GREQAM

Groupement de Recherche en Economie
Quantitative d'Aix-Marseille - UMR-CNRS 6579
Ecole des Hautes Etudes en Sciences Sociales
Universités d'Aix-Marseille II et III

Document de Travail
n°2009-12

GRAVITY, LOG OF GRAVITY AND THE "DISTANCE PUZZLE"

Clément BOSQUET
Hervé BOULHOL

May 2009

DT-GREQAM

Gravity, log of gravity and the "distance puzzle"*

Clément Bosquet† and Hervé Boulhol‡

this draft: may, 2009

Abstract

Estimations of gravity equations specified in logarithm generally conclude that the distance elasticity of trade has increased over time despite globalization. In contrast, building on Santos Silva and Tenreyro (2006), this elasticity is estimated to have been stable around 0.65-0.70 since the 1960s. Moreover, although FTAs tend to cover neighboring countries, this main result is robust to different treatments of FTA effects. The main estimated change refers to the impact of colonial linkages, which has been at least halved. This paper brings also several important methodological contributions to the analysis of gravity equations, including broad support for the Poisson PML estimator.

JEL Codes: F10, F15, C13, C21, C23

Keywords: distance puzzle, gravity equations, international trade, pseudo-maximum likelihood methods

*We are particulary grateful to Thierry Mayer who provided data as well as useful suggestions at an early stage. We also would like to greatly thank Joao Santos Silva, Pierre-Philippe Combes and Lionel Fontagné for useful comments.

†GREQAM, Université Aix-Marseille

‡OECD and Université Paris Panthéon-Sorbonne

1 Introduction

Despite globalization, the role of distance in shaping world trade across trading partners does not seem to have diminished over time. Indeed, according to the meta-analysis carried out by Disdier and Head (2008), trade decreases with distance by at least the same amount today than thirty years ago, with an increase in the (absolute value of) the trade elasticity to distance since the late eighties, a stylised fact framed as the "distance puzzle" by Buch, Kleinert and Toubal (2004) or the "missing globalization puzzle" by Coe, Subramanian and Tamirisa (2007). The latter argue that taking into account zero trade flows using a nonlinear estimator of gravity equation in levels, rather than the common practice of using a linear estimator based on the logarithm of flows, enables to resolve the puzzle: they find that the (absolute value of the) elasticity of trade to distance significantly decreased from roughly 0.5 in 1975 to 0.3 in 2000. In order to simplify the text, the current paper always discusses the evolution of this elasticity in absolute values since there is no ambiguity about its sign.

It is not clear in the first place why a stable elasticity would represent a puzzle. The rather vague presumption seems to be that the expansion of world trade associated with a fall in distance-related trade costs means that distance is having a lesser impact on the structure of trade. Noting that the elasticity of trade to distance is the product of the elasticity of trade to trade costs and of the elasticity of trade costs to distance, there are several reasons to be sceptical about this presumption.

First, the idea that a non-decreasing elasticity represents a puzzle seems somehow related to the "world-is-getting-flatter" hypothesis that Leamer (2007) questions, pointing out that trade remains mostly a neighbourhood phenomenon, as long-distance flows seem to have increased less than short-distance ones. Second, in a careful formalisation of gravity equation, Anderson and van Wincoop (2003) show that trade is actually a homogenous function of degree zero in trade costs due to the multilateral resistance terms. Therefore, even though a general decrease in tariffs spurs international relative to domestic trade, a uniform decrease in transport costs might not lead to increased trade. Third, an overall decrease in transport costs does not necessarily imply a lower distance elasticity of trade. For example, if trade costs, τ_{ijt}, between countries i and j in year t is a function of the bilateral distance d_{ij} such as $a_t\, d_{ij}^{\gamma_t}$, distance can become irrelevant over time through either a decrease in a_t or in γ_t, but a uniform decrease in distance-related transport costs would be associated with a fall in a_t with no implication for the elasticity γ_t (Buch *et al.*).[1] Fourth, the elasticity of trade with respect to trade costs might have rather increased. Based on theory (*e.g.* Anderson and van Wincoop), this elasticity is positively related to the elasticity of substitution between varieties, and it is often believed that globalization is associated with an increase in the degree of substitutability between varieties, thereby inducing an increase in the elasticity of trade to distance.

However, beyond the semantical debate about whether we are facing a puzzle, the analysis of how the elasticity of trade to distance has evolved in past decades is interesting in its own right. In that respect, the most important contribution of Coe *et al.* consists in highlighting that a nonlinear estimation of gravity equation specified in levels could lead to a radically different conclusion from that obtained using a linear estimation of the same equation in log.

Nevertheless, there is scope for improving the analysis of the "distance puzzle". Santos Silva and Tenreyro (2006) show that heteroskedasticity in trade levels is such that it biases the main parameters of interests in the log specification of the gravity equation, including the distance elasticity. They propose a Poisson Pseudo-Maximum Likelihood (PPML) estimator and argue that it is likely to be much more efficient than the nonlinear least squares (NLS) estimator. However, their study covers a single year (1990) only. With this in mind, the Coe *et al.*'s assessment could be revisited in several directions. First, the main result is established using NLS. A robustness check is performed using a Pseudo-Maximum Likelihood, but it is not totally clear which one is used. Second, the sample is restricted to 73 countries and start in 1975 only. Third, the sum of exports and imports is used as the dependent variable, but making a distinction between exports and imports, as one must according to Baldwin and Taglioni (2006), might lead to different results. Fourth, the data used for Free Trade Agreements (FTAs) are not well defined. However, because FTAs have mainly promoted regional integration, they are *de facto* inversely related to distance. Therefore, not properly controlling for FTAs might be misleading. The current paper addresses methodological issues related to the estimation of

[1] Even the presumption that transport costs have declined relative to the price of the goods being transported, *i.e.* mostly manufacturing goods, is far from obvious according to recent studies that provide direct measures of costs over different routes and modes of transportation (Hummels, 2007; Golub and Tomasik, 2008).

gravity equations and analyses carefully the sensitivity of the estimates to the treatment of FTAs. The main results are the following.

Methodology. The assumption that the conditional variance of flows is proportional to the conditional expectancy (Poisson hypothesis) cannot be rejected in any (post-1952) year. Second, the most efficient estimator might, however, be in between NLS and PPML, *i.e.* consistent with the variance of trade flows being about proportional to the square root of the conditional mean. Third, there remains a serious puzzle with respect to the whole methodology based on the class of PML estimators relying on the proportionality of the conditional variance to a power of the conditional mean. Indeed, the gamma PML estimates which should be consistent, albeit inefficient, under the proportionality assumption is significantly different from the Poisson PML ones, being actually closer to the biased OLS of the log specification. Fourth, there are some limitations to the type of estimators one can used. For example, the negative binomial PML estimator is not appropriate because it artificially depends on the unit chosen to measure the value of trade flows. Fifth, weighted least squares (WLS) of the log specification that uses observed flows as weights leads to estimates that are similar to PPML on the level specification (the same is true of iterated WLS that uses estimated flows as weights). Sixth, given the high level of serial correlation of trade flows, a "first-differencing" type of data transformation seems to be preferable to a fixed-effect estimator in a panel specification.

Empirics. Based on PPML and without controlling for FTAs, the distance elasticity of trade has been broadly stable within a 0.60-0.75 range since 1950, even though it has increased from the bottom to the top of that range since the late eighties. The gap between this PPML elasticity and that estimated based on the log specification has steadily increased over time, and this trend is shown to be related to the growing heterogeneity of trade flows. This result is consistent with the explanation proposed by Santos Silva and Tenreyro for the bias of the log specification. The most notable change refers to the effect of colonial linkages which is estimated to have basically vanished over time from a very high level in the fifties.[2] Taking into account the influence of FTAs to analyse the "distance puzzle" raises some intricate issues, as shown by Baier and Bergstrand (2007) who use a log specification. However, the influence of

[2]This is consistent with the detailed analysis of Head, Mayer and Ries (2008*a*).

FTAs appears somehow limited. In all carried out robustness checks, including panel estimates, the inclusion of FTAs does not alter the diagnosis of a broad stability of the elasticity, although within a tighter 0.60-0.70 range, still leading to a clear rejection of the rising elasticity obtained in usual log specification. Compared with Baier and Berstrand, all parameters related to the determinants of trade costs are allowed to vary over time using a panel specification. The effect of FTAs is then estimated at around 0.3 with few variations over time, which means that a trade agreement increases trade flows by about 35%.

The rest of the paper is organised as follows. Section 2 discusses the main methodological and empirical issues when estimating gravity equations to analyse the "distance puzzle". Section 3 presents the data and specifications, while Section 4 is devoted to the cross section results obtained when FTAs are not included in the list of explanatory variables. Section 5 focuses on the impact of accounting for FTAs, in both cross section and panel estimates. Section 6 briefly discusses the estimated trends in the other determinants of trade. Section 7 concludes.

2 The empirics of gravity equations

2.1 Microfoundations of gravity equations

There have been major advances in the formalisation of bilateral trade flows since the mid-nineties, as the traditional specifications of gravity equations were largely a-theoretic. In an effort to lay out the microfoundations of gravity equations, Deardoff (1998) shows that not only the bilateral distance between two countries but also their geographical positions relative to all other countries matter for the level of bilateral trade flows. Consequently, Wei (1996) and many researchers since then have added a remoteness indicator to the list of explanatory variables, approximating remoteness by the weighted average of distances from all trading partners, with trading partners' GDP as the weights.

The decisive methodological contribution of Anderson and van Wincoop (2003) consists in deriving an operational gravity model in which "multilateral resistance" that depends on all bilateral trade costs is a determinant of bilateral trade flows. The absence of the multilateral resistance terms in traditional gravity estimations leads to biased estimates of some key parameters such as the effect of a common border, as these missing terms are correlated to traditional explanatory variables. In Anderson and van Wincoop, the nominal value of exports from i to j, x_{ij}, depends on the total income, $Y_.$, of each country, world income, Y_W, the bilateral trade cost, τ_{ij}, the elasticity of substitution between all goods, σ, and the multilateral resistance, $P_.$, of each country, according to:

$$x_{ij} = \frac{Y_i Y_j}{Y_W} \left(\frac{\tau_{ij}}{P_i P_j} \right)^{1-\sigma} \tag{1}$$

where the $P_.$ terms are related to each other as follows:

$$P_j^{1-\sigma} = \sum_i P_i^{\sigma-1} \, \tau_{ij}^{1-\sigma} \, \frac{Y_i}{Y_W} \qquad \forall j \tag{2}$$

From this specification, Anderson and van Wincoop draw two implications that are especially relevant for the current study. First, the remoteness variables as commonly approximated are disconnected from the theory. Second, given a specification of trade costs, replacing the multilateral resistance terms by country fixed effect leads to consistent estimates of the gravity equation (1) in log form by ordinary least squares. Even though this fixed-effect estimator is less efficient than the nonlinear least-squares estimator that uses information on the full structure of the model, *i.e.* (1) and (2), it has the huge advantage of simplicity.

2.2 Log of gravity: consistency, efficiency and competing estimators

In turn, Santos Silva and Tenreyro (2006) highlight another typical bias of gravity equations that are estimated in log form, on top of the sample selection bias that results from the implicit exclusion of zero trade flows. Starting from a stochastic version of the gravity equation in levels such as (1), the log-linear specification generates biases as a consequence of Jensen's inequality ($E(ln\ x) \neq ln\ E(x)$), because the expected value of the logarithm of trade flows depends on higher moments, including the variance. Formally,

$$x_{ij} = exp(Z_{ij}\beta)u_{ij} \quad , \quad E(u|Z) = 1 \tag{3}$$

$$Var(u|Z) \neq 0 \ \Rightarrow \ E(ln\ u|Z) \neq 0 \tag{4}$$

where the Z explanatory variables include importer and exporter fixed effects, (log of) bilateral distances and other control variables influencing the trade costs. Since the variance of the residuals is likely to depend on explanatory variables such as importer and exporter characteristics (that cover observed ones like GDP), estimators using the log specification would bias the parameters of interest.[3] Thus, the magnitude of the bias depends on the structure of the variance of the residuals, and heteroskedasticity in the trade level equation could become a serious concern for inferences make from estimates based on the log-linear specification.

This problem can be overcome by estimating the level equation (3) using a nonlinear estimator. Santos Silva and Tenreyro (2006) propose the Poisson Pseudo-Maximum Likelihood estimator (PPML), assuming that the variance of x is proportional to its conditional expectancy, which is likely to make this estimator more efficient than the simple nonlinear least squares (NLS). Indeed, it is unrealistic to assume, as implicit with NLS, that the variance of estimated trade flows is the same for small/remote and large/central countries. Besides, whatever the specific choice of a nonlinear estimator, a level specification allows for the inclusion of zero trade flows, even though Santos Silva and Tenreyro show *ex post*, *i.e.* based on the empirical analysis, that including the zero flows does not make a material difference.

A natural extension consists in assuming other distributions than Poisson. This would include gamma distribution according to which the variance is proportional to the square of the conditional mean, and more generally any power of it. Some authors have also used the negative binomial distribution (e.g. Head, Mayer and Ries, 2008b), but this is inappropriate when applied to trade flows because such an estimator artificially depends on the choice of the nominal unit of the dependent variable. Indeed, the assumption of the negative binomial distribution is:

$$Var(x|Z) = E(x|Z) + \eta^2 E^2(x|Z) \tag{5}$$

where η is a scalar to be estimated. The problem arises because the ratio between the expectancy of x_{ij} and its square can be made either infinitely small or large depending on the unit choice. Formally, if the unit is changed such that the empirical analysis is conducted

[3] Indeed, in that case, the conditional variance depends on Z, and the bias is not limited to the constant (see the last paragraph of this sub-section).

on $X = \kappa\ x$, assuming that X follows a negative binomial distribution implies that:

$$Var(X|Z) = E(X|Z) + \eta^2 E^2(X|Z) = \kappa E(x|Z) + \kappa^2\eta^2 E^2(x|Z)$$

Hence, when $\kappa \rightarrow 0$, $Var(X|Z) \approx \kappa E(x|Z) = E(X|Z)$, and the negative binomial PML estimator tends towards PPML. Conversely, when $\kappa \rightarrow \infty$, $Var(X|Z) \approx \kappa^2\eta^2 E(x|Z) = \eta^2 E^2(X|Z)$, and the negative binomial PML estimator tends towards gamma PML.

In order to discriminate between the various *a priori* legitimate PML estimators, Manning and Mullahy (2001) suggest that if $Var(x_{ij}|Z) = \lambda_0\ E(x_{ij}|Z)^{\lambda_1}$, the choice of the appropriate estimator can be based on an asymptotically valid estimate of λ_1 from:

$$(x_{ij} - \tilde{x_{ij}})^2 = \lambda_0\ \tilde{x_{ij}}^{\lambda_1} + \zeta_{ij} \tag{6}$$

where $\tilde{x_{ij}}$ is the value of $E(x_{ij}|Z)$ estimated from an initially consistent estimator like PPML.[4]

A final comment refers back to the estimation of gravity equations in logarithm. A Taylor series that is limited to the second moment around the conditional mean gives:

$$Log\ x_{ij} \approx Log\ E(x_{ij}|Z) + \frac{x_{ij} - E(x_{ij}|Z)}{E(x_{ij}|Z)} - \frac{1}{2}\frac{[x_{ij} - E(x_{ij}|Z)]^2}{E^2(x_{ij}|Z)}$$

And therefore[5],

$$E(Log\ x_{ij}|Z) \approx Log\ E(x_{ij}|Z) - \frac{1}{2}\frac{Var(x_{ij}|Z)}{E^2(x_{ij}|Z)} \tag{7}$$

$$Var(Log\ x_{ij}|Z) \approx Var(x_{ij}|Z)/E^2(x_{ij}|Z) \tag{8}$$

On top of the possible selection bias due to the elimination of zero trade flows, these equations summarize two issues with the log of gravity. Equation (7) highlights the bias that is stressed by Santos Silva and Tenreyro. Beyond that bias, equation (8) shows that assuming that

[4]Santos Silva and Tenreyro actually suggest testing the adequacy of a particular value of λ_1 from a Taylor expansion of (6), which they apply in the empirical part of their paper. Unfortunately, this procedure is subject to the same problem as for the negative binomial estimator: it artificially depends on the unit choice of trade flows, and could therefore be misleading.

[5]The computation of the variance uses the Taylor series at the first order (Delta method). This is only an approximation of course, and going to the second degree sometimes improves the approximation of the variance importantly (*e.g.* Tiwari and Elston, 1999).

errors of the log specification are i.i.d., as implicit when estimating with OLS, is consistent with the conditional variance of the flow being proportional to the square of the conditional mean, *i.e.* with the gamma distribution. Therefore, if the true distribution were gamma, estimating the log level equation using OLS would only bias the intercept (ignoring the sample selection bias, see eq. 7), and not the other parameters of interest such as the distance coefficient. In other words, the magnitude of the biases (except the constant) of the gravity equation that is estimated using the log-linear specification depends on how far the distribution of trade flows is from the gamma distribution. If, however, the true distribution is Poisson, (7) and (8) become, where α is a constant, respectively:

$$E(Log\ x_{ij}|Z) \approx Log\ E(x_{ij}|Z) - \alpha/E(x_{ij}|Z)$$

$$Var(Log\ x_{ij}|Z) \approx 2\alpha/E(x_{ij}|Z)$$

In that case, the bias would be very severe for small flows. Moreover, OLS estimates of the log specification would ignore that the variance of the log is very large for small flows; in other words, it would give far too much weight to small flows.[6] Ignoring the bias, the "efficient" weighted least squares (WLS) is obtained in that case by weighting each observation by the inverse of the variance, *i.e.* in that case by the conditional mean. This WLS estimator of the log specification might be appealing because it reduces the bias mechanically, as low weights are given to observations that contribute the most to the bias.

2.3 Panel estimates and the "distance puzzle"

As argued by Baier and Bergstrand (2007), cross section estimates of gravity equation might be biased due to the endogeneity of free trade agreements (FTAs). In particular, properly controlling for the influence of FTAs might be important for the estimate of the evolution of the elasticity of trade with respect to distance. Indeed, FTAs cover an increasing share of world trade and are often agreements between neighbouring countries, hence an obvious

[6]In contrast, NLS of the trade level specification, although consistent, might be inefficient because they do not give enough weights to small flows. It has been checked that iterative weighted NLS of the level equation, where weights are the inverse ratios of the conditional mean, converges to PPML estimates (Davidson and MacKinnon, 2003, Chapter 11).

correlation between FTAs and distances. For example, Berthelon and Freund (2008) refer to regionalism as the most obvious explanation for the persistence of distance as a determinant of trade flows. However, finding an instrument for FTAs that does not influence trade by any other channel is extremely difficult. In this context, Baier and Bergstrand argue that country pair idiosyncrasies should be accounted for via so-called "dyadic fixed effects" to eliminate the bias due to the endogenity of FTAs; more generally panel specifications of gravity equations make it possible to control for a battery of fixed effects. Besides, multilateral resistance terms require the inclusion of origin and destination country dummies for each year in panel data. Importantly, when introducing dyadic fixed effects, the level of the distance elasticity is lost, and only the changes through time can be estimated. It is clear also that introducing the $i * j$, $i * t$, $j * t$ fixed effects in a nonlinear specification represents a numerical challenge that could be adressed *e.g.* by period-averaging.

3 Data and econometric specification

3.1 Data

Trade flow data come from the IMF Direction Of Trade Statistics (DOTS) database. This database provides trade flows for a long period, starting in 1948, which is important to study properly the distance puzzle, and for 205 trade partners. Moreover, DOTS includes zeros and differentiates them from missing values, avoiding some necessary but doubtful assumptions when otherwise. Figure 1 plots the number of zero and non-zero trade flows through time, thereby illustrating the risk of selection bias using log-linear OLS. Indeed, despite the decreasing share of zero trade flows from 80% in 1948 to 29% in 2006, it still represents an important proportion. The sample of striclty positive trade flows, used for comparison of the different estimators, has about 3,700 flows in 1948 and 22,000 in 2006.

The geographical variables, distance between countries, common border, common language and colonial linkage dummies, are taken from the CEPII database.[7] The FTA data is broadly the same as the one used in Baier and Bergstrand (2007). Specifically, the database used by

[7]http://www.cepii.fr/anglaisgraph/bdd/distances.htm, Centre d'Etudes Prospectives et d'Informations Internationales.

these authors has been corrected and improved by Fontagné and Zignago (2007) in their re-estimation of the impact of FTAs.[8] The proportion of the value of world trade covered by FTAs goes from 7% in 1958 to 31% in 2006.

Finally, it proved useful to work also with a balanced panel to account for the increasing number of trade flows as well as for the change in the sample over time. Hence, the largest possible balanced panel consits of the same 2550 pairs of countries between 1952 and 2006 covering 90 countries and 78% of world trade on average. In particular, the construction of the balanced panel drastically reduces the number of dyadic fixed effects, that are necessary to estimate the FTA effect in panel.

3.2 Specification

Following the discussion in section 2, the gravity equation is estimated in levels including importer and exporter fixed effects. Although the Poisson PML is used for the central estimates, NLS and gamma PML estimators are also computed. Moreover, the most efficient power of the conditional mean is estimated according to equation (6), which enables to test the Poisson assumption.

Formally in the cross section analysis, the following equation is estimated for each year:

$$x_{ij} = exp(\alpha_0 + \gamma\ ln\ d_{ij} + \alpha_1\ B_{ij} + \alpha_2\ L_{ij} + \alpha_3\ C_{ij} + \alpha_4\ FTA_{ij} + FX_i + FM_j)u_{ij} \qquad (9)$$

with $Var(x_{ij}|Z) = \lambda_0\ E(x_{ij}|Z)^{\lambda_1}$.

x_{ij} is the nominal US$ value of export from i to j, FX_i and FM_j are the fixed effects for exporting and importing countries, respectively. B_{ij}, L_{ij} and C_{ij} are the traditionnal control covariates: common border, common official language and colonial linkage dummies, respectively. u_{ij} are the multiplicative error terms of the nonlinear estimates. The log version is also estimated using ordinary and weighted least squares estimators (sub-section 2.2).

In order to separate the various factors influencing the analysis of the distance puzzle, the gravity equations are first estimated without controlling for FTAs (FTA_{ij}). Because these first

[8] Compared with Fontagné and Zignago, FTA data has been updated beyond 2000. In total, 47 FTAs are covered. The first FTA in the database is the European Economic Community. Its treaty was signed on March 25th, 1957, so it begins in 1958 in the database.

results might be subject to the omitted variable biases, the analysis focuses, in a second step, on the effect of controlling for FTAs. Finally, following the discussion in sub-section 2.3, a panel specification including dyadic fixed effect is estimated:

$$x_{ijt} = exp(\gamma_t \, ln \, d_{ij} + \alpha_{1t} \, B_{ij} + \alpha_{2t} \, L_{ij} + \alpha_{3t} \, C_{ij} + \alpha_{4t} \, FTA_{ijt} + FX_{it} + FM_{jt} + Dyadic_{ij}) u_{ijt} \quad (10)$$

4 Cross section results without controlling for FTAs

This section presents the cross section results obtained without controlling for free trade agreements. The focus is on the elasticity of trade with respect to bilateral distance, the other parameters of interest in the gravity equation are discussed in greater detail in Appendix A. As a benchmark, estimates using the PPML are presented in sub-section 4.1 while sub-section 4.2 shows why this benchmark is actually the baseline.

4.1 PPML

The gravity equation as specified in equation (9) is estimated for each year by PPML. Table 1 presents the results for six specific years between 1955 and 2005. The elasticity of trade to distance is estimated to have been broadly stable over the period within a (0.60, 0.75) range. This range is tight compared with those found in the literature based on log specifications. The estimated robust standard error has steadily declined from 0.040 to 0.025 indicating an improvement over time in the precision of the estimate.

Figure 2 presents the evolution of the trade elasticity to distance, estimated using either OLS in logs or PPML in levels, along with confidence intervals. The PPML estimates are not sensitive to whether the zero trade flows are included or not (confidence intervals are also similar), a result also found by Santos Silva and Tenreyro, and Coe *et al.* using NLS. Based on the log-linear specification, the elasticity would have steadily increased from 0.70 to 1.60, which characterizes the distance puzzle. As a result, the difference between "PPML" and "log-linear" elasticities has dramaticaly increased over time.

This increasing difference seems to be due to the *growing* heterogeneity of flows and induced heteroskedasticity, consistent with the idea introduced by Santos Silva and Tenreyro. The intuition behind such a link is illustrated in two ways. First, Figure 3 replicates Figure 2 and adds the elasticity estimated from the smaller albeit more homogenous balanced panel. While the PPML estimate is not sensitive to the choice of the sample, the reduced heterogeneity in the balanced panel leads to a lower estimated elasticity in the log-linear specification compared with that for the whole sample, the more so for the more recent years. Second, two measures of dispersion and one of heteroskedasticity were computed. The measures of dispersion are the interquartile ratio (ratio of 3rd to 1st quartile) of trade flows, and the coefficient of variation (standard deviation divided by mean). They are computed on the sample on which the log-linear specification is based, *i.e.* without zeros (again, the inclusion of zero flows has minor effects on the PPML estimates). The measure of heteroskedasticity related to the bias of the log-linear estimator is the share of the variance of $log\ \hat{u}$ explained by (the log of) distance, where $\hat{u}$ is the PPML estimated multiplicative residual. Indeed, according to Jensen's inequality, the bias of log OLS is due to the dependence of $log\ u$ on Z.[9]

Figure 4 represents these three indicators in addition to the difference in the distance elasticity between PPML and log-linear OLS. Over the period, the q3 / q1 ratio has increased by a factor of 12. This is due to the tremendous increase in small non-zero flows, as q1 decreased from $ 2 millions (US $ is deflated by US GDP deflator with 2000 as the base year) in the 1950s to $ 0.1 million since the mid-1990s. Within the same period, the average flow increased from $ 100 M to $ 400 M, and the standard deviation increased even faster as the coefficient of variation rised from 4 to 10. Since the small non-zero flows carry a disproportionate weight in log, the increase in its share is likely to contribute heavily to the widening of the gap between the PPML and the log-linear elasticities. Visually, the variations of the difference between log OLS and PPML elasticities are strikingly closely related to those of the contribution of distance to the variance of the residuals.

[9]The log of the residual is directly related to the bias as equation 3 and 7 imply that $E(log\ u|Z) \approx -\frac{1}{2}Var(u|Z)$.

4.2 Which pseudo-maximum likelihood estimator?

This part investigates whether the assessment of the distance puzzle is sensitive to the choice of the nonlinear estimator among the class that verifies $Var(x_{ij}|Z) = \lambda_0 E(x_{ij}|Z)^{\lambda_1}$, all of them being consistent under (3). This includes the nonlinear least squares (NLS, $\lambda_1 = 0$), the PPML ($\lambda_1 = 1$) and the gamma PML (GPML, $\lambda_1 = 2$).[10]

Figure 5 compares the NLS and PPML estimates of the distance elasticity. The level and evolution of the estimated elasticity is similar between the two estimators, even though the variations are greater with NLS (and the elasticity unrealistically low at the beginning of the period). On average, the NLS estimated standard error is twice as large as the PPML one, suggesting that PPML is more efficient. The main difference in the point estimates is that the elasticity is estimated to have fallen since the mid-1980s with NLS, while it is broadly stable with PPML.

The significant difference between the elasticity estimated by PPML and GPML is striking, as illustrated by Figure 6.[11] Actually, the trend in the GPML elasticity looks very similar to the log-linear OLS one, which is problematic: even though both GPML and log-linear OLS give a high weight to small flows, which might be a source of poor efficiency (and bias for log), the GPML should be consistent under (3). However, it is significantly different from PPML, another consistent estimator under (3): this is the main remaining puzzle of the whole approach.

Figure 6 also adds the negative binomial estimator for different unit values of the trade flows in order to illustrate the theoretical result established in sub-section 2.2: this estimator is indeed sensitive to the unit choice, converging at the limits either to the PPML or the GPML, which makes it inadequate to estimate gravity equations.

As discussed in Section 2, estimating equation (6) should help to select the most efficient estimator. However, this test actually proves to be deficient, as its conclusions depend on the choice of the estimator used to calculate $\tilde{x_{ij}}$. When PPML is used, the average estimate over the period is 1.03 with an average estimated standard error of 0.26 (Figure 7 and Table 2 for

[10]For $\lambda_1 = 0$, NLS or maximum likelihood leads to almost identical estimates.

[11]The standard errors of both estimators, not represented in the graph, are around 0.03, which makes them significantly different.

specific years). PPML is never rejected as optimal, whereas NLS and GPML always are at 95% confidence level (except for a few years for GPML). When NLS is used to calculate $\tilde{x_{ij}}$, λ_1 is estimated at 0.64 on average. So far, this seems to clearly discriminate in favour of PPML. Unfortunately, when GPML is used for $\tilde{x_{ij}}$, λ_1 is estimated at 2.31 on average, indicating that GPML should be preferred over both PPML and NLS. The reason why this test is inadequate is not totally clear; our hypothesis is that GPML is not consistent to start with, and this is left for further research.

GPML is a questionable estimator according to three types of arguments. The first is based on a judgement call that is illustrated by taking two trade flows of \$ 10,000 and \$ 1 Billion. While NLS gives the same importance to making an estimated error of \$ 10,000 on each of them, GPML gives the same importance to making an error of say 10% of each flow, i.e. of \$ 1,000 and \$ 100 Millions, respectively. PPML seems to be a good compromise between these two extremes. Second, Figure 8 presents the scatter plots of the observed flows (x-axis) and estimated ones (y-axis) using NLS, PPML and GPML. The fit for GPML is strikingly very bad. As a matter of fact, the sum of squares of the residual is equal to 4.3% of the total variance with NLS, 9.8% with PPML and 4200% with GPML. Although NLS does better by definition based on this indicator, the poor performance of GPML is extreme. Finally, and perhaps even more convincingly, for 2000 as an example, the total sample has been split in two using the median of trade flows as a cut-off. As shown in Table 3, NLS and PPML produce similar estimates on the whole and split samples. In contrast, with GPML (and log OLS), the elasticity is lower with both the below- and above-median samples compared with the whole sample, and closer to PPML. This suggests that GPML does not account for heterogeneity properly. We take all these as evidence against GPML.

Finally, as PPML seems to be the preferred estimator, one is tempted, following the discussion in sub-section 2.2, to compute the WLS estimator of the log specification, where the weights are either observed or iterated estimated (starting with OLS) trade flows. Both WLS linear estimates are amazingly close to the PPML ones, as shown in Figure 9. This highlights that using WLS to improve efficiency is also powerful to reduce, or even eliminate, the bias of the OLS log specification.

5 Results with FTAs, and panel specification

Sub-section 5.1 discusses whether the baseline obtained in section 4 without controlling for FTAs is sensitive to different values of the FTA parameter when it is constrained in cross sections. Sub-section 5.2 analyses the "distance puzzle" using panel data.

5.1 Sensitivity to the effect of FTAs in cross section

Accounting for FTAs has a small impact on the assesment of the "distance puzzle". The sensitivity of the distance elasticity of trade to the FTA parameter is shown in Figure 10, the parameter being equal to 0, 0.3 or 0.6 using PPML in cross sections. These values are based on the different estimates of the FTA effect in the panel approach (sub-section 5.2). Naturally, the first case (parameter = 0) corresponds to the results presented in the previous section. As expected, because FTAs are negatively correlated with distance, taking into account the effect of FTAs reduces the estimated distance coefficient, although the difference is never greater than 0.12.

Whether the FTA parameter is contrained to 0, 0.3 or 0.6 does not make any difference in the evolution of the elasticity of trade to distance until 1972. From 1973, the gap between the different estimates increases, as does the coverage of trade flows by FTAs (right scale).[12] When the FTA parameter is constrained to 0.3, for example, the estimated distance coefficient remains stable around 0.65 (average standard error around 0.03) and within a 0.60-0.70 range. From the mid-eighties the evolution is hump-shaped with a recent increase from 0.60 to 0.70 between 1994 and 2006.

In sum, the main difference between these three estimates lies in the end point value for the elasticitiy. It is 0.76, 0.71 and 0.66 in 2006 when the FTA parameter is constrained to 0, 0.3 and 0.6 respectively. The shape of the evolution is affected accordingly, but without changing the broad assessment of a stable elasticity over the whole period. In particular, these differences are small compared with the magnitude of the bias indentified by Santos Silva and Tenreyro.

[12] 1973 is an important year for FTAs: the United Kingdom, Danemark and Ireland join the European Union. An agreement between the EU and the European Free Trade Association is created, as well as the Carribean Community.

5.2 Panel analysis of the distance puzzle

Following section 2.3, the panel estimation of gravity equations requires both time-varying importer and exporter fixed effects as well as "dyadic fixed effects" in order to account for bilateral heterogeneity. All these fixed effects generate computational difficulties. As a result, the estimation of equation (10) has been carried out using 5-year averaging.[13] Also, working on the perfectly balanced panel dataset presented in section 3 (covering 90 countries) allows to drastically reduce the number of fixed effects.

With dyadic fixed effects, only the evolution through time can be estimated, the levels of the parameters that are fixed through time such as geographic characteristics being wiped out. However, unlike Baier and Bergstrand, elasticities are here allowed to vary. As shown in Figure 11, estimation of equation (10) using PPML leads to a similar evolution of the trade elasticity to distance to that obtained in the cross section analysis, even though the estimated elasticity has increased by about 0.10 since the mid-60s.

Depending on the autocorrelation level of the residuals, the fixed-effect estimator might not be the most efficient (Woolridge, 2001, chapter 10). As robustness checks, two other panel estimators, that mirrors first-differencing, have been implemented to deal with the dyadic fixed effects. The first uses x_{ijt}/x_{ijt-1} as the dependent variable.[14]

$$\frac{x_{ijt}}{x_{ijt-1}} = exp(\Delta\gamma_t \; ln \; d_{ij} + \Delta\alpha_{1t} \; B_{ij} + \Delta\alpha_{2t} \; L_{ij} + \Delta\alpha_{3t} \; C_{ij} + \Delta(\alpha_{4t} \; FTA_{ijt}) + \Delta FX_{it} + \Delta FM_{jt}) v_{ijt} \tag{11}$$

where $v_{ijt} = u_{ijt}/u_{ijt-1}$.

Due to heteroskedasticity, this specification is estimated efficiently by weighted NLS where the weights are the inverse of the variance of x_{ijt}/x_{ijt-1}. Appendix B shows that under the Poisson assumption $Var(x_{ijt}/x_{ijt-1}|Z)$ is proportional to the inverse of $E(x_{ijt-1}|Z)$. Hence, the

[13] To be consistent with the gravity specification in levels such as (9) or (10), the geometric mean of trade flows is used as the dependant variable. By comparison, Baier and Bergstrand use a specification in logs with elasticities with respect to distance, border, colonial link, etc. that are constant through time, and reduce the number of fixed effects by keeping only one out of five years.

[14] The balanced panel does not include zero flows which are eliminated by this transformation. The analysis in section 4 has shown that zero flows do not matter anyway.

ratio x_{ijt}/x_{ijt-1} is weighted by x_{ijt-1}.[15] The second alternative estimator, following the results established at the end of section 4, consists in a linear WLS of equation (11) specificied in logs

$$Log(\frac{x_{ijt}}{x_{ijt-1}}) = \Delta\gamma_t \, ln \, d_{ij} + \Delta\alpha_{1t} \, B_{ij} + \Delta\alpha_{2t} \, L_{ij} + \Delta\alpha_{3t} \, C_{ij} + \Delta(\alpha_{4t} \, FTA_{ijt}) + \Delta FX_{it} + \Delta FM_{jt} + \epsilon_{ijt} \quad (12)$$

Appendix B shows that the efficient estimator also uses x_{ijt-1} as weights.

As shown in Figure 11, the broad picture is not affected by the choice of estimators, supporting the view of an overall stability of the trade elasticity to distance through time. However, although the choice of the parameter has basically no impact on the assesment of the distance puzzle, some other parameters are affected, as shown in Section 6. Testing serial correlation of the residuals (Wooldridge, equation 10.71) enables to discriminate between the fixed-effect and the "first-difference" estimator, depending on whether the auto-correlation parameter of $log \, v_{ijt}$ is close to 0 and 1. Based on this test, the "first-difference" estimators (eq. 11 and 12) should clearly be prefered over the level estimator (eq. 10).[16]

6 Other trade determinants

6.1 FTA parameter

Although Baier and Bergrstrand constrain the elasticities to be constant over time, the panel approach allows for time-varying elasticities, as required for the analysis of the "distance puzzle". However, for comparison purposes, estimates of equations (10) and (11) are also reported holding the FTA parameter α_{4t} to be constant over time, and Figure 12 presents the evolution of the FTA parameter across the different estimators. By contrast to the elasticity of distance, the effect of FTAs is sensitive to the choice of the panel estimator. Holding the FTA parameter constant leads to an estimate of 0.52 and 0.27 using equations (10) and (11), while Baier and Bergstrand obtain 0.49 and 0.36, respectively, using a specification of trade flows in

[15]Ideally one would like to use an iterated weighted NLS, but only the ratios are estimated, not the level of flows. Fortunately, results in section 4 indicate that WLS and iterated WLS estimates are very close to each other.

[16]This auto-correlation parameter is estimated to be very close to 0 in both weighted NLS and log-linear WLS (eq. 11 and 12, respectively).

logarithm. While the "first-difference" transformation (eq. 11) generates only small variations of the FTA parameter through time, the estimated coefficient using the level specification (eq. 10) varies from 0.27 to 0.74 between 1952 and 1975, and steadily decreases to 0.50. As discussed above, the tests for the serial correlation of residuals discriminate in favour of the "first-difference" transformation. A point estimate of 0.27 means that FTA increase trade by 35%.[17]

6.2 Contiguity, common language and colonial linkages dummies

Because the dummies for colonial linkages, contiguity and common official language do not vary through time, only the changes of the corresponding coefficients can be estimated in panel specifications. Figures 13(a-c) present these evolutions using the three panel estimates, as well as those obtained using PPML in cross sections. The latter is particulary useful as it provides estimates of the parameters in levels. For the first period (1952-1956), the cross section point estimates for colonial linkages, common border and common official language dummies are 1.11, 0.48 and 0.14, respectively, with estimated robust standard errors of 0.15, 0.14 and 0.13.

All estimators point to a notable decrease in the effect of colonial linkages over the years, especially pronounced through the mid-eighties, even though the amplitude varies importantly across estimators (Figure 13a). Based on cross sections, the effect sharply decreases from 1.1 to about 0 over the whole period, while the decrease is of "only" 0.5 with PPML in panel and of 0.9 with the nonlinear "first-difference" panel estimator. Figure 13b shows that the effect of contiguity is broadly stable in cross sections, around an average value of 0.47. This elasticity decreases by 0.20 on average across panel estimators over the whole period. Finally, the effect of a common official language is broadly stable for all estimators, except for the nonlinear "first-difference" panel one leading to an estimated increase of 0.2 throughout the whole period (Figure 13c).

[17] $exp(0.27) - 1$.

7 Conclusion

Evolution through time of the main determinants of trade flows has been mostly studied through "log of gravity" specifications, generally pointing to an increasing role of distance in shaping world trade. This paper focuses on this "distance puzzle" starting with the Poisson Pseudo-Maximum Likelihood (PPML) estimator recently proposed by Santos Silva and Tenreyro, which they apply to one single year. The main result is that this elasticity has been broadly stable since the 1960s at around 0.65-0.70. Despite the increasing coverage of free trade agreements, taking FTAs into account has a small effect on the distance elasticity. Another interesting finding is that the effect of colonial linkages has been vanishing.

The paper brings additional methodological contributions by comparing various estimators to PPML; OLS when the trade equation is specified in logs, nonlinear least-squares (NLS), negative binomial and gamma PML when specified in levels, as well as panel estimates based on levels and "first-difference" specifications. First, the bias identified by Santos Silva and Tenreyro between log OLS and PPML is shown to have increased over time in relation with the growing heterogeneity of trade flows. Weighting the logs by the (estimated) level of trade flows is sufficient to eliminate this difference. Second, although the PPML estimator is the most satisfactory of those tested, the most efficient one might be in between NLS and PPML, corresponding to the proportionality of the variance of trade flows to the square root of the conditional expectancy. PML based on negative binomial is not an option as it artificially depends on the unit chosen to measure trade flows. Third, there remains a methodological puzzle. The gamma PML estimator is not only inefficient, giving too much weight on small flows, it also leads to significantly different estimates than PPML, contrary to what one should expect. Finally, auto-correlation tests of residuals discriminate the panel estimates in favour of the "first-difference" estimator relative to the fixed effects one.

APPENDIX

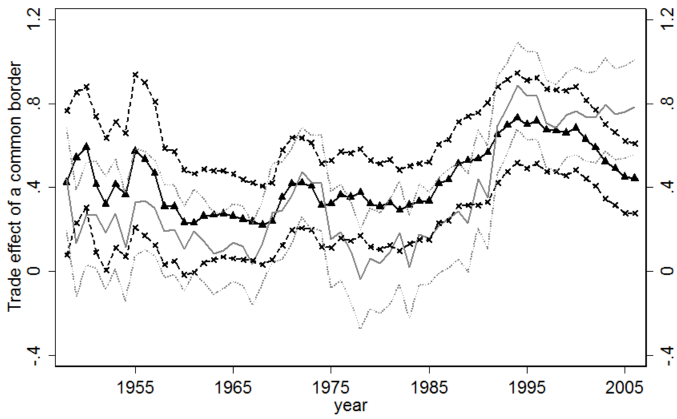
A Trade determinants in cross sections

In this Appendix, equation (9) is estimated for each year in cross section constraining the FTA parameter to 0.3. PPML and log OLS lead to similar estimates of the impact of a common border on international trade (Figure A1a). The parameter is estimated at about 0.3 between 1948 and 1985. Then the coefficient increases for the two specifications to about 0.7 in the early nineties.

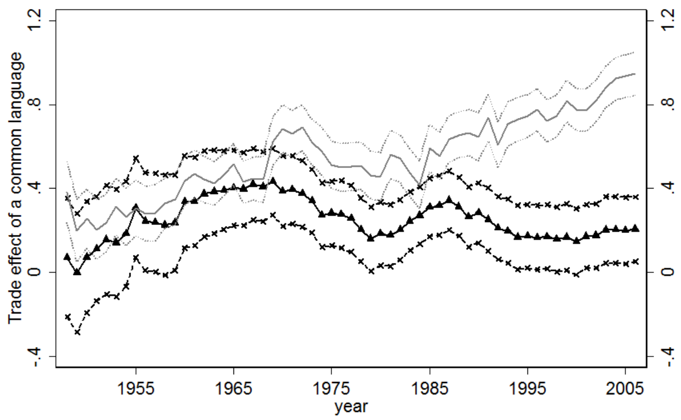
The effect of a common official language is broadly stable around 0.2 when estimated by PPML (Figure A1b). In contrast, log OLS leads to a steady increase, somehow unrealistically, of the estimated impact up to 0.95 in 2006.

The evolution of the estimated effect of having colonial linkages is illustrated in Figure A1c. While the two methods highlight the declining importance of colonial links, PPML seems to produce here also more realistic results. According to PPML, the impact of colonial linkages has been stable at about 0.2 since the mid-eighties, whereas it is still around 1.0 based on log OLS.

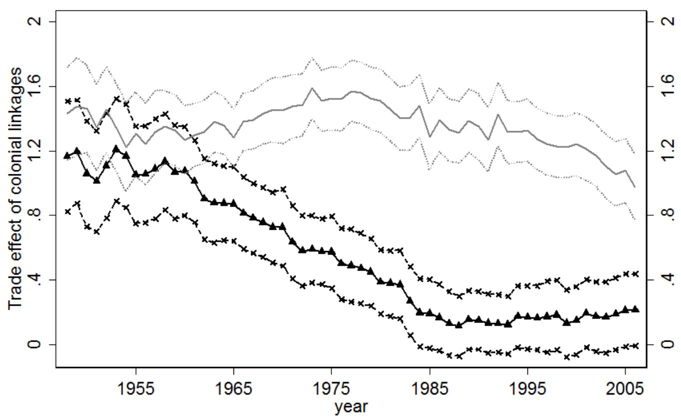
Figure A.1: Impact of trade determinants other than distance



(a) Impact of sharing a common border on trade



(b) Impact of sharing a common official language on trade



(c) Impact of colonial linkages on trade

Notes: C.I. = Confidence Interval. A gravity equation is estimated for each year both in levels with the Poisson Pseudo-Maximum Likelihood estimator and in log with OLS.

B "First-difference" transformation in panel estimates

Estimation of the panel specification (10) with pseudo-maximum likelihood might be inefficient depending on the structure of the residuals. An appealing transformation that gets rid of the dyadic fixed effects and is consistent with the level specification is the ratio x_t/x_{t-1} as given by (11).

$$\frac{x_{ijt}}{x_{ijt-1}} = exp(\Delta\gamma_t \; ln \; d_{ij} + \Delta\alpha_{1t} \; B_{ij} + \Delta\alpha_{2t} \; L_{ij} + \Delta\alpha_{3t} \; C_{ij} + \Delta(\alpha_{4t} \; FTA_{ijt}) + \Delta FX_{it} + \Delta FM_{jt}) v_{ijt} \tag{13}$$

Assuming that the conditional variance of trade flows is proportional to the conditional mean (Poisson assumption), equation (13) can be estimated efficiently by weighted (nonlinear) least-squares where the weights are the inverse of the conditional variance of x_t/x_{t-1}. The latter can be computed from a Taylor-series of the ratio around the means. Indeed, for any two variables X and Y,

$$\frac{Y}{X} - E\Big(\frac{Y}{X}\Big) \approx -\frac{EY}{E^2X}(X - EX) + \frac{1}{EX}(Y - EY) \tag{14}$$

It follows that, noting ρ the linear correlation coefficient between X and Y:

$$Var\left(\frac{Y}{X}\right) = \frac{E^2Y}{E^4X} Var(X) + \frac{1}{E^2X} Var(Y) - 2\rho\frac{EY}{E^3X}\sqrt{Var(X)Var(Y)} \tag{15}$$

where κ is a constant. Equation (15) is used with $X = x_{t-1}|Z$ and $Y = x_t|Z$, assuming that the conditional variance is proportional to the conditional mean and that the conditional mean of x_t is proportional to the conditional mean of x_{t-1}:

$$Var\left(\frac{x_{ijt}}{x_{ijt-1}}|Z\right) = \frac{\kappa}{E(x_{ijt-1}|Z)}$$

Hence, based on these assumptions, the efficient estimator of (11) uses $E(x_{ijt-1}|Z)$ as weights.[18]

[18] κ is proportional to $(g + g^2 - 2\rho g^{3/2})$ where $g = E(x_{ijt}|Z)/E(x_{ijt-1}|Z)$.

A similar exercise is replicated to compute the weights for the log specification in eq. (12):

$$Log\frac{Y}{X} - E\left(Log\frac{Y}{X}\right) \approx \frac{1}{EY}(Y - EY) - \frac{1}{EX}(X - EX)$$

from which one derives:

$$Var\left(Log\frac{x_{ijt}}{x_{ijt-1}}|Z\right) = \frac{\tilde{\kappa}}{E(x_{ijt-1}|Z)}$$

where $\tilde{\kappa}$ is a constant.[19]

[19] $\tilde{\kappa}$ is proportional to $(1 + g - 2\rho/g)$.

References

Anderson, J. and van Wincoop, E. (2003), 'Gravity with Gravitas: A Solution to the Border Puzzle', *American Economic Review* **93**(1), 170–192.

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.

Baldwin, R. and Taglioni, D. (2006), 'Gravity for Dummies and Dummies for Gravity Equations', *CEPR Discussion Paper* (5850).

Berthelon, M. and Freund, C. (2008), 'On the Conservation of Distance in International Trade', *Journal of International Economics* **75**(2), 310–320.

Buch, C. M., Kleinert, J. and Toubal, F. (2004), 'The Distance Puzzle: On the Interpretation of the Distance Coefficient in Gravity Equations', *Economics Letters* **83**(3), 293–298.

Coe, D. T., Subramanian, A. and Tamirisa, N. (2007), 'The Missing Globalization Puzzle: Evidence of the Declining Importance of Distance', *IMF Staff Papers* **54**(1).

Davidson, R. and MacKinnon, J. G. (2003), *Econometric Theory and Methods*, New York: Oxford University Press.

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: University of Chicago Press.

Disdier, A.-C. and Head, K. (2008), 'The Puzzling Persistence of the Distance Effect on Bilateral Trade', *The Review of Economics and Statistics* **90**(1), 37–48.

Fontagné, L. and Zignago, S. (2007), 'A Re-evaluation of the Impact of Regional Agreements on Trade Patterns', *Economie Internationale* **109**(1), 31–51.

Golub, S. and Tomasik, B. (2008), 'Measures of International Transport Cost for OECD Countries', *OECD Economics Department Working Papers* (609).

Head, K., Mayer, T. and Ries, J. (2008*a*), 'The Erosion of Colonial Trade Linkages After Independence', *CEPR Discussion Paper* (6951).

Head, K., Mayer, T. and Ries, J. (2008*b*), 'How Remote is the Offshoring Threat?', *forthcoming the in European Economic Review* .

Hummels, D. (2007), 'Transportation Costs and International Trade in the Second Era of Globalization', *Journal of Economic Perspectives* **21**(3), 131–154.

Leamer, E. (2007), 'A Flat World, a Level Playing Field, a Small World After All, or None of the Above? Review of Thomas L. Friedman's *The World is Flat*', *Journal of Economic Literature* **45**, 83–126.

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

Santos Silva, J. and Tenreyro, S. (2006), 'The Log of Gravity', *The Review of Economics and Statistics* **88**(4), 641–658.

Tiwari, H. K. and Elston, R. C. (1999), 'The Approximate Variance of a Function of Random Variables', *Biometrical Journal* **41**(3), 351–357.

Wei, S. (1996), 'Intra-national versus International Trade: How Stubborn are Nation States in Globalization?', *NBER working paper* (5331).

Wooldridge, J. M. (2002), *Econometric Analysis of Cross Section and Panel Data*, Cambridge MA.: The MIT Press.

Tables

Table 1: Gravity equations estimated with PPML, cross sections

year	1955	1965	1975	1985	1995	2005
Log distance	-0.53***	-0.62***	-0.70***	-0.72***	-0.66***	-0.75***
	(0.06)	*(0.04)*	*(0.03)*	*(0.03)*	*(0.03)*	*(0.03)*
Contiguity	0.57***	0.32***	0.35***	0.33***	0.65***	0.43***
	(0.19)	*(0.11)*	*(0.11)*	*(0.09)*	*(0.10)*	*(0.09)*
Language	0.31**	0.40***	0.29***	0.32***	0.18**	0.20**
	(0.12)	*(0.09)*	*(0.08)*	*(0.07)*	*(0.08)*	*(0.08)*
Colony	1.05***	0.85***	0.53***	0.14	0.11	0.16
	(0.15)	*(0.12)*	*(0.12)*	*(0.11)*	*(0.10)*	*(0.12)*
Country FE	yes	yes	yes	yes	yes	yes
Observations	4558	7449	11649	13063	19973	22201

Notes: A gravity equation is estimated for each year, where the dependant variable is the level of bilateral trade flows. Standard errors are in parentheses; ***, ** and * are significance levels at the 1%, 5% and 10% thresholds, respectively.

Table 2: λ_1 is estimated from equation (6) using the different estimators as starting points for $\tilde{x_{ij}}$

year	1955	1965	1975	1985	1995	2005
using PPML ($\lambda_1 = 1$)	1.33	1.06	0.80	1.07	0.96	1.06
	(0.29)	*(0.30)*	*(0.14)*	*(0.35)*	*(0.22)*	*(0.28)*
using NLS ($\lambda_1 = 0$)	0.66	0.65	0.64	0.55	0.63	0.63
	(0.12)	*(0.10)*	*(0.08)*	*(0.06)*	*(0.07)*	*(0.07)*
using GPML ($\lambda_1 = 2$)	2.33	2.92	1.93	2.58	2.16	2.10
	(0.28)	*(0.43)*	*(0.31)*	*(0.22)*	*(0.06)*	*(0.05)*

Notes: Standard errors are in parentheses. Equation (6) is estimated for each year. Taking the year 2005 as an example, λ_1 is estimated at 1.06 using PPML for $\tilde{x_{ij}}$, at 0.63 using NLS and at 2.10 using GPML

Table 3: Sample analysis of the different estimators, above and below the median of trade flows. Trade elasticity to distance, 2000

	PPML	NLS	GPML	OLS log
Whole sample	-0.69	-0.53	-1.26	-1.54
	(0.03)	*(0.05)*	*(0.02)*	*(0.02)*
Big flows	-0.69	-0.53	-1.00	-1.03
	(0.03)	*(0.05)*	*(0.02)*	*(0.02)*
Small flows	-0.57	-0.57	-0.69	-1.00
	(0.02)	*(0.02)*	*(0.03)*	*(0.04)*

Notes: A gravity equation is estimated for the year 2000 with different estimators. Small and big flows refer to the trade flows which are above and below the median, respectively. Standard errors are in parentheses.

Figures

Figure 1: Number of total trade flows and strictly positive ones in the DOTS database

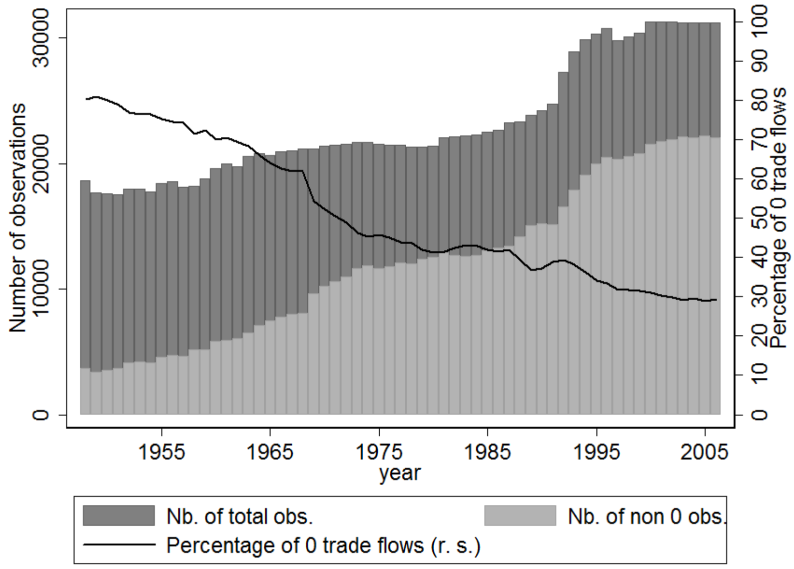


Figure 2: Evolution of trade elasticity to geographic distance : PPML vs log-linear estimates, cross sections

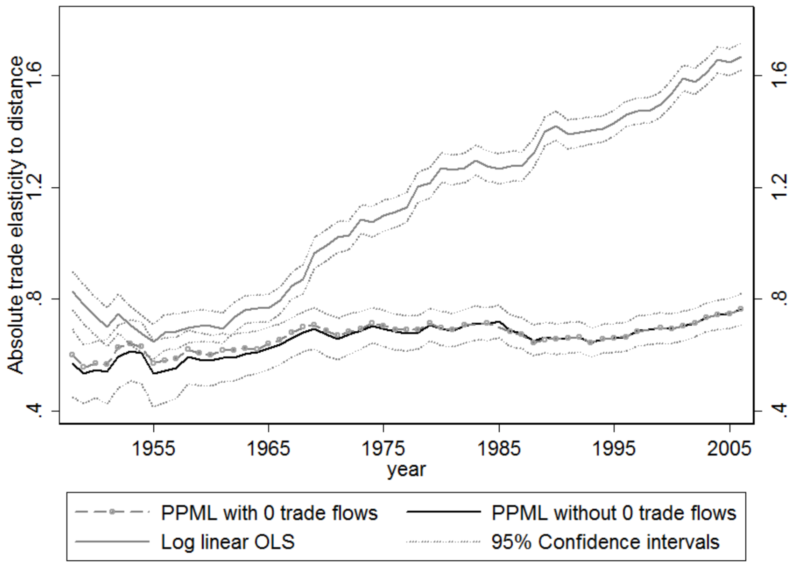
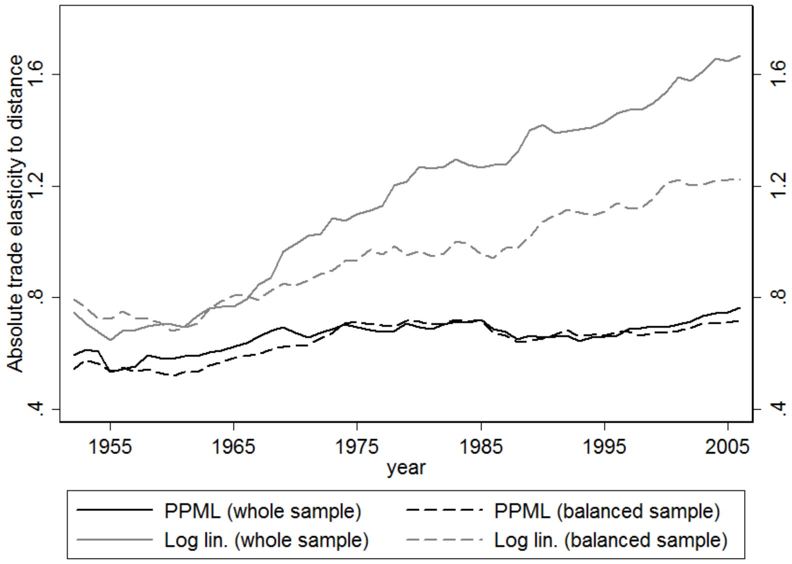
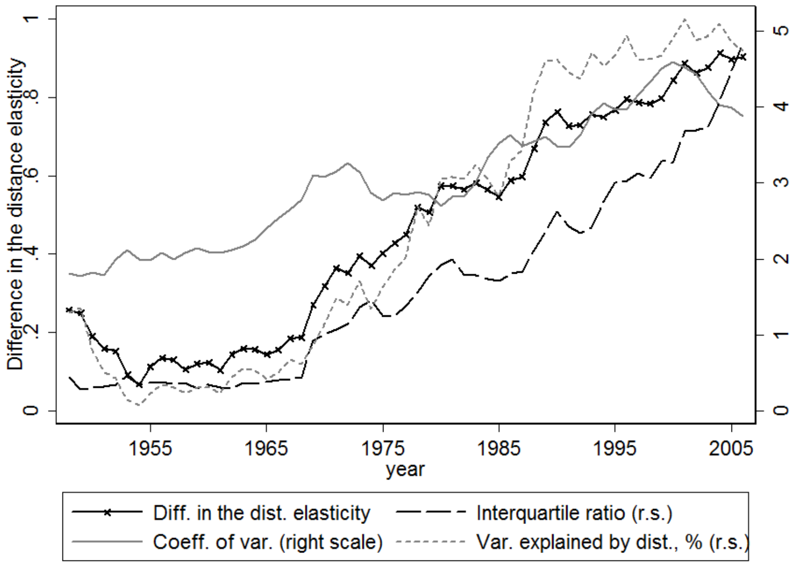


Figure 3: Evolution of trade elasticity to geographic distance : sample analysis



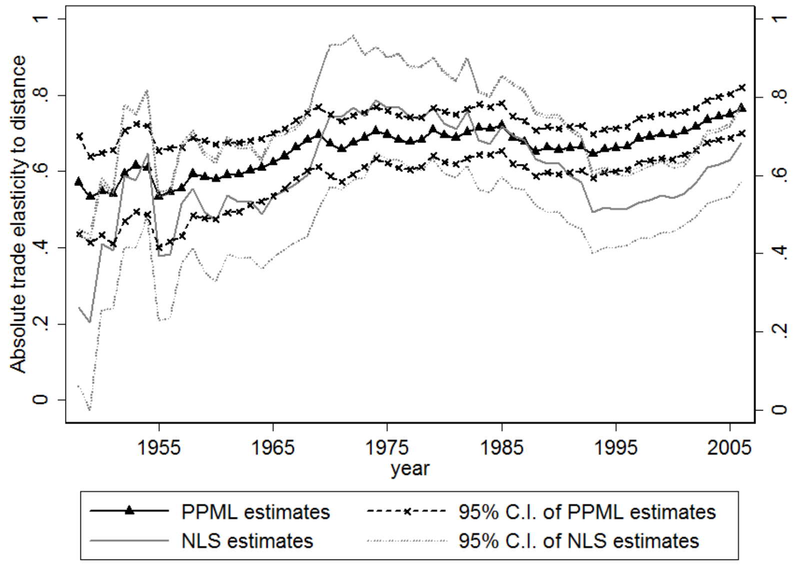
Notes: A gravity equation is estimated for each year both in levels with the PPML estimator and in log with OLS. The balanced sample contains 2550 observations.

Figure 4: Illustration of the heteroskedasticity issue



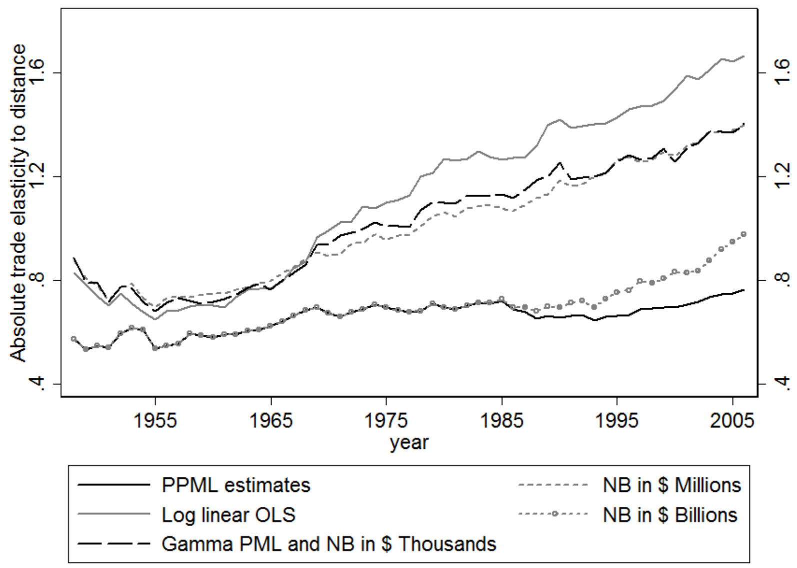
Notes: The difference in the distance elasticity is the gap between Poisson Pseudo-Maximum Likelihood and OLS log estimates of the trade elasticity to distance. The interquartile ratio if the ratio of the third over the first quartile of trade flows. The coefficient of variation is the standard deviation of trade flows divided by the mean. The variance of *log* $\hat{u}$ explained by distance is computed as the difference between the adjusted-R^2 of the regression of *log* $\hat{u}$ on the explicative variables and the adjusted-R^2 of this same regression omitting (the log of) distance as explanatory variable. To fit the right scale, the interquartile ratio and the coefficient of variation have been divided by 70 and 2.5, respectively.

Figure 5: Evolution of trade elasticity to geographic distance : PPML vs NLS estimates, cross sections



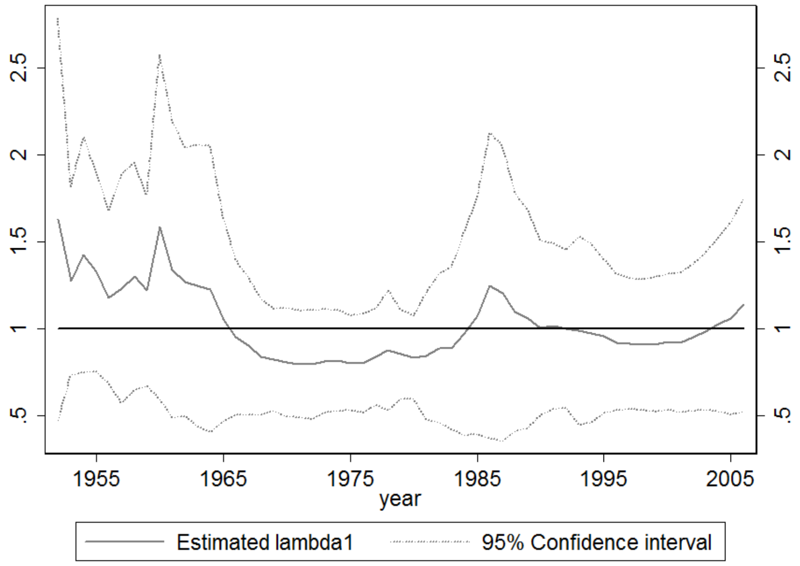
Notes: C.I. = Confidence Interval. A gravity equation is estimated in levels for each year both with Poisson Pseudo-Maximum Likelihood and Nonlinear Least Squares.

Figure 6: Which PML estimator?



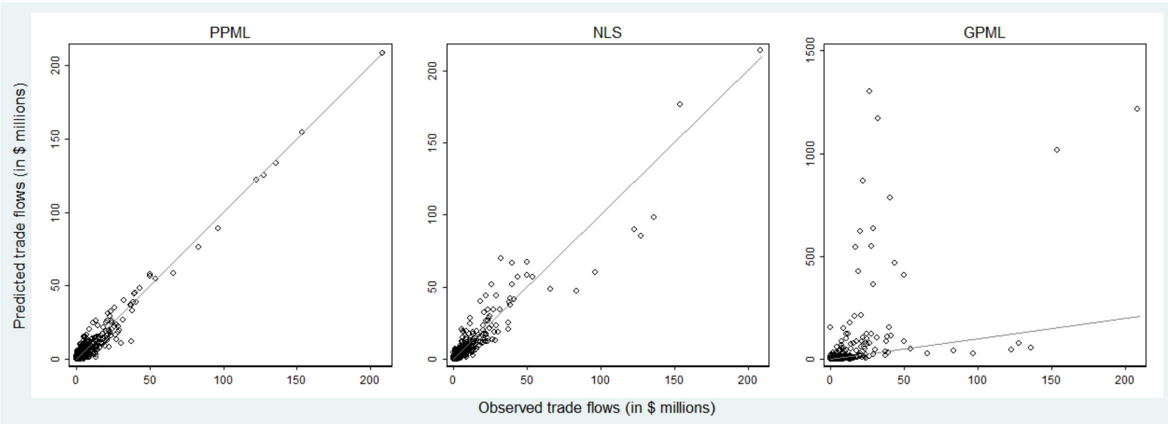
Notes: A gravity equation is estimated for each year with different estimators. In levels, using Poisson Pseudo-Maximum Likelihood, gamma PML and negative binomial PML and, in log using OLS. The negative binomial estimator is computed using three different $ units to measure trade flows. When flows are measured in thousands of $, the negative binomial and the gamma estimators coincide visually.

Figure 7: Estimated λ_1



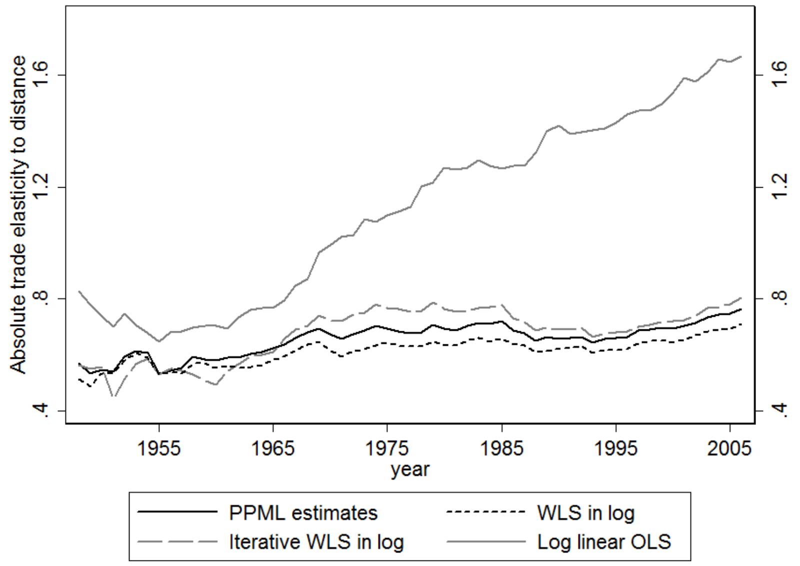
Notes: λ_1 is estimated for each year with equation (6) using PPML for $\tilde{x_{ij}}$.

Figure 8: Observed and predicted trade flows using different estimators, 2000
Mind the y-axis scale for GPML!



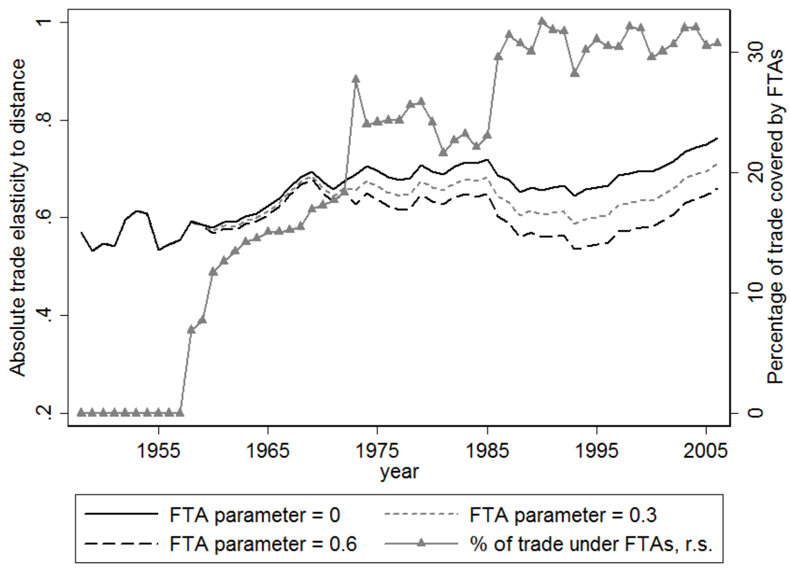
Notes: The straight line is the 45° line.

Figure 9: Log-linear WLS and PPML



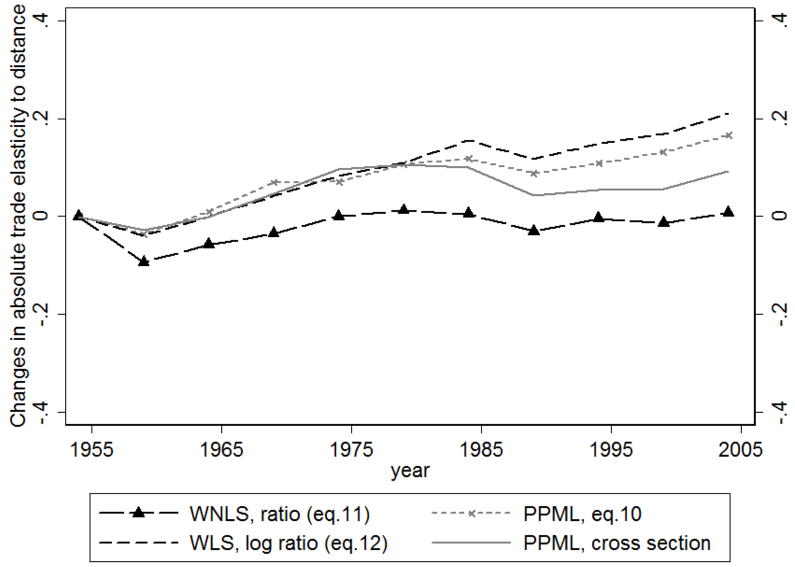
Notes: A gravity equation is estimated for each year with different estimators. In levels with Poisson Pseudo-Maximum Likelihood and in log with OLS, Weighted Least Squares and iterative WLS. WLS uses observed trade flows as weights and iterative WLS uses estimated ones starting with OLS.

Figure 10: Trade elasticity to distance based on different values of the FTA parameter, cross sections



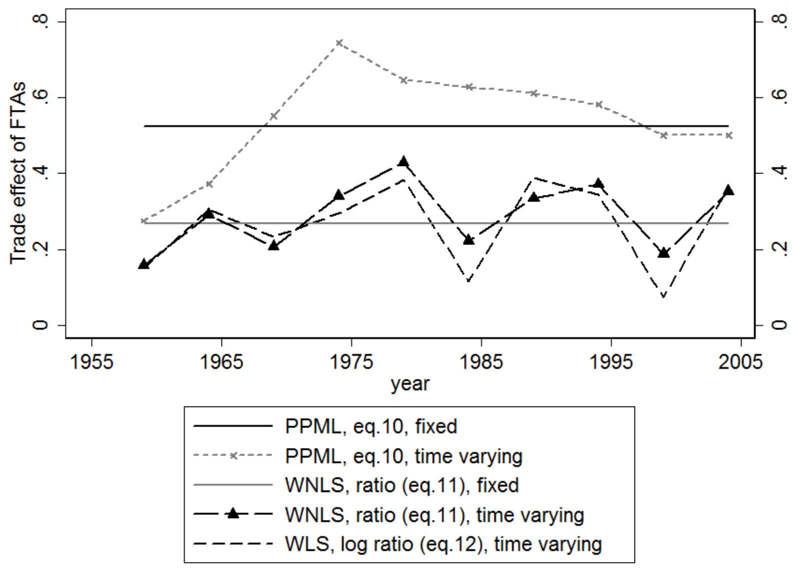
Notes: A gravity equation is estimated for each year in levels with the Poisson Pseudo-Maximum Likelihood estimator. The parameter associated to FTAs is constrained to different values.

Figure 11: Evolution of the trade elasticity to distance (first period, 1952-1956, = 0)



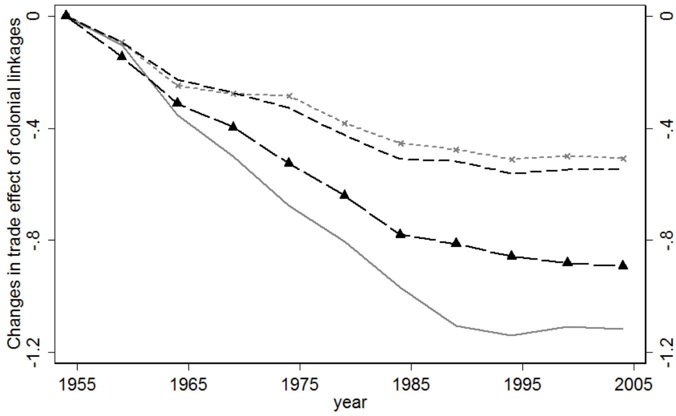
Notes: Except for the Poisson Pseudo-Maximum Likelihood estimates in cross section where a gravity equation is estimated for each year, the gravity equation is estimated in panel over the whole period with different estimators. The weights of the Weighted Least Squares are the trade flows of previous period. The PPML cross section estimates is the one with the FTA parameter constrained to 0.3.

Figure 12: Trade effect of FTAs, panel estimates

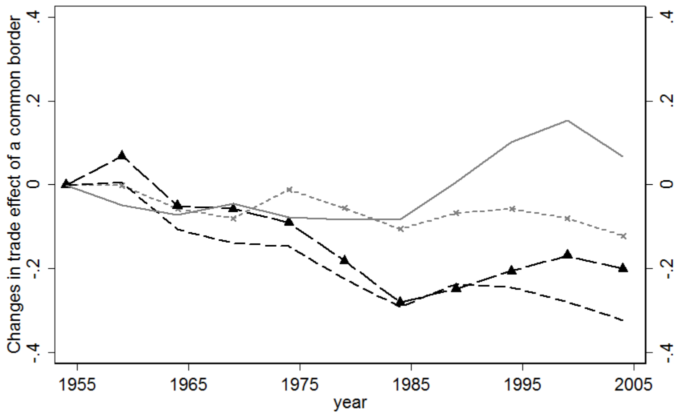


Notes: A gravity equation is estimated in panel over the whole period with different estimators. The FTA parameter is supposed to be either fixed through time or time-varying. The weights of the Weighted Least Squares are the trade flows of previous period.

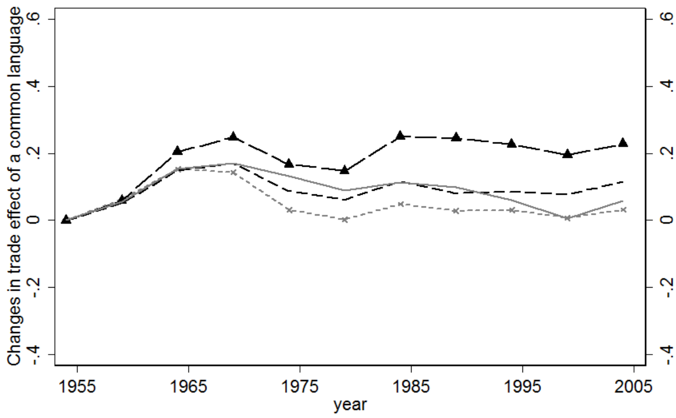
Figure 13: Evolution of other trade determinants (first period, 1952-1956, = 0)



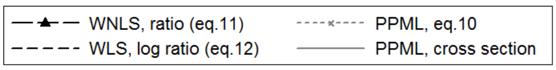
(a) Evol. of the trade effect of colonial linkages



(b) Evol. of the trade effect of a common border



(c) Evol. of the trade effect of a common official language



Notes: Except for the Poisson Pseudo-Maximum Likelihood estimates in cross section, where a gravity equation is estimated for each year, a gravity equation is estimated in panel over the whole period with different estimators. The weights of the Weighted Least Squares are the trade flows of previous period. The PPML cross section estimates is the one with the FTA parameter constrained to 0.3.