Model: From Intuition to Theory

---

This section provides a general grounding in the gravity model and its relationship to the data. We start the analysis intuitively, and state the basic ideas lying behind the gravity model. We then use descriptive statistics and graphical techniques to investigate whether those ideas are in fact reflected in a dataset of bilateral trade in services. In the second part of the section, we discuss some of the limitations of the traditional, or intuitive, gravity model. This discussion leads into the third part of the section, in which we discuss gravity theory. Over the last decade, theory has become an increasingly important part of gravity modeling, and applied researchers now have the choice among a number of commonly-used theoretically-grounded gravity models. We provide the intuition behind such models in the main text, and then present full derivations of the most commonly used one – the Anderson and Van Wincoop (2003) “gravity with gravitas” model – in an appendix.

### 2.1 The Gravity Intuition

As noted above, the gravity model was initially presented as an intuitive way of understanding trade flows. In its most basic form, the gravity model can be written as follows:

$$\log X_{ij} = c + b_1 \log GDP_i + b_2 \log GDP_j + b_3 \log \tau_{ij} + e_{ij} \quad (1a)$$

$$\log \tau_{ij} = \log(\text{distance}_{ij}) \quad (1b)$$

where  $X_{ij}$  indicates exports from country  $i$  to country  $j$ , GDP is each country's gross domestic product,  $\tau_{ij}$  represents trade costs between the two countries, distance is the geographical distance between them – as an observable proxy for trade costs – and  $e_{ij}$  is a random error term. The  $c$  term is a regression constant, and the  $b$  terms are coefficients to be estimated. The name “gravity” comes from the fact that the nonlinear form of equation 1a resembles Newton's law of gravity: exports are directly proportional to the exporting and importing countries' economic “mass” (GDP), and inversely proportional to the distance between them (not the square of the distance between them, as in physics). In other words, gravity says that we expect larger country pairs to trade more, but we expect countries that are further apart to trade less, perhaps because transport costs between them are higher.

Before moving to the details of gravity modeling using econometric methods, we can use graphical techniques to examine the basic intuition underlying the model. For all empirical examples in this user guide, we use a dataset on bilateral trade in services compiled by Francois et al. (2009). Their original data has been re-aggregated using a programme supplied by the authors so that the sectoral definition follows the one used by the Global Trade Analysis Project (GTAP). We will usually be interested in total (aggregate) services trade, which is in the sector SER. Readers are free to use the remaining sectors to extend the examples presented here. To facilitate the replication of examples, this user guide provides all necessary Stata code, and it can be used jointly with the dataset accompanying it, which also includes standard gravity model control variables from standard sources such as CEPII, as well as additional variables explained later in the text.

A first step in examining the intuition behind the gravity model is to examine the correlations among the variables. To do that, we first use Stata's *generate* command to transform the dataset into logarithms. We then use the *correlate* command to produce a correlation matrix, as in Table 1. By adding the *if* option, we limit the calculation to the SER sector, i.e. total services trade.

**Table 1: Correlation matrix for basic gravity variables**

```
. gen ln_trade = ln(trade)
(19012 missing values generated)

. gen ln_distance = ln(dist)
(606 missing values generated)

. gen ln_gdp_exp = ln(gdp_exp)
(991 missing values generated)

. gen ln_gdp_imp = ln(gdp_imp)
(992 missing values generated)

. correlate ln_trade ln_gdp_exp ln_gdp_imp ln_distance if sector == "SER"
(obs=3884)
```

	ln_trade	ln_gd~xp	ln_gd~mp	ln_dis~e
ln_trade	1.0000			
ln_gdp_exp	0.3643	1.0000		
ln_gdp_imp	0.3731	-0.3103	1.0000	
ln_distance	-0.2648	0.0518	0.0431	1.0000

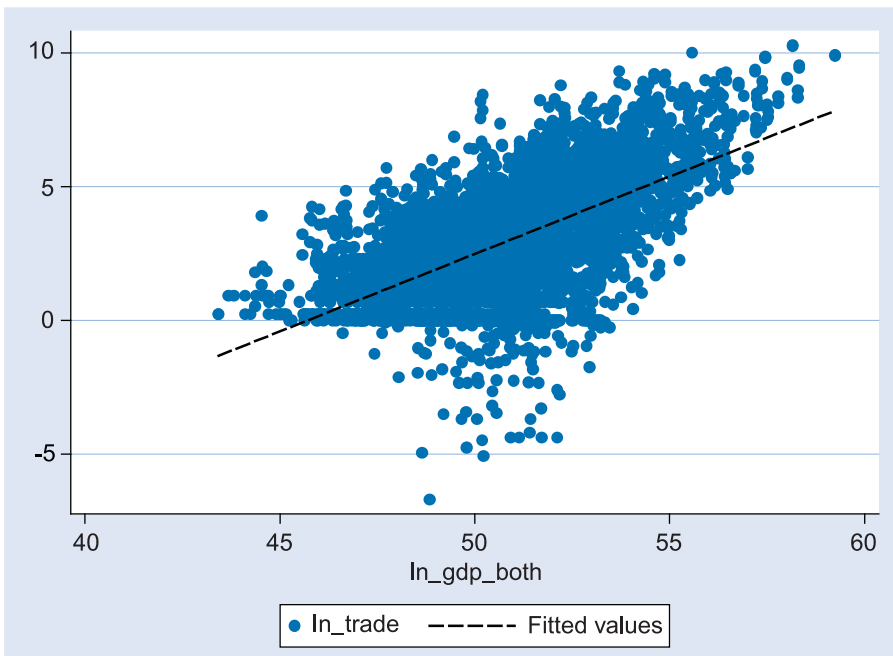
The correlations of interest are contained in the non-diagonal elements of the matrix. We see that trade and GDP are strongly positively correlated, and that the correlation is approximately the same for exporter and importer GDP. This finding supports the basic intuition that bigger countries tend to trade more. By contrast, we find a strong negative correlation between trade and distance: country pairs that are further apart tend to trade less. Again, this finding is in line with the basic intuition of the gravity model.

We can use Stata's *tway* command to present the same information graphically. That command allows us to overlay graphs one on top of the other, and we use it to combine a scatter plot of the variables of interest with a linear line of best fit, which reflects the correlation between them. Figure 1 shows the result using the combined economic mass of the exporting and importing countries, i.e. the product of their GDPs, as the explanatory variable. The scatter plot shows a clear positive association between the two variables, in line with the correlation analysis. Similarly, the line of best fit is strongly upward sloping. Graphical evidence therefore also supports the basic intuition that larger country pairs tend to trade more than smaller ones.

**Figure 1: Scatter plot and line of best fit for trade versus combined GDP**

```
. gen ln_gdp_both = ln(gdp_exp*gdp_imp)
(1983 missing values generated)

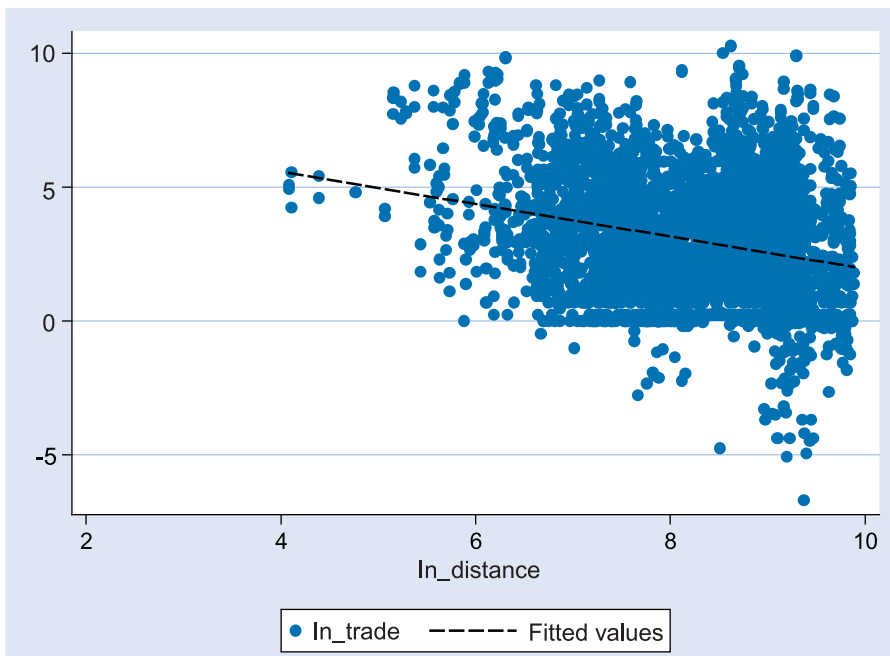
. twoway (scatter ln_trade ln_gdp_both if sector == "SER") (lfit ln_trade
ln_gdp_both if sector == "SER")
```



We can also use graphical methods to investigate the association between trade and distance. Figure 2 presents results using the same approach as for Figure 1. In this case, the scatter plot is suggestive of a negative association. That impression is reinforced by the line of best fit, which is strongly downward sloping. Graphical evidence again confirms the basic gravity intuition that country pairs that are further apart tend to trade less.

**Figure 2: Scatter plot and line of best fit for trade versus distance**

```
. twoway (scatter ln_trade ln_distance if sector == "SER") (lfit ln_trade
ln_distance if sector == "SER")
```



## 2.2 Problems with the Intuitive Gravity Model

The previous section shows that the basic gravity model picks up two important stylized facts of international trade: bigger countries trade more, and more distant countries trade less. These regularities are almost uniformly reflected in the early gravity literature, which applies the model to all regions of the world, covering both developed and developing economies, and various products and time periods. The model is clearly a useful starting point in applied international trade research.

However, the intuitive gravity model is not free of difficulties once more advanced concepts from the trade literature are introduced. As one example, consider the impact on trade between countries  $i$  and  $j$  of a change in trade costs between countries  $i$  and  $k$ . An example of such a change might be that countries  $i$  and  $k$  enter into a preferential trade agreement that lowers tariffs on their respective goods. Basic economic theory suggests that such a move may well impact the trade of country  $j$ , even though it is not party to the agreement. The well-known

concepts of trade creation and trade diversion are examples of such effects. However, the intuitive gravity model does not account for this issue at all. As is clear from equation 1a,  $\frac{\partial \log X_{ij}}{\partial \log t_{ik}} = 0$ . Reducing trade costs on one bilateral route therefore does not affect trade on other routes in the basic model, which is at odds with standard economic theory.

A second problem with the basic model arises if we consider equal decreases in trade costs across all routes, including domestic trade (goods that a country sells internally, or  $X_{ii}$ ). An example could be a fall in the price of oil, which lowers transport costs everywhere, including within countries. In the basic model, this move would result in proportional increases in trade across all bilateral routes, including domestic trade. However, such an outcome sits ill with the observation that despite the change in trade costs, relative prices have not changed at all. In the absence of a change in relative prices, we would expect consumption patterns to remain constant for a given amount of total production (GDP). This is a second instance in which the basic gravity model makes predictions that are at odds with standard economic theory.

### 2.3 “Theoretical” Gravity Models

Issues such as those identified in the previous section led some researchers to turn to theory to provide a basis for a gravity-like model of trade. Clearly, the basic model needs to be altered in some fundamental dimensions if it is to deal with the issues raised in the previous section. The first example of such an approach is Anderson (1979). However, the applied literature has only paid serious attention to theoretically-grounded gravity models since the famous “gravity with gravitas” model of Anderson and Van Wincoop (2003). In this section, we present the basic intuition behind that model, with the mathematical details left to the Appendix.

At its most basic, the Anderson and Van Wincoop (2003) gravity model is essentially a demand function. It owes much of its final form to the constant elasticity of substitution structure chosen for consumer preferences. Consumers have “love of variety” preferences, which means that their utility increases both from consuming more of a given product variety, or from consuming a wider range of varieties without consuming more of any one.

On the production side, the model makes assumptions that are standard following Krugman (1979). Each firm produces a single, unique product variety under increasing returns to scale. By assuming a large number of firms, competitive interactions disappear and firms engage in constant markup pricing: in equilibrium, the difference between price and marginal cost is just enough to cover the fixed cost of market entry.

A producer in one country can sell goods in any country, either the one where it is located, or an overseas country. To simplify the model, selling goods locally is assumed to involve no transport

costs. However, selling goods internationally does involve transport costs. Consumers therefore consume product varieties from all countries, but the prices of non-domestically produced varieties are adjusted upwards to take account of the cost of moving goods between countries.

These building blocks make it possible to derive an equilibrium in which firms both produce for the local market and engage in international trade, and in which consumers consume accordingly. The basic model provides expressions for the volume of exports by each firm. Aggregating across firms within an economy then makes it possible to derive an expression for the total value of a country's exports, which is the dependent variable in the gravity model.

All that is required to produce a gravity-like model from these foundations is to impose some macroeconomic accounting identities. Essentially, these identities flow from the fact that in a single sector economy like the one being modeled, where there are no input-output relationship, the sum of all production must be equal to GDP. Performing the aggregation in an appropriate way makes it possible to derive the "gravity with gravitas" model:

$$\log X_{ij}^k = \log Y_i^k + \log E_j^k - \log Y^k + (1 - \sigma_k) [\log \tau_{ij}^k - \log \Pi_i^k - \log P_j^k] \quad (2a)$$

$$\Pi_i^k = \sum_{j=1}^C \left\{ \frac{\tau_{ij}^k}{P_j^k} \right\}^{1-\sigma_k} \frac{E_j^k}{Y^k} \quad (2b)$$

$$P_j^k = \sum_{i=1}^C \left\{ \frac{\tau_{ij}^k}{\Pi_i^k} \right\}^{1-\sigma_k} \frac{Y_i^k}{Y^k} \quad (2c)$$

where  $X$  is exports indexed over countries ( $i$  and  $j$ ) and sectors ( $k$ ),  $Y$  is GDP,  $E$  is expenditure (which is not necessarily the same as GDP on a sectoral basis),  $Y^k = \sum_{i=1}^C Y_i^k$  (i.e., world GDP),  $\sigma_k$  is the intra-sectoral elasticity of substitution (between varieties), and  $\tau_{ij}^k$  is trade costs.

The first notable feature of the Anderson and Van Wincoop (2003) model is its inclusion of two additional variables,  $\Pi_i^k$  and  $P_j^k$ . The first is called outward multilateral resistance, and it essentially captures the fact that exports from country  $i$  to country  $j$  depend on trade costs across all possible export markets. The second is called inward multilateral resistance, and it likewise captures the dependence of imports into country  $i$  from country  $j$  on trade costs across all possible suppliers. Together, these terms are the key to the model, and they resolve both issues identified as problems with the intuitive gravity model in the previous section. In particular, it is immediately apparent that because the multilateral resistance terms involve trade costs across all bilateral

routes,  $\frac{\partial \log X_{ij}}{\partial \log t_{ik}} \neq 0$ . In other words, this model picks up the fact that changes in trade cost on one bilateral route can affect trade flows on all other routes because of relative price effects. Because the intuitive model does not include these two multilateral resistance variables but they are, by construction, correlated with trade costs, there is a classic case of omitted variables bias in the intuitive model. Finding a way to correct for this problem will be the main thrust of the estimation approaches discussed in the remainder of this user guide.

The second point to note is that the theoretical gravity model has a number of implications for the way in which a gravity model should be set up, and the types of data that should be used. In the early gravity model literature, some authors used dependent variables such as the logarithm of total trade for a country pair (the sum of exports and imports) or the average of exports in both directions. The theoretical gravity model suggests that such an approach is likely to produce misleading results. The model applies to unidirectional export flows, which means that each line in a gravity database should represent a single flow. Thus, exports from Australia to New Zealand are recorded in one line of the database, and exports from New Zealand to Australia are recorded in a separate one.

Another question that has arisen in the literature is whether trade values should be expressed in nominal or real terms. For the standard cross-sectional gravity model, of course, nothing turns on this issue: data for a single year will give equivalent results regardless of any uniform scaling factor applied. In a time series context, however, this question can be important. The answer given by theory is clear: trade flows should be in nominal, not real, terms. The reason is that exports are effectively deflated by the two multilateral resistance terms, which are special price indices. Deflating exports using different price indices, such as the CPI or the GDP deflator, would not adequately capture the unobserved multilateral resistance terms, and could produce misleading results.

A similar analysis applies to the GDP data used in the model: they too should be in nominal, not real, terms. The reason is again that they are effectively deflated by the multilateral resistance terms, which are unobserved price indices. Deflating by some other factor, such as a readily observable price index, is likely to be misleading. In addition, the sectoral version of the gravity model makes clear that ideally we would like to include data on sectoral expenditure and output rather than GDP as such. However, this is usually impossible in an empirical context – particularly when developing countries are included in the sample – so GDP is used as an acceptable proxy. Finally, the model makes clear that it is aggregate GDP, not per capita GDP, which should be used. Expedients sometimes encountered in the older gravity literature, such as using population and per capita GDP as separate regressors, should therefore be avoided.

The last part of the model that needs to be specified for estimation purposes is the trade costs function  $\tau_{ij}^k$ . The literature typically specifies this function in terms of observable variables that



are believed to influence trade costs, using a simple log-linear specification. In the examples in this user guide, we generally specify the trade costs function as follows:

$$\log \tau_{ij}^k = b_1 \log distance_{ij} + b_2 contig + b_3 comlang\_off + b_4 colony + b_5 comcol \quad (3)$$

Distance is the geographical distance between countries  $i$  and  $j$ , *contig* is a dummy variable equal to unity for countries that share a common land border, *comlang\_off* is a dummy variable equal to unity for country pairs that share a common official language, *colony* is a dummy variable equal to unity if countries  $i$  and  $j$  were once in a colonial relationship, and *comcol* is a dummy variable equal to unity for country pairs that were colonized by the same power. This formulation is typical of the gravity model literature, in which each of these factors has been found to be a significant determinant of bilateral trade. However, this specification is by no means exhaustive. We use it as a baseline for the examples in this user guide, but researchers will generally need to augment it in applied work to include other variables, especially trade-related policies.

One point that needs to be emphasized in relation to the trade costs function is that it is impossible to separate the elasticity of substitution  $\sigma_k$  from the trade cost elasticities (the  $b$  terms) when estimating the model. The two are always multiplied together. This fact has important implications for the way in which gravity models with multiple sectors are estimated, which we discuss in the next section. More broadly, it suggests that we should be wary of interpreting differences in estimated coefficients across sectors as evidence of different levels of sensitivity to particular trade cost factors: we might simply be observing differences in the elasticities of substitution. In order to obtain pure trade cost elasticities, we would need to interact the trade cost variables with estimates of the substitution elasticities, either using a general assumption, or model-based estimates such as those in Broda and Weinstein (2006). However, this path is generally not followed in applied work using the gravity model.

Although we have focused here on the Anderson and Van Wincoop (2003) gravity model, the literature provides a variety of other theoretically-grounded models. For example, Chaney (2008) and Helpman et al. (2008) develop gravity-like equations based on underlying models of trade in which firms are heterogeneous in productivity. Although there are of course important differences among the exact forms of gravity produced by these models, they all retain some fundamental similarities with the basic model set out at the beginning of this section. The same is true of the Ricardian model due to Eaton and Kortum (2002). Applied researchers are therefore free to choose from among a number of theoretical gravity models when developing an estimating model for particular purposes. In the current literature, however, it is increasingly important that a model be as theoretically grounded as possible. It is becoming more and more difficult to justify atheoretical gravity models, based on ad hoc specifications.

# Appendix 2.1: Derivation of the Anderson and Van Wincoop (2003) Gravity Model

## Consumption Side

Consider a world of  $C$  countries indexed by  $i$ . We assume from the start that countries can trade with each other, and thus that consumers in one country can potentially purchase varieties from any other country. For the moment, trade is costless. Consumers are identical in each country, and maximize CES utility across a continuum of varieties (index  $v$ ) in  $K$  sectors (indexed by  $k$ ) with the following form:

$$U_i = \sum_{k=1}^K \left\{ \int_{v \in V_i^k} [x_i^k(v)]^{1-\frac{1}{\sigma_k}} dv \right\}^{\frac{1}{1-\frac{1}{\sigma_k}}} \quad (4)$$

The set  $V_i$  defines the range of varieties that is consumed in country  $i$ . As usual, we use  $x_i^k(v)$  to indicate the quantity of variety  $v$  from sector  $k$  consumed in country  $i$ , and  $p_i^k(v)$  to indicate its unit price. We use function notation because of the continuum of varieties. In the version of the model with a discrete number of varieties,  $v$  becomes a subscript, and the integrals are replaced with sums.

The utility function is simply the sum of the sectoral sub-utilities, each of which is weighted equally. That restriction can easily be relaxed by aggregating the sectoral sub-utilities via a Cobb-Douglas utility function, and allowing for different weights. So long as the shares are exogenous to the model, however, the basic results stay the same. See Chaney (2008) for an example of what the alternative expressions look like. Anderson and Van Wincoop (2003) and Helpman et al. (2008) consider, in effect, a single sector so as to avoid cluttering up the algebra with additional indices. But nothing turns on this, and in the present case it is useful to retain some sectoral disaggregation so that we can examine a couple of important data implications that flow from the model in a multi-sector context.

The budget constraint in country  $i$  is:

$$E_i = \sum_{k=1}^K \left\{ \int_{v \in V_i^k} p_i^k(v) x_i^k(v) dv \right\} \equiv \sum_{k=1}^K E_i^k \quad (5)$$

where  $E_i$  is total expenditure in that country, and  $E_i^k$  is country  $i$ 's total expenditure in sector  $k$ .

The consumer's problem is to choose  $x_i^k(v)$  for all  $v$  so as to maximize (4) subject to (5). The Lagrangian is:

$$\mathcal{L} = \sum_{k=1}^K \left\{ \int_{v \in V_i^k} [x_i^k(v)]^{1-\frac{1}{\sigma_k}} dv \right\}^{\frac{1}{1-\frac{1}{\sigma_k}}} - \lambda \sum_{k=1}^K \left\{ \int_{v \in V_i^k} p_i^k(v) x_i^k(v) dv \right\} \quad (6)$$

Taking the first order condition with respect to quantity and setting it equal to zero gives:

$$\frac{\partial \mathcal{L}}{\partial x_i^k(v)} = \frac{1}{1-\frac{1}{\sigma_k}} \left\{ \int_{v \in V_i^k} [x_i^k(v)]^{1-\frac{1}{\sigma_k}} dv \right\}^{\frac{1}{1-\frac{1}{\sigma_k}}-1} \left(1-\frac{1}{\sigma_k}\right) [x_i^k(v)]^{\frac{1}{\sigma_k}} - \lambda p_i^k(v) = 0 \quad (7)$$

Define  $X^k = \left\{ \int_{v \in V_i^k} [x_i^k(v)]^{1-\frac{1}{\sigma_k}} dv \right\}^{\frac{1}{1-\frac{1}{\sigma_k}}}$ , regroup terms, and rearrange to get:

$$\frac{[x_i^k(v)]^{\frac{1}{\sigma_k}}}{\int_{v \in V_i^k} [x_i^k(v)]^{1-\frac{1}{\sigma_k}} dv} X^k = \lambda p_i^k(v) \quad (8)$$

Now rearrange again, multiply through by prices, aggregate over all varieties in a given sector, and then solve for the Lagrangian multiplier:

$$p_i^k(v) x_i^k(v) = \lambda^{-\sigma_k} [p_i^k(v)]^{1-\sigma_k} (X^k)^{\sigma_k} \left\{ \int_{v \in V_i^k} [x_i^k(v)]^{1-\frac{1}{\sigma_k}} dv \right\}^{-\sigma_k} \quad (9)$$

$$\begin{aligned} & \int_{v \in V_i^k} p_i^k(v) x_i^k(v) dv \equiv E_i^k \\ & = \lambda^{-\sigma_k} [p_i^k(v)]^{1-\sigma_k} (X^k)^{\sigma_k} \left\{ \int_{v \in V_i^k} [x_i^k(v)]^{1-\frac{1}{\sigma_k}} dv \right\}^{-\sigma_k} \int_{v \in V_i^k} [p_i^k(v)]^{1-\sigma_k} dv \end{aligned} \quad (10)$$

$$\lambda = \left\{ \frac{\int_{v \in V_i^k} [p_i^k(v)]^{1-\sigma_k} dv}{E_i^k} \right\}^{\frac{1}{\sigma_k}} \frac{X_k}{\int_{v \in V_i^k} [x_i^k(v)]^{1-\frac{1}{\sigma_k}} dv} \quad (11)$$

To get the direct demand function, substitute this expression for the Lagrangian multiplier back into the first order condition (8):

$$\frac{[x_i^k(v)]^{-\frac{1}{\sigma_k}}}{\int_{v \in V_i^k} [x_i^k(v)]^{1-\frac{1}{\sigma_k}} dv} X^k = \left\{ \frac{\int_{v \in V_i^k} [p_i^k(v)]^{1-\sigma_k} dv}{E_i^k} \right\}^{\frac{1}{\sigma_k}} \frac{X_k}{\int_{v \in V_i^k} [x_i^k(v)]^{1-\frac{1}{\sigma_k}} dv} p_i^k(v) \quad (12)$$

$$\therefore x_i^k(v) = \frac{[p_i^k(v)]^{-\sigma_k}}{\int_{v \in V_i^k} [p_i^k(v)]^{1-\sigma_k} dv} E_i^k \equiv \left\{ \frac{p_i^k(v)}{P_i^k} \right\}^{-\sigma_k} \frac{E_i^k}{P_i^k} \quad (13)$$

where  $P_i^k = \left\{ \int_{v \in V_i^k} [p_i^k(v)]^{1-\sigma_k} dv \right\}^{\frac{1}{1-\sigma_k}}$  is the ideal CES price index for sector k in country i.

## Production Side

The producer's problem in this model is to maximize profit. Assuming a continuum of firms, i.e. an uncountably large number of them, makes this problem much easier to solve. It turns out that strategic interactions disappear, and firms charge a constant markup. In terms of the overall model, this section gives us an equilibrium pricing equation which, with the equilibrium demand equation derived in the previous section, is just about all we need to generate gravity.

Each country i has a measure  $N_i^k$  of active firms in sector k. Each firm makes a unique product, so the total worldwide measure of products in each sector is  $\sum_{i=1}^C N_i^k$ . To produce one unit of its product, a firm must pay a fixed cost  $f_i^k$  and a variable cost  $a_i^k$ . With the wage rate equal to w, a typical firm's profit function is therefore:

$$\pi_i^k(v) = p_i^k(v)x_i^k(v) - wa_i^kx_i^k(v) - wf_i^k \quad (14)$$

With a continuum of varieties, it does not matter at this point whether we assume Bertrand (price) or Cournot (quantity) competition. If the firms play Bertrand, the first order condition is:

$$\frac{\partial \pi_i^k(v)}{\partial p_i^k(v)} = x_i^k(v) + p_i^k(v) \frac{\partial x_i^k(v)}{\partial p_i^k(v)} - wa_i^k \frac{\partial x_i^k(v)}{\partial p_i^k(v)} = 0 \quad (15)$$

Solving for prices gives:

$$p_i^k(v) = wa_i^k - \frac{x_i^k(v)}{\frac{\partial x_i^k(v)}{\partial p_i^k(v)}} \quad (16)$$

To do something with that expression, we need to know more about the partial  $\frac{\partial x_i^k(v)}{\partial p_i^k(v)}$ . In fact, it can be directly evaluated using the demand function (13) and noting that due to the large group assumption (a continuum of firms)  $\frac{\partial P_i^k}{\partial p_i^k(v)} = 0$ . In other words, a small change in one firm's price does not affect the overall level of prices in the sector because so many firms are competing. In light of this, we can write:

$$\frac{\partial x_i^k(v)}{\partial p_i^k(v)} = -\sigma_k [p_i^k(v)]^{-\sigma_k - 1} \left\{ \frac{1}{P_i^k} \right\}^{-\sigma_k} \frac{E_i^k}{P_i^k} = -\frac{\sigma_k x_i^k(v)}{p_i^k(v)} \quad (17)$$

The first order condition for profit maximization can therefore be rewritten as:

$$p_i^k(v) = wa_i^k + x_i^k(v) \frac{p_i^k(v)}{\sigma_k x_i^k(v)} \quad (18)$$

then rearranged and solved for prices to give:

$$p_i^k(v) - \frac{1}{\sigma_k} p_i^k(v) \equiv p_i^k(v) \left( 1 - \frac{1}{\sigma_k} \right) = wa_i^k \quad (19)$$

$$\therefore p_i^k(v) = \left( \frac{\sigma_k}{\sigma_k - 1} \right) wa_i^k \quad (20)$$

The second term on the right hand side in equation (20) is simply the firm's marginal cost of production. The term in brackets is a constant (within sector) markup: since the numerator must be greater than the denominator, there is a positive wedge between the firm's factory gate ("mill") price and its marginal cost. Since the wedge only depends on the sectoral elasticity of substitution, it is constant across all firms in the sector.

## Trade Costs

Thus far, we have not considered the conditions under which international trade takes place. At the moment, the model simply consists of a set of demand functions and pricing conditions for all countries and sectors. As it is, the model describes trade in a frictionless world, in which goods produced in country  $i$  can be shipped to country  $j$  at no charge. Arbitrage ensures, therefore, identical prices in both countries.

To introduce trade frictions, we can use the common “iceberg” formulation. When a firm ships goods from country  $i$  to country  $j$ , it must send  $\tau_{ij}^k \geq 1$  units in order for a single unit to arrive. The difference can be thought of as “melting” (like an iceberg) en route to the destination. Equivalently, the marginal cost of producing in country  $i$  a unit of a good subsequently consumed in the same country  $i$  is  $wa_i^k$ , but if the same product is consumed in country  $j$  then the marginal cost is instead  $\tau_{ij}^k wa_i^k$ . Using this definition, costless trade corresponds to  $\tau_{ij}^k = 1$ , and  $\tau_{ij}^k$  corresponds to one plus the ad valorem tariff rate. Since the size of the trade friction associated with a given iceberg coefficient does not depend on the quantity of goods shipped, we can treat iceberg costs as being variable (not fixed) in nature.

Taking any two countries  $i$  and  $j$ , the presence of iceberg trade costs means that the price in country  $j$  of goods produced in country  $i$  is (from equation (20) above):

$$p_j^k(v) = \left( \frac{\sigma_k}{\sigma_k - 1} \right) \tau_{ij}^k wa_i^k = \tau_{ij}^k p_i^k(v) \quad (21)$$

This result allows us to rewrite the country price index in a more useful (and general) form:

$$P_j^k = \left\{ \int_{v \in V_j^k} [\tau_{ij}^k p_i^k(v)]^{1-\sigma_k} dv \right\}^{\frac{1}{1-\sigma_k}} \quad (22)$$

Note that this index includes varieties that are produced and consumed in the same country: all the  $\tau_{ii}^k$  terms are simply set to unity, so as to reflect the absence of internal trade barriers.

## Model Closure: Gravity with Gravititas

These are all the ingredients required to put together a gravity model with gravitas. The trick is in combining them in the right way.

The gravity model is usually concerned with the value of bilateral trade ( $x_{ij}^k$ ), i.e. exports from country  $i$  to country  $j$  of a particular product variety. Combining the price equation (##) with the demand function (##) gives:

$$x_{ij}^k(v) = p_{ij}^k(v)x_j^k(v) = \tau_{ij}^k p_i^k(v) \left\{ \frac{\tau_{ij}^k p_i^k(v)}{P_j^k} \right\}^{-\sigma_k} \frac{E_j^k}{P_j^k} \equiv \left\{ \frac{\tau_{ij}^k p_i^k(v)}{P_j^k} \right\}^{1-\sigma_k} E_j^k \quad (23)$$

The above expression gives us bilateral exports of a single product variety. To derive something that looks more obviously like a gravity equation, we need to aggregate this expression to give total sectoral exports from  $i$  to  $j$ , i.e.  $X_{ij}^k$ . From the production side of the model, it is clear that all firms in a given country-sector are symmetrical in terms of marginal cost, sales, price, etc. Using the measure  $N_i$  of firms active in country  $i$ , we can therefore write total sectoral exports very simply:

$$X_{ij}^k = N_i \left\{ \frac{\tau_{ij}^k p_i^k(v)}{P_j^k} \right\}^{1-\sigma_k} E_j^k \quad (24)$$

Now comes the important part: introducing a general equilibrium accounting identity. It must be the case that sectoral income in country  $i$ ,  $Y_i^k$ , is the income earned from total worldwide sales of all locally made varieties in that sector. Thus:

$$Y_i^k = \sum_{j=1}^C X_{ij}^k = N_i [p_i^k(v)]^{1-\sigma_k} \sum_{j=1}^C \left\{ \frac{\tau_{ij}^k}{P_j^k} \right\}^{1-\sigma_k} E_j^k \quad (25)$$

Solving for  $N_i [p_i^k(v)]^{1-\sigma_k}$  gives:

$$N_i [p_i^k(v)]^{1-\sigma_k} = \frac{Y_i^k}{\sum_{j=1}^C \left\{ \frac{\tau_{ij}^k}{P_j^k} \right\}^{1-\sigma_k} E_j^k} \quad (26)$$

Next, substitute that expression back into the sectoral exports equation (24):

$$X_{ij}^k = \frac{Y_i^k E_j^k}{\sum_{j=1}^C \left\{ \frac{\tau_{ij}^k}{P_j^k} \right\}^{1-\sigma_k} E_j^k} \left\{ \frac{\tau_{ij}^k}{P_j^k} \right\}^{1-\sigma_k} \quad (27)$$

## 2 The Basis for the Gravity Model: From Intuition to Theory

For convenience, define  $\Pi_i^k = \sum_{j=1}^C \left\{ \frac{\tau_{ij}^k}{P_j^k} \right\}^{1-\sigma_k} \frac{E_j^k}{Y^k}$  where  $Y^k$  is total world output in sector  $k$ . Dividing the above expression through by  $Y^k$  and substituting  $\Pi_i^k$  gives the Anderson and Van Wincoop (2003) gravity model:

$$X_{ij}^k = \frac{Y_i^k E_j^k}{Y^k} \left\{ \frac{\tau_{ij}^k}{\Pi_i^k P_j^k} \right\}^{1-\sigma_k} \quad (28)$$

or in the more common log-linearized form:

$$\log X_{ij}^k = \log Y_i^k + \log E_j^k - \log Y^k + (1-\sigma_k)[\log \tau_{ij}^k - \log \Pi_i^k - \log P_j^k] \quad (29)$$