Gravity-Defying Trade*

J.M.C. Santos Silva[†] Silvana Tenreyro[‡]

March 26, 2003

*The authors are grateful to Francesco Caselli, Juan Carlos Hallak, Daniel Mota, Paulo Parente, Manuela Simarro and Kim Underhill for helpful suggestions and comments on previous versions of this paper. The usual disclaimer applies. Jiaying Huang provided excellent research assistance. Santos Silva gratefully acknowledges the partial financial support from Fundação para a Ciência e a Tecnologia, program POCTI, partially funded by FEDER.

[†]ISEG/Universidade Técnica de Lisboa, R. do Quelhas 6, 1200 Lisboa, Portugal. Fax: +351 213922781. E-mail: jmcss@iseg.utl.pt.

[‡]Federal Reserve Bank of Boston, 600 Atlantic Avenue, Boston, MA 02106-2076, USA. Fax: +1 6179733957. E-mail: Silvana.Tenreyro@bos.frb.org.

Gravity-Defying Trade

Abstract

Although economists have long been aware of Jensen's inequality, many econometric applications have neglected an important implication of it: estimating economic relationships in logarithms can lead to significant biases in the presence of heteroskedasticity. This paper explains why this problem arises and proposes an appropriate estimator. Our criticism to conventional practices and the solution we propose extends to a broad range of economic applications where the equation under study is log-linearized. We develop the argument using one particular illustration, the gravity equation for trade, and use the proposed technique to provide novel estimates of this equation. Three results stand out. First, contrary to general belief, income elasticities are significantly smaller than 1. Second, standard estimators greatly exaggerate the roles of distance and colonial links. Finally, bilateral trade between countries that have signed a free-trade agreement is 30 percent larger than that between other countries, a magnitude remarkably different from that predicted by conventional methods (above 100 percent).

Key words: Gravity equation, Free-trade agreements, Heteroskedasticity, Poisson regression.

JEL Codes: C21, F10, F11, F12, F15.

1. Introduction

Economists have long been aware of Jensen's inequality. In particular, it is well known that $E(\ln y) \neq \ln E(y)$, i.e., the expected value of the logarithm of a random variable is different from the logarithm of its expected value. This basic fact, however, has been neglected in many econometric applications. Indeed, one important implication of Jensen's inequality is that, in the presence of heteroskedasticity, the standard practice of using least squares to estimate economic relationships in logarithms (instead of levels) can lead to significant biases. This paper shows how this problem arises and proposes an appropriate estimator. We develop the argument using one particular illustration: the gravity equation for trade. However, our criticism to the conventional practice and the solution we propose extends to a broad range of economic applications where the equations under study are log-linearized, or, more generally, transformed by a non-linear function. A short list of examples includes the estimation of Mincerian equations for wages, production functions, and Euler equations, which are typically estimated in logarithms.

There is a vast theoretical and empirical literature on the gravity equation for trade, initiated by the pioneering work of Jan Tinbergen (1962). Theories based on different foundations for trade, including endowment and technological differences, increasing returns to scale, and "Armington" demands, all predict a gravity relationship for trade flows analogous to Newton's "Law of Universal Gravitation."[1] In its simplest form, the gravity equation for trade states that exports from country i to country j, denoted by T_{ij}, are proportional to the product of the two countries' GDPs, denoted by Y_i and Y_j, and inversely proportional to their distance, D_{ij}, broadly construed to include all factors that might create trade resistance. That is,

$$T_{ij} = \alpha_0 Y_i^{\alpha_1} Y_j^{\alpha_2} D_{ij}^{\alpha_3}, \tag{1}$$

[1] See, for example, Anderson (1979), Helpman and Krugman (1985), Bergstrand (1985), Davis (1995), Deardoff (1998), and Anderson and van Wincoop (2003). A feature common to these models is that they all assume complete specialization: each good is produced in only one country. However, Haveman and Hummels (2001), Feenstra, Markusen, and Rose (1999), and Eaton and Kortum (2001) derive the gravity equation without relying on complete specialization.

where α_0, α_1, α_2, and α_3 are parameters to be estimated.

The analogy between trade and the physical force of gravity, however, clashes with the observation that there is no set of parameters for which equation (1) will hold exactly. To account for deviations from the theory, stochastic versions of the equation are used in empirical studies. Typically, the stochastic version of the gravity equation has the form

$$T_{ij} = \alpha_0 Y_i^{\alpha_1} Y_j^{\alpha_2} D_{ij}^{\alpha_3} \eta_{ij}, \tag{2}$$

where η_{ij} is an error term with $E(\eta_{ij}|Y_i, Y_j, D_{ij}) = 1$, assumed to be statistically independent of the regressors.

There is a long tradition in the trade literature of log-linearizing (2) and estimating the parameters of interest by ordinary least squares (OLS) using the equation

$$\ln(T_{ij}) = \ln(\alpha_0) + \alpha_1 \ln(Y_i) + \alpha_2 \ln(Y_j) + \alpha_3 \ln(D_{ij}) + \ln(\eta_{ij}). \tag{3}$$

The validity of this procedure depends critically on the assumption that η_{ij}, and therefore $\ln(\eta_{ij})$, are statistically independent of the regressors. It is very unlikely, however, that the variance of η_{ij} will be independent of the countries' GDPs and of the various measures of distance between them. In other words, the error term η_{ij} will, in general, be heteroskedastic. This implies that the standard estimation method will generate inconsistent estimates. To see why this is so, notice that the expected value of the logarithm of a random variable depends both on its mean and on higher-order moments of the distribution. Hence, whenever the variance of the error term η_{ij} in equation (1) depends on Y_i, Y_j, or D_{ij}, the expected value of $\ln(\eta_{ij})$ will also depend on the regressors, violating the condition for consistency of OLS.[2] Clearly, in this instance, homoskedasticity is critical not only for the efficiency of the estimator, but also for its consistency.

A related problem with the analogy between Newtonian gravity and trade is that gravitational force can be very small, but never zero, whereas trade between several pairs of

[2] As an illustration, consider the case in which η_{ij} follows a log-normal distribution, with $E(\eta_{ij}|Y_i, Y_j, D_{ij}) = 1$ and variance $\sigma_{ij}^2 = f(Y_i, Y_j, D_{ij})$. The error term in the log-linearized representation will then follow a normal distribution, with $E\left[\ln(\eta_{ij})|Y_i, Y_j, D_{ij}\right] = -\frac{1}{2}\ln(1+\sigma_{ij}^2)$, which is also a function of the covariates.

countries is literally zero. These observations pose an additional problem to the use of the log-linear form of the model, and the empirical literature developed after Tinbergen has either ignored it or addressed it in unsatisfactory ways.

To address these issues, we suggest the use of a simple pseudo-maximum likelihood estimation technique. Besides estimating the gravity equation in its multiplicative form, this method provides great efficiency gains relatively to the standard non-linear least squares estimator (NLS). Using Monte Carlo simulations, we compare the performance of our estimator with that of OLS (in the log-linear specification) and NLS. The results are striking. In the presence of heteroskedasticity, the estimation methods used in empirical applications of the gravity equation can be severely biased, generating highly distorted estimates of the model. These biases might be critical for the comparative assessment of competing theories underlying the gravity equation, as well as for the use of the gravity equation as a framework to evaluate the effects of different policies on trade.[3] In contrast, our method is robust to different patterns of heteroskedasticity and, in addition, it provides a natural way to deal with zero values of the dependent variable.

We next use the proposed method to provide new estimates of the gravity equation and, in particular, to reassess the impact of free-trade agreements on the volume of international trade. Our estimation method paints a very different picture of the determinants of international trade. The coefficients on GDP are clearly not, as generally believed, close to 1. Instead, they are significantly smaller, calling for modifications to the simple gravity models.[4] Incidentally, the smaller estimated elasticities help reconcile the gravity equation with the observation that the trade-to-GDP ratio decreases with total GDP (or,

[3]Examples of empirical studies framed on the gravity equation include the evaluation of trade protection (e.g., Harrigan, 1993), regional trade agreements (e.g., Frankel, Stein, and Wei, 1995; Frankel, 1997), exchange rate variability (e.g., Frankel and Wei, 1993; Eichengreen and Irwin, 1995), and currency unions (e.g., Rose, 2000; Frankel and Rose, 2002; and Tenreyro and Barro, 2002). See also the various studies on "border-effects" influencing the patterns of intranational and international trade, including McCallum (1995), and Anderson and van Wincoop (2003), among others.

[4]Note that a more complex – and complete – model of gravity, like the one proposed by Anderson and van Wincoop (2003) can rationalize our results, as their model is consistent with smaller income elasticities.

in other words, that smaller countries tend to be more open to international trade). In addition, OLS greatly exaggerates the roles of colonial ties and geographical proximity. Perhaps more interesting, we find that, on average, bilateral trade between countries that have signed a free-trade agreement is 30 percent larger than trade between pairs without agreement, in contrast to the 117 percent predicted by OLS regressions. This striking contrast in estimates suggests that inferences drawn on the standard regressions used in the literature can produce misleading conclusions and confound policy decisions.

The remainder of the paper is organized as follows. Section 2 studies the econometric problems raised by the estimation of gravity equations and introduces the pseudo-maximum likelihood estimator. Section 3 presents the Monte Carlo simulations. Section 4 provides new estimates of the gravity equation, revisiting the role of free-trade agreements in international trade. Section 5 contains concluding remarks.

2. The econometrics of gravity equations

Despite their immense popularity, and even after the recent interest on the appropriate specification of gravity equations (see Mátyás, 1997 and 1998; Egger, 2000 and 2002; and Cheng and Wall, 2002), there are still important econometric flaws in empirical studies involving gravity equations. This section examines how the deterministic gravity equation suggested by economic theory can be used in empirical studies of the determinants of trade.

2.1 The gravity equation in a stochastic context

In their non-stochastic form, the relationship between the multiplicative model and its log-linear additive formulation is trivial. The problem, of course, is that the gravity equation for trade flows does not hold with the accuracy of the physical law. All that can be expected is that it holds on average. Indeed, we interpret the gravity equation as the expected value of trade for a given value of the explanatory variables. That is, denoting the volume of trade by y and the set of covariates by x, the gravity equation is interpreted as the conditional expectation of y_i given x, denoted $E[y_i|x]$ (see Goldberger, 1991, p. 5). Because trade between any pair of countries is necessarily non-negative and the explanatory variables are all positive (measures of distance and size), the multiplicative form of the gravity equation is well suited to describe $E[y_i|x]$, which must be positive. A multiplicative model with positive covariates can always be written in the exponential form, leading to $E[y_i|x] = \exp(x_i\beta)$, which emphasizes that $E[y_i|x]$ is always positive.[5]

Since the gravity law holds on average, but not for each pair of countries, there is an error term associated with each observation, which is defined as $\varepsilon_i = y_i - E[y_i|x]$.[6] Therefore, the stochastic model for trade can be formulated as

$$y_i = \exp(x_i\beta) + \varepsilon_i, \tag{4}$$

with $y_i \geq 0$ and $E[\varepsilon_i|x] = 0$.

As we mentioned in the introduction, the standard practice of log-linearizing equation (4) and estimating β by OLS is inappropriate for a number of reasons. First of all, y_i can be zero, in which case log-linearization is unfeasible. (In spite of this, most authors have opted for dropping the zeroes out of the sample, and estimated β from the truncated sample.) Second, even if all observations of y_i are strictly positive, the expected value of the log-linearized error will in general depend on the covariates and hence OLS will be inconsistent. To see the point more clearly, notice that equation (4) can be expressed as

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

with $\eta_i = 1 + \varepsilon_i / \exp(x_i\beta)$ and $E[\eta_i|x] = 1$. Assuming for the moment that y_i is positive, the model can be made linear in the parameters by taking logarithms of both sides of the equation, leading to

$$\ln(y_i) = x_i\beta + \ln(\eta_i). \tag{5}$$

To obtain a consistent estimator of the slope parameters in equation (4) estimating (5) by OLS, it is necessary that $E[\ln(\eta_i)|\,x]$ does not depend on x_i.[7] This condition is met

[5] Using the notation in the introduction, the multiplicative gravity relationship can be written as the exponential function $E(T_{ij}|Y_i, Y_j, D_{ij}) = \exp[\ln(\alpha_0) + \alpha_1 \ln(Y_i) + \alpha_2 \ln(Y_j) + \alpha_3 \ln(D_{ij})]$.

[6] Whether the error term enters additively or multiplicatively is irrelevant for our purposes, as explained below.

[7] Consistent estimation of the intercept would also require $E[\ln(\eta_i)|\,x] = 0$.

only if ε_i can be written as $\varepsilon_i = \exp(x_i\beta)\,\nu_i$, where ν_i is a random variable statistically independent of x_i. In this case, η_i is also statistically independent of x_i, implying that $E[\ln(\eta_i)|\,x]$ is constant and the conditional variance of y_i (and ε_i) is proportional to $\exp(2x_i\beta)$. Thus, the log-linear representation of the gravity equation is only useful as a device to estimate the parameters of interest under very specific conditions on the error term.

Economic theory provides no information on the variance of ε_i, but we can infer some of its properties from the characteristics of trade data. Because y_i is non-negative and can be zero with positive probability, when $E[y_i|x]$ approaches zero, the probability of y_i being positive must also approach zero. This implies that $V[y_i|x]$, the conditional variance of y_i, tends to vanish as $E[y_i|x]$ passes to zero.[8] On the other hand, when the expected value of trade is far away from the lower bound of the dependent variable, it is possible to observe large deviations from the conditional mean in either direction, leading to greater dispersion. Thus, in practice, ε_i will generally be heteroskedastic but there is no reason to assume that $V[y_i|x]$ is proportional to $\exp(2x_i\beta)$. Therefore, in general, regressing $\ln(y_i)$ on x_i by OLS will lead to inconsistent estimates of β.

It may be surprising that the pattern of heteroskedasticity and, indeed, the form of all higher-order moments of the conditional distribution of the error term can affect the consistency of an estimator, rather than just its efficiency. The reason is that the non-linear transformation of the dependent variable in equation (5) changes the properties of the error term in a non-trivial way since the conditional expectation of $\ln(\eta_i)$ depends on the shape of the conditional distribution of η_i. Hence, unless very strong restrictions on the form of this distribution are imposed, it is not possible to recover information about the conditional expectation of y_i from the conditional mean of $\ln(y_i)$ simply because $\ln(\eta_i)$ is correlated with the regressors. Although this problem has been neglected in the

[8] Heuristically, when $E[y_i|x]$ is close to its lower bound (i.e., for pairs of small and distant countries), it is unlikely that large values of trade are observed since they cannot be offset by equally large deviations in the opposite direction simply because trade cannot be negative. Therefore, for these observations, dispersion around the mean tends to be small.

literature on international trade, and indeed in most econometric studies and textbooks, it has been recognized by a small group of authors (see, for example, Manning and Mullahy, 2001).

In short, even assuming that all observations on y_i are positive, it is not advisable to estimate β from the log-linear model. Instead, the non-linear model has to be estimated.

2.2 Estimation of the non-linear model

Although most empirical studies in international trade use the log-linear form of the gravity equation, some authors (e.g., Frankel and Wei, 1993) have estimated the multiplicative gravity equation using non-linear least squares (NLS), which is an asymptotically valid estimator for (4). However, the NLS estimator can be very inefficient in this context, as it ignores the heteroskedasticity that, as discussed before, is characteristic of this type of model.

The NLS estimator of β is defined by

$$\hat{\beta} = \arg\min_b \sum_{i=1}^{n} \left[y_i - \exp\left(x_i b\right)\right]^2 ,$$

which implies the following set of first order conditions:

$$\sum_{i=1}^{n} \left[y_i - \exp\left(x_i \hat{\beta}\right)\right] x_i \exp\left(x_i \hat{\beta}\right) = 0. \tag{6}$$

These equations give more weight to observations where $\exp\left(x_i \hat{\beta}\right)$ is large because that is where the curvature of the conditional expectation is more pronounced. However, these are generally also the observations with larger variance, which implies that NLS gives more weight to noisier observations. Thus, this estimator may be very inefficient, depending heavily on a small number of observations.

If the form of $V\left[y_i|x\right]$ was known, this problem could be eliminated using a weighted NLS estimator. However, in practice, all we know about $V\left[y_i|x\right]$ is that, in general, it goes to zero as $E\left[y_i|x\right]$ passes to zero. Therefore, an optimal weighted-NLS estimator cannot be used without further information on the distribution of the errors. A possible way of obtaining a more efficient estimator is to follow McCullagh and Nelder (1989) and estimate the parameters of interest using a pseudo-maximum likelihood estimator based on some assumption on the functional form of $V[y_i|x]$.[9]

Among the many possible specifications for the conditional variance, the hypothesis $V[y_i|x] = E[y_i|x]$ is particularly appealing. This hypothesis is characteristic of the Poisson regression model which is often used to describe count data.[10] Like trade data, counts cannot be negative and they have a positive probability of being zero. (Additionally, counts are necessarily integers, but that is not strictly required to use Poisson regression.)

In Poisson regressions, it is customary to specify $E[y_i|x] = V[y_i|x] = \exp(x_i\beta)$, and, under the assumptions of the model, β can be estimated by maximum likelihood. This estimator is defined by

$$\tilde{\beta} = \arg\max_b \sum_{i=1}^{n} \{y_i(x_i b) - \exp(x_i b)\},$$

which is equivalent to solving the following set of first order conditions:

$$\sum_{i=1}^{n} \left[y_i - \exp\left(x_i\tilde{\beta}\right)\right] x_i = 0. \tag{7}$$

The form of (7) makes clear that all that is needed for this estimator to be consistent is for the conditional mean to be correctly specified, i.e., $E[y_i|x] = \exp(x_i\beta)$. Therefore, the data do not have to be Poisson at all and, what is more important, y_i does not even have to be an integer, for the estimator based on the Poisson likelihood function to be consistent. This is the well-known pseudo-maximum likelihood result first noted by Gourieroux, Monfort and Trognon (1984).

Comparing equations (6) and (7), it is clear that, unlike the NLS estimator, the Poisson estimator gives the same weight to all observations, rather than emphasizing those for which $\exp(x_i\beta)$ is large. This is because, under the assumption that $E[y_i|x] = V[y_i|x]$, all observations have the same information on the parameters of interest as the additional

[9]See also Manning and Mullahy (2001). A related estimator is proposed by Papke and Wooldridge (1996) for the estimation of models for fractional data.

[10]See Cameron and Trivedi (1998) and Winkelmann (2000) for more details on the Poisson regression and on more general models for count data.

information on the curvature of the conditional mean coming from observations with large $\exp(x_i\beta)$ is exactly offset by their larger variance. Even if $E[y_i|x]$ is not equal to $V[y_i|x]$, the Poisson estimator is likely to be more efficient than the NLS estimator when the heteroskedasticity increases with the conditional mean.

Of course, if $V[y_i|x]$ is a function of higher powers of $E[y_i|x]$, a more efficient estimator could be obtained down-weighing even more the observations with large conditional mean. However, in the case of trade data, this class of models may have an important drawback. Trade data for larger countries (as gauged by GDP per capita) tend to be of higher quality (see Frankel and Wei, 1993, and Frankel 1997); hence, models assuming that $V[y_i|x]$ is a function of higher powers of $E[y_i|x]$ might give excessive weight to the observations that are more prone to measurement errors.[11] The Poisson regression emerges as a reasonable compromise, giving less weight to the observations with larger variance than the standard NLS estimator, without giving too much weight to observations more seriously contaminated by measurement error.

The implementation of the Poisson pseudo-maximum likelihood estimator is straightforward since there are standard econometric programs with commands that permit the estimation of Poisson regression, even when the dependent variables are not integers. Of course, because the assumption $E[y_i|x] = V[y_i|x]$ is unlikely to hold, this estimator does not fully account for the heteroskedasticity in the model and all inference has to be based on an Eicker-White (Eicker, 1963; and White, 1980) robust covariance matrix estimator.

[11] Frankel and Wei (1993) and Frankel (1997) suggest that larger countries should be given more weight in the estimation of gravity equations. This would be appropriate if the errors in the model were just the result of measurement errors in the dependent variable. However, if it is accepted that the gravity equation does not hold exactly, measurement errors account for only part of the dispersion of trade data around the gravity equation.

2.3 Zero Gravity

A well-known characteristic of trade data is that the value of trade flows between some pairs of countries, or regions within a country, is zero. This can happen at the aggregate level or, very typically, at the sectoral level.

In many cases, these zeros occur simply because some pairs of countries did not trade in a given period. For example, it would not be surprising to find that Tajikistan and Togo did not trade in a certain year.[12] These zero observations pose no problem at all for the estimation of gravity equations in their multiplicative form. In contrast, the existence of observations for which the dependent variable is zero creates an additional problem for the use of the log-linear form of the gravity equation. A number of methods has been developed to deal with this problem (see Frankel, 1997, for a description of the various procedures). The approach followed by the large majority of empirical studies is simply to drop the pairs with zero trade from the data set and estimate the log-linear form by OLS. This truncation, however, makes the OLS estimator of β inconsistent. The severity of the problem will depend on the particular characteristics of the sample and model used, but there is no reason to believe that it will always be negligible.

Naturally, there are other reasons for observing pairs of countries with zero trade. For example, zeroes may be the result of rounding errors. If trade is measured in thousands of dollars, it is possible that for pairs of countries for which bilateral trade did not reach a minimum value, say $500, the value of trade is registered as zero. If these rounded-down observations were partially compensated by rounded-up ones, the overall effect of these errors would be relatively minor. However, because there is a large number of pairs of countries for which the value of bilateral trade is expected to be very small, it is likely that the rounding down will not be totally offset. Moreover, the rounding down is more likely to occur for small or distant countries and, therefore, the probability of rounding down

[12] The absence of trade between small and distant countries might be explained, among other factors, by large variable costs (e.g., bricks are too costly to transport) or large fixed costs (e.g., information on foreign markets). At the aggregate level, these costs can be best proxied by the various measures of distance and size entering the gravity equation.

will depend on the value of the covariates, leading to the inconsistency of the estimators. Finally, the zeros can just be missing observations which are wrongly recorded as zero. This problem is more likely to occur when small countries are considered and, again, measurement error will depend on the covariates, leading to inconsistency.

Besides the problems mentioned above, trade data can suffer from many other forms of errors, as described in Feenstra, Lipsey, and Bowen (1997). Of course, in any empirical study, the quality of the results will largely depend on the richness of the data. Therefore, these measurement problems are not specific to the models of trade and they do not play an important role in the choice of the form of the gravity equation. However, as discussed before, they were important in the choice of the particular pseudo-maximum likelihood estimator adopted here. The Poisson pseudo-maximum likelihood estimator will also be affected by the presence of errors in the data, but the consequences of this problem, as we will show in the next section, are likely to be less severe than for estimators based on log-linear versions of the gravity equation. In Section 3, we investigate the robustness of various estimators in the presence of heteroskedasticity and rounding errors.

3. A simulation study

This section reports the results of a small simulation study designed to assess the performance of different methods to estimate the gravity equation in the presence of heteroskedasticity and rounding errors. These experiments are centered around the following multiplicative model:

$$E\left[y_i|x\right] = \mu\left(x_i\beta\right) = \exp\left(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i}\right), \quad i = 1, \ldots, 1000. \tag{8}$$

Since, in practice, gravity models often include a mixture of continuous and dummy variables, we replicate this feature in our experiments: x_{1i} is drawn from a standard normal and x_2 is a binary dummy variable that equals 1 with a probability of 0.4.[13] The two covariates are independent and a new set of observations of all variables is generated

[13] Examples of continuous variables (which are all strictly positive) include income and geographical distance. In equation (8), x_1 can be interpreted as (the logarithm of) one of these variables. Examples of

in each replication using $\beta_0 = 0$, $\beta_1 = \beta_2 = 1$. Data on y are generated as

$$y_i = \mu(x_i\beta)\,\eta_i, \tag{9}$$

where η_i is a log-normal random variable with mean 1 and variance σ_i^2. As noted before, the slope parameters in (8) can be estimated using the log-linear form of the model only when σ_i^2 is constant. That is, when $V[y_i|x]$ is proportional to $\mu(x_i\beta)^2$.

We studied the performance of the following estimators:

PML: The Poisson pseudo-maximum likelihood estimator;

NLS: Standard non-linear least squares estimator;

OLS: Ordinary least squares estimator of the log-linear model.

To assess the performance of the estimators under different patterns of heteroskedasticity, we considered the four following specifications of σ_i^2:

Case 1: $\sigma_i^2 = 1$; $V[y_i|x] = \mu(x_i\beta)^2$;

Case 2: $\sigma_i^2 = \mu(x_i\beta)^{-1}$; $V[y_i|x] = \mu(x_i\beta)$;

Case 3: $\sigma_i^2 = \exp(x_{2i}) + \mu(x_i\beta)^{-1}$; $V[y_i|x] = \mu(x_i\beta) + \exp(x_{2i})\,\mu(x_i\beta)^2$;

Case 4: $\sigma_i^2 = \mu(x_i\beta)$; $V[y_i|x] = \mu(x_i\beta)^3$.

Case 1 is the special case in which OLS estimation of the log-linear model is consistent for the slope parameters of (8). Moreover, in this case the log-linear model corrects the heteroskedasticity in the data, and, because η_i is log-normal, it coincides with the maximum likelihood estimator. In Case 2, the conditional variance of y_i equals its conditional mean, as in the Poisson distribution. The pseudo-likelihood approach based on the Poisson distribution is optimal in this situation. Case 3 is the only one in which the conditional variance does not depend exclusively on the mean. The variance is a quadratic

binary variables include dummies for free-trade agreements, common language, colonial ties, contiguity and access to land.

function of the mean, as in Case 1, but it is not proportional to the square of the mean. Finally, in Case 4 the variance is a cubic function of the mean, generating highly dispersed data for $\mu(x_i\beta) > 1$, but very little dispersion for small values of $\mu(x_i\beta)$. Notice that the standard NLS estimator is not optimal in any of the cases considered. For this estimator to be optimal, the variance of ε_i would have to be constant, which, as argued before, is unlikely in models of bilateral trade.

We carried out two sets of experiments. The first set was aimed at studying the performance of the estimators of the multiplicative and the log-linear models under different patterns of heteroskedasticity. The second set studied the estimators' performance in the presence of rounding errors in the dependent variable. For that purpose, a new random variable was generated rounding to the nearest integer the values y_i obtained as before. This procedure mimics the rounding errors in official statistics and generates a large number of zeros, a typical feature of trade data. Because the model considered here generates a large proportion of observations close to zero, rounding down is much more frequent than rounding up. As the probability of rounding up or down depends on the covariates, this procedure will necessarily bias the estimates, as discussed before. The purpose of the study is to gauge the magnitude of these biases. Naturally, the log-linear model cannot be estimated in these conditions because the dependent variable equals zero for some observations. Following what is the usual practice in these circumstances, estimation using the log-linear model was performed dropping the observations for which the dependent variable equals zero.

The biases of the three estimators of β were estimated with $10,000$ replicas of the simulation procedure described above. The results are displayed in Table 1.[14] The first three columns report the biases in the absence of rounding errors; the last three report the biases when rounding errors are present.

[14] Results for the variance of the different estimators were also obtained and these results are available from the authors on request.

Table 1: Estimated Bias under Different Patterns of Heteroskedasticity.

	Without rounding error			With rounding error		
Estimator:	PML	NLS	OLS	PML	NLS	OLS
Case 1:	$V[y_i \vert x] = \mu(x_i\beta)^2$					
β_0	−0.00314	−0.92062	−0.50046	−0.06697	−0.93359	0.30732
β_1	−0.00493	0.27306	0.00016	0.02379	0.27701	−0.41019
β_2	−0.00153	0.11522	0.00033	0.02900	0.12065	−0.41152
Case 2:	$V[y_i \vert x] = \mu(x_i\beta)$					
β_0	−0.00106	−0.00338	−0.40323	−0.04898	−0.01031	0.25750
β_1	0.00010	0.00057	0.21095	0.02216	0.00272	−0.24408
β_2	0.00050	0.00146	0.19964	0.02394	0.00446	−0.23868
Case 3:	$V[y_i \vert x] = \mu(x_i\beta) + \exp(x_{2i})\,\mu(x_i\beta)^2$					
β_0	−0.00198	−1.10371	−0.60117	−0.06096	−1.13154	0.35092
β_1	−0.00620	0.32746	0.13279	0.02112	0.33734	−0.45187
β_2	−0.00602	0.13104	−0.12524	0.01904	0.13484	−0.46131
Case 4:	$V[y_i \vert x] = \mu(x_i\beta)^3$					
β_0	0.00356	−5.43860	−0.40330	−0.06881	−5.81393	0.28541
β_1	−0.03638	1.66370	−0.28932	−0.00340	1.77359	−0.49566
β_2	−0.02024	0.40587	−0.29992	0.01407	0.45870	−0.49279

As expected, the estimator based on the log-linear model only performs well in Case 1 when no rounding error is present.[15] In all other cases, this estimator is clearly inadequate, being often badly biased. Moreover, the sign and magnitude of the bias vary considerably. Thus, these results clearly indicate that estimation based on the log-linear model cannot be recommended, except under very special circumstances. Otherwise, the estimates obtained provide very little information on the parameters of interest.

[15] Notice that only the results for β_1 and β_2 are of interest in this case since it is well know that the estimator of β_0 is inconsistent.

One remarkable result of this set of experiments is the extremely poor performance of the standard NLS estimator. Indeed, when the heteroskedasticity is more severe (cases 1, 3, and, especially, 4) this estimator leads to very poor results because of its erratic behavior.[16] Therefore, it is clear that the loss of efficiency caused by some of the forms of heteroskedasticity considered in these experiments is strong enough to render this estimator useless in practice.

As for the performance of the Poisson pseudo-maximum likelihood estimator, the results are very encouraging. In fact, when no rounding error is present, its performance is reasonably good in all cases, and it does not deteriorate markedly as one moves away from Case 2, in which it is an optimal estimator.[17] Of course, under extreme heteroskedasticity (Case 4) its performance is less satisfactory, but even in this instance, it clearly outperforms its competitors. Moreover, the results obtained with rounded data suggest that the Poisson based pseudo-maximum likelihood estimator is relatively robust to this form of measurement error of the dependent variable. Indeed, the bias introduced by the rounding-off errors in the dependent variable is relatively small and, in some cases, it even compensates the bias found in the first set of experiments.

Obviously, the sign and magnitude of the bias of the estimators studied here depend on the particular specification considered. Therefore, the results of these experiments cannot serve as an indicator of what can be expected in other situations. However, it is clear that, apart from the pseudo-maximum likelihood method, all estimators are potentially very misleading.[18]

[16] Manning and Mullahy (2001) report similar results.

[17] These results are in line with those reported by Manning and Mullahy (2001).

[18] We also studied the performance of Tobit models (with constant and estimated cut-off points as well as the semi-logarithmic Tobit), and OLS models with alternative transformations of the dependent variable, finding very poor results. The outcome of these experiments is available at request.

4. The gravity equation and the role of free-trade agreements

In this section, we use the Poisson method to quantitatively assess the determinants of bilateral trade flows, uncovering significant differences in the roles of various measures of size and distance from those predicted by the "logarithmic tradition." We focus particular attention on the role of free-trade agreements, since this policy instrument has been the object of intense debate (see, for example, Frankel, 1997, and Bhagwati and Panagariya, 1996).

4.1 The data

The analysis covers a cross section of 137 countries in 1990. Hence, our data set consists of $18,632$ observations of bilateral export flows (137×136 country pairs). The list of countries is reported in Table A1 in the Appendix. Information on bilateral exports comes from Feenstra et al. (1997). Data on real GDP per capita and population come from the World Bank's *World Development Indicators* (2002). Data on location and dummies indicating contiguity, common language, colonial ties, and access to water are constructed from the CIA's *World Factbook*. Bilateral distance is computed using the great circle distance algorithm provided by Andrew Gray (2001). Remoteness – or relative distance – is calculated as the (log of) GDP-weighted average distance to all other countries (see Wei, 1996). Finally, information on free-trade agreements come from Frankel (1997), complemented with data from the World Trade Organization.[19] Table A2 in the Appendix provides a description of the variables and displays the summary statistics.

[19]The free-trade agreements existing in 1990, the year of our cross section, include EEC, US-Canada, EFTA, US-Israel, CACM, CARICOM, and CER. We also include SPARTECA, as this preferential-trade agreement went very far in the process of trade liberalization. There are some custom unions – stronger forms of free-trade agreement with common external barriers – for which data on trade are consolidated (e.g., the countries in the Southern African Custom Union and Belgium-Luxembourg) and, hence, the effect of these arrangements cannot be captured by our free-trade dummy.

4.2 Results

Table 2 presents the benchmark estimation outcomes resulting from OLS and Poisson regressions. The first column reports OLS estimates using the logarithm of trade as the dependent variable; as noted before, this regression leaves out pairs of countries with zero bilateral exports (only $9,637$ country pairs, or 52 percent of the sample, exhibit positive export flows). For comparison, the second column reports Poisson estimates using only the subsample of positive-trade pairs. Finally, the third column shows the Poisson results for the whole sample (including zero-trade pairs).

The first point to notice is that Poisson-estimated coefficients are remarkably similar using both the whole sample and the positive-trade subsample.[20] However, most coefficients differ – oftentimes significantly – using OLS. This suggests that in this case, heteroskedasticity (rather than truncation) can distort results in a material way. Poisson estimates reveal that the coefficients on importer's and exporter's GDPs are not, as generally believed, close to 1. The estimated GDP elasticities are just above 0.7 (s.e.= 0.03). OLS generates significantly larger estimates, especially on exporter's GDP (0.94, s.e.= 0.01). These findings suggest that the simpler models of gravity equation (those that predict unit-income elasticities typically as a result of specialization in production and homothetic preferences) should be modified to feature a less-than-proportional relationship between trade and GDP.[21] (See Anderson and van Wincoop (2003), who provide a model consistent with smaller elasticities). It is worth pointing out that unit-income elasticities in the simple gravity framework are at odds with the observation that the trade-to-GDP ratio decreases with total GDP, or, in other words, that smaller countries tend to be more open to international trade.[22]

[20] The reason why truncation has little effect in this case is that observations with zero trade correspond to pairs for which the estimated value of trade is close to zero. Therefore, the corresponding residuals are also close to zero and their elimination from the sample has little effect.

[21] This result holds when one looks at the subsample of OECD countries. It is also robust to the exclusion of GDP per capita from the regressions.

[22] Note also that Poisson predicts almost equal coefficients for the GDPs of exporters and importers.

Table 2. The Gravity Equation. OLS and Poisson Estimations

	OLS	Poisson Trade > 0	Poisson
Log of exporter's GDP	0.944**	0.721**	0.732**
	(0.012)	(0.027)	(0.027)
Log of importer's GDP	0.806**	0.731**	0.741**
	(0.011)	(0.027)	(0.027)
Log of exporter's per capita GDP	0.183**	0.142*	0.147**
	(0.016)	(0.055)	(0.056)
Log of importer's per capita GDP	0.081**	0.121**	0.126**
	(0.017)	(0.043)	(0.044)
Log of distance	−1.165**	−0.785**	−0.788**
	(0.033)	(0.053)	(0.053)
Contiguity dummy	0.297*	0.158	0.153
	(0.128)	(0.096)	(0.096)
Common-language dummy	0.678**	0.778**	0.775**
	(0.067)	(0.128)	(0.129)
Colonial-tie dummy	0.370**	−0.046	−0.036
	(0.070)	(0.151)	(0.151)
Landlocked-exporter dummy	−0.061	−0.754**	−0.748**
	(0.062)	(0.136)	(0.136)
Landlocked-importer dummy	−0.660**	−0.585**	−0.581**
	(0.060)	(0.131)	(0.131)
Exporter's remoteness	0.478**	0.722**	0.719**
	(0.077)	(0.120)	(0.121)
Importer's remoteness	−0.193*	0.625**	0.621**
	(0.083)	(0.117)	(0.118)
Free-trade agreement dummy	0.776**	0.264**	0.262**
	(0.128)	(0.081)	(0.080)
Constant	−28.787**	−32.668**	−33.245**
	(1.054)	(1.881)	(1.837)
Observations	9637	9637	18632
RESET test, p-values	0.000	0.909	0.286

Note: In the OLS regression, the dependent variable is ln(trade). In the Poisson estimation, the dependent variable is trade (the gravity equation is estimated in its multiplicative form). Results for the restricted sample (with positive trade) and the whole sample are reported. The equations use data for 1990. Rejection of the RESET tests indicates that the model is misspecified (see text). Robust standard errors in parentheses. * significant at 5%; ** significant at 1%

The role of geographical distance as trade deterrent is significantly larger under OLS; the estimated elasticity is -1.17 (s.e. $= 0.03$), whereas the Poisson estimate is -0.79 (s.e. $= 0.05$). Our lower estimate suggests a smaller role for transport costs in the determination of trade patterns. Furthermore, Poisson estimates indicate that, after controlling for bilateral distance, sharing a border does not influence trade flows, while OLS, instead, generates a substantial effect: It predicts that trade between two contiguous countries is 35 percent larger than trade between countries that do not share a border.[23]

We control for remoteness to account for the hypothesis that larger distances to all other countries might increase bilateral trade between two countries.[24] Poisson regressions support this hypothesis, whereas OLS estimates suggest that only exporter's remoteness increases bilateral flows. Access to water appears to be important for trade flows, according to Poisson regressions; the negative coefficients on the land-locked dummies can be interpreted as an indication that ocean transportation is significantly cheaper. In contrast, OLS results suggest that whether or not the exporter is landlocked does not influence trade flows, whereas a landlocked importer experiences lower trade; this asymmetry is hard to interpret. We also explore the role of colonial heritage, obtaining, as before, significant discrepancies: Poisson indicates that colonial ties play no role in determining trade flows, once a dummy variable for common language is introduced. OLS regressions, instead, generate a sizeable effect (countries with a common colonial past trade almost 45 percent more than other pairs). Language is statistically and economically significant under both estimation procedures.

Strikingly, free trade agreements play a much smaller – although still substantial – role according to Poisson regressions. OLS estimates suggest that free-trade agreements rises expected bilateral trade by 117 percent, whereas Poisson estimates indicate an average enhancement effect of 30 percent. The contrast in estimates suggests that the biases

[23] The formula to compute this effect is $(e^{b_i} - 1) \times 100\%$, where b_i is the estimated coefficient.

[24] To illustrate the role of remoteness, consider two pairs of countries, (i, j) and (k, l), and assume that the distance between the countries in each pair is the same $D_{ij} = D_{kl}$, however, i and j are closer to other countries. In this case, the most remote countries, k and l, will tend to trade more between each other because they do not have alternative trading partners. See Deardoff (1998).

generated by standard regressions can be substantial, leading to misleading inferences and, perhaps, erroneous policy decisions.[25]

Free-trade agreements might also cause trade diversion; if this is the case, the coefficient on the free-trade dummy will not reflect net effect of free-trade agreements. To account for the possibility of diversion, we include an additional dummy, "openness," similar to that used by Frankel (1997). This dummy takes the value 1 whenever one (or both) of the countries in the pair is part of a free-trade agreement and, thus, it captures the extent of trade between members and non-members of a free-trade agreement. The sum of the coefficients on the free-trade agreement and the openness dummies give the net effect of free-trade agreements. We report the new results in Table 3. Consistently, both OLS and Poisson regressions provide no significant evidence of trade diversion. However, the point estimates are larger under the Poisson method. Hence, on average, the Poisson method estimates a smaller, yet significant, effect of free-trade agreements.

To check the adequacy of the estimated models, we performed a heteroskedasticity-robust RESET test (Ramsey, 1969). This is essentially a test for the correct specification of the conditional expectation, which is performed by checking the significance of an additional regressor constructed as $(x'b)^2$, where b denotes the vector of estimated parameters. The corresponding p-values are reported at the bottom of the tables. In all OLS regressions, the test rejects the hypothesis that the coefficient on the test variable is zero. This means that the models estimated using the logarithmic specification are inappropriate. In contrast, the models estimated using the Poisson regressions pass the RESET test, i.e., the RESET test provides no evidence of misspecification of the gravity equations estimated using the Poisson method.

[25] It is interesting to remark that there is a pattern in the direction of the bias generated by OLS. The bias tends to be positive for the coefficients on variables that relate to larger volumes of trade and, presumably, to larger variance. It tends to be negative for variables that deter trade and, possibly, reduce the variance.

Table 3. The Gravity Equation.
OLS and Poisson Estimations Accounting for Trade Diversion.

	OLS	Poisson Trade > 0	Poisson
Log of exporter's GDP	0.944**	0.720**	0.732**
	(0.012)	(0.027)	(0.027)
Log of importer's GDP	0.806**	0.731**	0.740**
	(0.011)	(0.028)	(0.027)
Log of exporter's per capita GDP	0.188**	0.153**	0.155**
	(0.017)	(0.053)	(0.054)
Log of importer's per capita GDP	0.087**	0.131**	0.133**
	(0.018)	(0.045)	(0.045)
Log of distance	−1.162**	−0.762**	−0.772**
	(0.034)	(0.052)	(0.052)
Contiguity dummy	0.295*	0.175	0.166
	(0.128)	(0.100)	(0.100)
Common-language dummy	0.680**	0.786**	0.781**
	(0.067)	(0.131)	(0.132)
Colonial-tie dummy	0.371**	−0.031	−0.025
	(0.070)	(0.148)	(0.148)
Landlocked-exporter dummy	−0.060	−0.752**	-0.747**
	(0.062)	(0.136)	(0.136)
Landlocked-importer dummy	−0.659**	−0.583**	-0.579**
	(0.060)	(0.131)	(0.130)
Exporter's remoteness	0.469**	0.667**	0.679**
	(0.078)	(0.131)	(0.130)
Importer's remoteness	−0.201*	0.570**	0.582**
	(0.084)	(0.117)	(0.116)
Free-trade agreement dummy	0.787**	0.277**	0.272**
	(0.129)	(0.080)	(0.078)
Openness dummy	−0.045	−0.133	−0.097
	(0.051)	(0.129)	(0.127)
Constant	−28.683**	−31.937**	−32.728**
	(1.063)	(2.080)	(1.982)
Observations	9637	9637	18632
RESET test, p-values	0.000	0.815	0.262

Note: In the OLS regression, the dependent variable is ln(trade). In the Poisson estimation, the dependent variable is trade (the gravity equation is estimated in its multiplicative form). Results for the restricted sample (with positive trade) and the whole sample are reported. The equations use data for 1990. Rejection of the RESET tests indicates that the model is misspecified (see text). Robust standard errors in parentheses. * significant at 5%; ** significant at 1%

5. Conclusions

In this paper, we argue that the standard empirical methods used to estimate gravity equations are inappropriate. The basic problem is that log-linearization (or, indeed, any non-linear transformation) of the empirical model in the presence of heteroskedasticity leads to inconsistent estimates. This is because the expected value of the logarithm of a random variable depends on higher-order moments of its distribution. Therefore, if the errors are heteroskedastic, the transformed errors will be generally correlated with the covariates. An additional problem of log-linearization is that it is incompatible with the existence of zeroes in trade data, which led to several unsatisfactory solutions, including truncation of the sample (i.e., elimination of zero-trade pairs) and further non-linear transformations of the dependent variable.

To address the various estimation problems, we propose a simple pseudo-maximum likelihood method and assess its performance using Monte Carlo simulations. We find that the standard methods, in the presence of heteroskedasticity, can severely bias the estimated coefficients, casting doubt on previous empirical findings. Our method, instead, is robust to different patterns of heteroskedasticity and, in addition, provides a natural way to deal with zeroes in trade data.

We use our method to re-estimate the gravity equation and document significant differences from the results obtained using the log-linear method. Among other differences, income elasticities are systematically smaller than those obtained with log-linearized OLS regressions. In addition, OLS estimation exaggerates the role of geographical proximity and colonial ties. Finally, and perhaps more interesting, bilateral trade between countries that have signed a free-trade agreements is, on average, 30 percent larger than that between pairs of countries without agreement, in contrast to the 117 percent predicted by OLS regressions. Our results suggest that heteroskedasticity (rather than truncation of the data) is responsible for the main differences.

Log-linearized equations estimated by OLS are of course used in many other areas of empirical economics and econometrics. Our Monte Carlo simulations and the regression outcomes indicate that in the presence of heteroskedasticity this practice can lead to significant biases. These results suggest that, at least when there is evidence of heteroskedasticity, the pseudo maximum likelihood estimator should be used as a substitute for the standard log-linear model.

References

Anderson, J. (1979). "A Theoretical Foundation for the Gravity Equation," *American Economic Review*, 69, 106-116.

Anderson, J. and E. van Wincoop (2003). "Gravity with Gravitas: A Solution to the Border Puzzle," *American Economic Review*, forthcoming.

Bhagwati, J. and A. Panagariya, eds. (1996). *Free Trade Areas or Free Trade? The Economics of Preferential Trading.* Washington, DC: The AEI Press.

Bergstrand, J. (1985). "The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence," *Review of Economics and Statistics*, 69, 474-481.

Cameron, A. C. and P. K. Trivedi (1998). *Regression Analysis of Count Data*, Cambridge, MA: Cambridge University Press.

Cheng, I-H. and H.J. Wall (2002). "Controlling for Heterogeneity in Gravity Models of Trade Integration," Federal Reserve Bank of St. Louis, Working Paper No. 1999-010C.

Davis, D. (1995). "Intra-industry Trade: A Hecksher-Ohlin-Ricardo Approach," *Journal of International Economics*, 39, 201-226.

Deardoff, A. (1998). "Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical World?" in Jeffrey Frankel (ed.), *The Regionalization of the World Economy.* Chicago, IL: University of Chicago Press.

Eaton, J. and S. Kortum (2001). "Technology, Geography and Trade," NBER Working Paper No. 6253.

Eicker, F. (1963). "Asymptotic Normality and Consistency of the Least Squares Estimators for Families of Linear Regressions," *The Annals of Mathematical Statistics*, 34, 447-456.

Egger, P. (2000). "A Note on the Proper Econometric Specification of the Gravity Equation," *Economic Letters*, 66, 25-31.

Egger, P. (2002). "An Econometric View on the Estimation of Gravity Models and the Calculation of Trade Potentials," *World Economy*, 25, 297-312.

Feenstra, R. C., R. E. Lipsey and H. P. Bowen (1997), "World Trade Flows, 1970-1992, with Production and Tariff Data," NBER Working Paper No. 5910.

Feenstra, R., J. Markusen and A. Rose (1999). "Using the Gravity Equation to Differentiate Among Alternative Theories of Trade," *http://www.econ.ucdavis.edu/~feenstra*.

Frankel, J. (1997), *Regional Trading Blocs in the World Economic System*. Washington, DC: Institute for International Economics.

Frankel, J. and A. Rose (2002). "An Estimate of the Effect of Common Currencies on Trade and Income," *Quarterly Journal of Economics*, 117, 409-466.

Frankel, J. and S. Wei (1993). "Trade Blocs and Currency Blocs," NBER Working Paper No. 4335.

Gourieroux, C., A. Monfort and A. Trognon (1984). "Pseudo Maximum Likelihood Methods: Applications to Poisson Models," *Econometrica*, 52, 701-720.

Goldberger, A. (1991). *A Course in Econometrics*. Cambridge, MA: Harvard University Press.

Gray, A. (2001). http://argray.fateback.com/dist/formula.html.

Harrigan, J. (1993). "OECD Imports and Trade Barriers in 1983," *Journal of International Economics,* 35, 95-111.

Haveman, J. and D. Hummels (2001). "Alternative Hypotheses and the Volume of Trade: The Gravity Equation and the Extent of Specialization," mimeo Purdue University.

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

McCullagh, P. and J. A. Nelder (1989). *Generalized Linear Models*, 2nd ed., London: Chapman and Hall.

Manning, W.G. and J. Mullahy (2001). "Estimating Log Models: To Transform or Not to Transform?," *Journal of Health Economics* 20, 461-494.

Mátyás, L. (1997). "Proper Econometric Specification of the Gravity Model," *World Economy*, 20, 363-368.

Mátyás, L. (1998). "The Gravity Model: Some Econometric Considerations," *World Economy*, 21, 397-401.

McCallum, J. (1995). "National Borders Matter: Canada-US Regional Trade Patterns," *American Economic Review*, 85, 615-623.

Papke, L.E. and J.M. Wooldridge (1996). "Econometric Methods for Fractional Response Variables with an Application to 401(k) Plan Participation Rates," *Journal of Applied Econometrics*, 11, 619-632

Ramsey, J.B. (1969). "Tests for Specification Errors in Classical Linear Least Squares Regression Analysis," *Journal of the Royal Statistical Society B*, 31, 350-371.

Rose, A. (2000). "One Money One Market: Estimating the Effect of Common Currencies on Trade," *Economic Policy*, 15, 7-46.

Tenreyro, S. and R. Barro (2002). "Economic Effects of Currency Unions," FRB Boston Series, Working Paper No. 02-4.

Tinbergen, J. (1962). *The World Economy. Suggestions for an International Economic Policy.* New York, NY: Twentieth Century Fund.

Wei, S. (1996), "Intra-National versus International Trade: How Stubborn Are Nation States in Globalization?," NBER Working Paper No. 5331.

White, H., (1980). "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica*, 48, 817-838.

Winkelmann, R. (2000). *Econometric Analysis of Count Data*, 3rd ed. Berlin: Springer-Verlag.

World Development Indicators CD-rom (2002). The World Bank.

World Factbook (2002). http://www.cia.gov/cia/publications/factbook/

Appendix

Table A1: List of countries.

Albania	Djibouti	Korea Rp.	Saudi Arabia
Algeria	Dominican Rp.	Laos P. Dem. Rp.	Senegal
Angola	Ecuador	Lebanon	Seychelles
Argentina	Egypt	Madagascar	Sierra Leone
Australia	El Salvador	Malawi	Singapore
Austria	Eq. Guinea	Malaysia	Solomon Islands
Bahamas	Ethiopia	Maldives	South Africa
Bahrain	Fiji	Mali	Spain
Bangladesh	Finland	Malta	Sri Lanka
Barbados	France	Mauritania	St. Kitts and Nevis
Belgium-Luxemburg	Gabon	Mauritius	St. Helena
Belize	Gambia	Mexico	Sudan
Benin	Germany	Mongolia	Suriname
Bhutan	Ghana	Morocco	Sweden
Bolivia	Greece	Mozambique	Switzerland
Brazil	Guatemala	Nepal	Syrian Arab Rp.
Brunei	Guinea	Netherlands	Tanzania
Bulgaria	Guinea-Bissau	New Caledonia	Thailand
Burkina Faso	Guyana	New Zealand	Togo
Burundi	Haiti	Nicaragua	Trinidad and Tobago
Cambodia	Honduras	Niger	Tunisia
Cameroon	Hong Kong	Nigeria	Turkey
Canada	Hungary	Norway	Uganda
Central African Rp.	Iceland	Oman	United Arab Em.
Chad	India	Pakistan	U.K.
Chile	Indonesia	Panama	U.S.A.
China	Iran	Papua New Guinea	Uruguay
Colombia	Ireland	Paraguay	Venezuela
Comoros	Israel	Peru	Vietnam
Congo Dem. Rp.	Italy	Philippines	Yemen
Congo Rp.	Jamaica	Poland	Zambia
Costa Rica	Japan	Portugal	Zimbabwe
Cote D'Ivoire	Jordan	Romania	
Cyprus	Kenya	Russian Federation	
Denmark	Kiribati	Rwanda	

Table A2. Summary Statistics.

Variable	Mean	Standard Deviation
Trade	169620.7	1815440
Log of trade	8.42509	3.27013
Log of GDP	23.22588	2.40468
Log of per capita GDP	7.51025	1.63485
Log of distance	8.78662	0.74276
Contiguity dummy	0.01932	0.13766
Common-language dummy	0.21168	0.40851
Colonial-tie dummy	0.17293	0.37820
Landlocked dummy	0.15328	0.36027
Remoteness	8.94763	0.26323
Free-trade agreement dummy	0.01524	0.12252
Openness dummy	0.51052	0.49990

Note: $N = 18,632$ for the whole sample and $9,632$ for the subsample of positive trade.