Università degli Studi del Molise
Facoltà di Economia
Dipartimento di Scienze Economiche, Gestionali e Sociali
Via De Sanctis, I-86100 Campobasso (Italy)

ECONOMICS & STATISTICS DISCUSSION PAPER
No. 35/07

Reciprocal trade agreements in gravity models: A meta-analysis

by

Maria Cipollina
University of Molise, Dept. SEGeS

and

Luca Salvatici
University of Molise, Dept. SEGeS

Reciprocal trade agreements in gravity models: a meta-analysis

Maria Cipollina and Luca Salvatici*

(University of Molise)

Abstract

Over the time a large number of reciprocal preferential trade agreements (RTAs) have been concluded among countries. Recently many studies have used gravity equations in order to estimate the effect of RTAs on trade flows between partners. These studies report very different estimates, since they differ greatly in data sets, sample sizes, and independent variables used in the analysis. So, what is the "true" impact of RTAs? This paper combines, explains, and summarizes a large number of results (1460 estimates included in 75 papers), using a meta-analysis (MA) approach. Notwithstanding quite an high variability, studies consistently find a positive RTAs impact on bilateral trade: the hypothesis that there is no effect of trade agreements on trade is easily and robustly rejected at standard significance levels. We provide pooled estimates, obtained from fixed and random effects models, of the increase in bilateral trade due to RTAs. Finally, information collected on each estimate allows us to test the sensitivity of the results to alternative specifications and differences in the control variables considered.

JEL classification: C10; F10.
Keywords: Free Trade Agreements; Gravity equation; Meta-regression analysis; Publication bias.

* This work was in part financially supported by the "Agricultural Trade Agreements (TRADEAG)" project, funded by the European Commission (Specific Targeted Research Project, Contract no. 513666); and in part supported by the Italian Ministry of University and Technological Research ("The new multilateral trade negotiations within the World Trade Organisation (Doha Round): liberalisation prospects and the impact on the Italian economy").

1. Introduction.

Preferential agreements are discriminatory policies entailing trade liberalization with respect to a subset of trading partners. The world trading system is characterized by a wide variety of preferential agreements, which can be broadly categorized into two major types: reciprocal (bilateral), entailing symmetric trade liberalization, and nonreciprocal (unilateral), entailing asymmetric trade liberalization aimed at providing support to the country which gains improved market access without being required to open up its own domestic market. The latter, as it well known, have been widely utilized as an instrument for integrating the developing countries into the world trading system.

Traditionally, reciprocal preferential agreements occurred between geographically contiguous countries with already established trading patterns. However, the configuration of these agreements is presently diverse and becoming increasingly complex with overlapping agreements spanning within and across continents in what Bhagwati calls a “spaghetti bowl” of trade relationships.[1] The world has witnessed a veritable explosion of reciprocal preferential trade agreements (RTAs) in the past 15 years. More than half of world trade now occurs within actual or prospective trading blocs, and nearly every country in the world is a member of one or more agreements (Clarete et al., 2003).

RTAs take many forms. The most common are the free trade area (FTA)—where trade restrictions among member countries are removed, but each member maintains its own trade policies towards nonmembers—and the customs union—a FTA where members adopt a common external trade policy. Deeper forms of integration include a common market—a customs union that also allows for the free movement of factors of production—and economic unions, which involve some degree of harmonization of national economic policies. New RTAs, indeed, place considerable emphasis on liberalisation of services, investments and labour markets, government procurement, strengthening of technological and scientific cooperation, environment, common competition policies or monetary and financial integration.

In the literature there are numerous studies analysing the economic impacts of RTAs. The focus of this paper is on estimates of the effects on trade. RTAs might be expected to increase trade between partners, since cheaper imports within the agreement may replace domestic production – trade creation – or crowd out imports from the rest of the world – trade diversion (Viner, 1950; Meade, 1955). However, in many cases intra-agreement trade flows have been strongly increasing even before the formal signature of the agreement, the increases reflecting the impact of unilateral and multilateral liberalization, as well as the simple fact that agreements may be due to, rather than allow for, growing trade relationships.[2]

The purpose of this review is to use a Meta-Analysis (MA) approach to summarize and analyse the RTAs trade effects estimated in the literature, mostly through gravity models assessing the difference between potential and actual trade flows (see Appendix 1 for details on the agreements considered). The approach takes as individual observations the point estimates of relevant parameters from different studies. An MA can improve the assessment of the parameter describing the RTAs impact by combining all of the estimates, investigating the sensitivity of the overall estimate to variations in underlying assumptions, identifying and filtering out publication bias, and by explaining the diversity in the study results in relation to the heterogeneity of study features through meta-regression analysis (MRA).

In this paper, we firstly consider all point estimates provided in the literature, i.e. including multiple estimates coming from a single study. We test for correlation within and between studies, and estimate meta-regression models using weighted least squares (WLS), checking the robustness and sensitivity of our results. Then, we focus on the effect on bilateral trade of specific trade

[1] As a consequence we decided not to use the term “regional”, which is traditionally used as a convenient shortcut, but is inconsistent with the plethora of agreements linking countries around the globe.

[2] Also in the case of multilateral agreements, recent empirical work (Rose, 2004) does not find significant differences between the trade patterns of countries belonging to the GATT/WTO and those of the.

agreements. Finally, we run a probit regression in order to assess what are the most important factors explaining a positive (and significant) impact of the agreements on bilateral trade flows.

The paper is structured as follows. In Section 2 we briefly review the literature studying the impacts of RTAs on trade, while in Section 3 we present some methodological issues regarding the MA approach. In Section 4 we discuss the explanatory variables and present the econometric results. Finally, Section 5 concludes.

2. The impact on trade of reciprocal preferential trade agreements

Empirical research applies econometric approaches to historical trade data in order to assess the impact of trade agreements on bilateral trade flows. Usually, these approaches use gravity models, based upon Newton's Law of Gravitation, predicting that the variation in volume of trade between two economies increase with their size (proxies are real GDP, population, land area) and decrease with transaction costs measured as bilateral distance, adjacency, cultural similarities (Baldwin, 1994; Eichengreen and Irwin, 1996; Feenstra, 1998; Anderson and van Wincoop, 2003).

The standard formulation expresses the bilateral trade between country i and country j as:

$$\ln T_{ij} = \beta_0 + \beta_1 \ln(Y_i) + \beta_2 \ln(Y_j) + \beta_3 \ln(Dist_{ij}) + \varepsilon_{ij} \tag{1}$$

where T_{ij} is the country pairs' trade flow, $Y_{i(j)}$ indicate GDP or GNP of i and j, $Dist_{ij}$ is the distance between i and j, finally ε_{ij} is the error term. Most applications of the gravity model search for evidence of actual or potential effects by adding dummy variables for common languages, for common land borders and for the presence or absence of a RTA. Then, the gravity model is estimated as:

$$\ln T_{ij} = \beta_0 + \beta_1 \ln(Y_i) + \beta_2 \ln(Y_j) + \beta_3 \ln(Dist_{ij}) + \beta_4 Adj_{ij} + \beta_5 Lang_{ij} + \gamma RTA_{ij} + \varepsilon_{ij} \tag{2}$$

where Adj_{ij} is a binary variable assuming the value 1 if i and j share a common land border and 0 otherwise, $Lang_{ij}$ is a binary variable assuming the value 1 if i and j share a common language and 0 otherwise, RTA_{ij} is a binary variable assuming the value 1 if i and j have a reciprocal trade agreement in place and 0 otherwise. A positive coefficient for the RTA variable indicates that it tends to generate more trade among its members. In MA, the parameter of interest (estimate of γ) is commonly referred to as the "effect size".

Many papers find positive and statistically significant RTAs dummies, although they are not primarily interested in estimating the RTA effect, i.e. the existence of an RTA is only included as a control variable. On the other hand, if there is a particular interest on specific RTAs, different dummies may be introduced for each agreement (see Appendix 1).

Some authors distinguish between the increase in the volume of trade within the bloc and the decrease in trade from countries outside the bloc (i.e., trade diversion) by including two dummies for intrabloc and extrabloc trade. An example of a gravity equation that takes in account the trade creation and diversion effects is:

$$\ln T_{ij} = \beta_0 + \beta_1 \ln(Y_i) + \beta_2 \ln(Y_j) + \beta_3 \ln(Dist_{ij}) + \beta_4 Adj_{ij} + \beta_5 Lang_{ij} + \gamma_1 RTA_{kij} + \gamma_2 RTA_{ki-j} \varepsilon_{ij} \tag{3}$$

where RTA_{kij} is a dummy taking value 1 if both i and j are members of bloc k and zero otherwise, and $RTA_{ki\text{-}j}$ is a dummy taking value 1 if i is a member of the bloc but j is not. Accordingly, γ_1 is the coefficient measuring the extent to which trade is influenced by the agreement between i and j (intrabloc trade), and γ_2 is the coefficient associated with extrabloc trade.

Greenaway and Milner (2002) claim that although the impact of any trade agreement is a combination of trade creation and diversion effects, gravity modellers rarely tried to decompose these effects by using dummy variables for members of trade blocs and for non-members, with the expectation of negative coefficients for the latter. Frankel, Stein and Wei (1995) and Frankel and Wei (1997) find evidence of trade creation in European trading blocs from 1970 to 1990, as well as Martìnez-Zarzoso et at (2003), and Mayer and Zignano (2005) for EU and MERCOSUR during the 1990s. Also, Jayasinghe and Sarker (2004) show positive effects for NAFTA in trade of selected agrifood products. Rauch (1996) and Sapir (2001) find negative and significant effect for EFTA.

Other RTAs as LAIA and MERCOSUR appear to have been net trade creating in some studies (Gosh and Yamarik, 2002; Elliott and Ikemoto, 2004; Soloaga and Winters, 2000) and net trade diverting in some others (Carrère, 2006; Krueger, 1999). More recent works (Gosh and Yamarik, 2002; Elliott and Ikemoto, 2004; Cheng and Tsai, 2005; Lee and Park, 2005; Martìnez-Zarzoso and Horsewood, 2005; Carrère, 2006) support the idea that free trade arrangements are generally trade creating.

Recent works investigate on the robustness of the determinants of international of trade by means of extreme-bounds analysis.[3] Ghosh and Yamarik (2004) show that the trade-creating effect is highly sensitive to the choice of other variables included or excluded from the gravity model. Thus, the empirical evidence seems to be rather "fragile". Another work by Baxter and Kouparitsas (2006) tests the robustness of the RTA dummy in gravity equations using three different extreme-bounds approaches. Their analysis gives a mixed view of the relationship between free trade areas and the level of bilateral trade: different methods lead to different outcomes, so results are inconclusive.

The standard estimation method used in gravity equations is the ordinary least squares (OLS). A recent work by Egger (2005) compares four different estimators with respect to their suitability for cross-section gravity models. He recommends a Hausman–Taylor approach that provides consistent parameter estimates, while OLS or the traditional random-effects model are biased.

Most of the articles run regressions from cross-section data either for a single year or for multiple years. Even if panel data allow to pin down the estimates of persistent effects with more accuracy, only very recently gravity equations have been estimated using panel data techniques.

Usually, empirical studies do not take account the endogeneity problem, since countries might enter into a RTA for reasons unobservable to the econometrician and possibly correlated with the level of trade. Baier and Bergstrand (2005) address the endogeneity problem using instrumental variables, Heckman's control-function (CF) techniques (Heckman, 1997), and panel-data estimates. They find that the best method to estimate the effect of RTAs on bilateral trade flows is through differenced panel data, while instrumental variables applied to cross-section data are biased and underestimated.

The Global Economic Prospects (2005) of the World Bank provides a meta-analysis of the literature on the impact of regional trade agreements on intra- and extra-regional trade. It finds that although these agreements typically have a positive impact on intra-regional trade, their overall impact is uncertain. The analysis considers 17 research studies providing 362 estimates of the impact on the level of trade between partners. The mean value of these estimates is positive, but there is a high degree of variance about the mean.

In this study we collect papers that: (1) use gravity models for analysing bilateral trade flows; (2) include dummy variables for the presence of RTAs; (3) estimate coefficients through cross-section or panel analyses.

3. Methodological issues

MA is a set of quantitative techniques for evaluating and combining empirical results from different studies (Rose and Stanley, 2005). The central concern of MA is to test the null hypothesis that different point estimates, treated as individual observations (γ), are equal to zero when the findings from this entire area of research are combined.[4]

[3] Econometric technique testing the robustness of estimated coefficients to changes in a conditioning set of information (Levine and Renelt, 1992).

[4] Under the null hypothesis of no effect ($\gamma = 0$), no publication selection and independence, the statistic minus twice the sum of the logarithms of the p-values is distributed approximately as a χ^2 with $2n$ degrees of freedom (Fisher, 1932).

MA has recently been growing in popularity in economics.[5] Empirical economists have increasing employed meta-analysis methods to summarize regression results particularly in environmental economics (van den Bergh et al, 1997; Florax, 2002, Jeppesen et al 2002), labour economics (Card and Krueger, 1995; Jarrel and Stanley, 1990; Stanley and Jarrel, 1998; Ashenfelter et al., 1999; Longhi et al, 2005; Nijkamp and Poot, 2005; and Weichselbaumer and Winter-Ebmer, 2005), monetary economics (Knell and Stix, 2005) and international trade (Disdier and Head, 2004; Rose and Stanley, 2005).

Although MA is an appealing technique for evaluating and combining empirical results, there is a risk to analyze completely different outcome variables or different explanatory variables (the "Apples and Oranges Problem" as referred to by Glass et al, 1981). In this respect, it is crucial the first step of any MA, namely the construction of a database of estimates. In this application, we only used papers written in English. Papers were selected via extensive search in Google and in databases, such as EconLit and Web of Science. EconLit provides coverage since 1969 to the economics literature including 750 journals. Web of Science provides access to current and retrospective multidisciplinary information from approximately 8700 of the most prestigious, high impact research journals in the world (199 journals in the field of economics), covering the time period from 1992 to the present. With the search in Google, we get papers and working papers that are not published in academic journals. Finally, we traced some specific papers cross-referenced in other works.

The keywords searched for were: "trade agreements", "gravity equation or gravity model" in the title, the abstract or the subject. The first keyword permits to get the papers dealing with trade agreements, while the second keyword sorts out papers using a gravity approach. Among the first group of papers we select the papers analyzing trade agreements focusing on bilateral trade flows; in the second group, we selected those studies including trade agreements as a control variable in the gravity equation.

The final sample includes 75 papers (38 published in academic journals, 37 are working papers or unpublished studies) providing 1460 point estimates of the impact of RTAs on bilateral trade: i.e., the coefficient γ or γ_1 in equations (2) and (3), respectively (see Appendices 2 and 3 for details). In case some agreements changed their nature from "unilateral" to "reciprocal" over time, we did not consider the estimates referring to periods when there were only preferential tariff reductions.

It happens quite often that a study provide multiple estimates of the effect under consideration. The presence of more that one estimated reported per study is problematic, because the assumption that multiple observations from the same study are independent draws becomes too strong. On the other hand, counting all estimates equally would tend to overweight studies with many estimates (Stanley, 2001).

Various solutions have been suggested in the literature. Some authors include a dummy variable (fixed effect) for each study that provided more than one observation (Jarrell and Stanley, 1990), others use a panel specification (Jeppesen et al., 2002, Disdier and Head, 2004). Alternatively, one may decide to represent each study with a single observation, identifying a "preferred" estimate, using averages or medians of the estimates from each paper, or randomly selecting one estimate (Card and Krueger, 1995; Stanley, 2001; and Rose and Stanley, 2005). In this case, though, important information is lost in the grouping process and it is not clear which estimate one should use (Jeppensen et al, 2002).

Pooling different estimates into a large sample for meta-analysis raises the question of within-study versus between study heterogeneity. In order to take this into account, a distinction between a fixed effect (FE) and a random effect (RE) models can be made: the former assume that differences across studies are only due to within-variation; the latter consider both between study and within-study variability, and assume that the studies are a random sample from the universe of all possible studies (Sutton et al., 2000).

[5] In 2005, the Journal of Economic Surveys dedicated a Special Issue (Vol.19, No. 3) to the use of meta-regression analysis.

More specifically, the fixed-effects model assumes that a single, "true" effect ($\hat{\theta}_F$)underlies every study. Following Higgins and Thompson (2002), the $\hat{\theta}_F$ is calculated as a weighted average of the study estimates, using the precisions as weights:

$$\hat{\theta}_F = \frac{\sum_{i=1}^{n} \hat{\theta}_i w_i}{\sum_{i=1}^{n} w_i} \tag{4}$$

where $\hat{\theta}_i$ is the individual estimate of the RTA effect (our γ_i), and the weights w_i are inversely proportional to the square of the standard errors:

$$w_i = \frac{1}{Se(\hat{\theta}_i)^2} \tag{5}$$

So that studies with smaller standard errors have greater weight that studies with larger standard errors.

A field of the literature showing high heterogeneity cannot be summarized by the fixed-effects estimate under the assumption that a single "true" effect underlies every study. As a consequence, the fixed-effects estimator is inconsistent and the random effects model is more appropriate.

The random-effects model assumes that there are real differences between all studies in the magnitude of the effect. Unlike the fixed effects model, the individual studies are not assumed to be estimating a true single effect size, rather the true effects in each study are assumed to have been sampled from a distribution of effects, assumed to be Normal with mean _0 and variance τ^2. The weights incorporate an estimate of the between-study heterogeneity, $\hat{\tau}^2$, so that the random effects estimate ($\hat{\theta}_R$) is equal to (Higgins and Thompson, 2002):

$$\hat{\theta}_R = \frac{\sum_{i=1}^{n} \hat{\theta}_i w^*_i}{\sum_{i=1}^{n} w^*_i} \tag{6},$$

where the weights are equal to:

$$w_i^* = (w_i^{-1} + \hat{\tau}^2)^{-1} \tag{7}.$$

Allowing for the extra between-study variation has the effect of reducing the relative weighting given to the more precise studies. Hence, the random effects model produces a more conservative confidence interval for the pooled effect estimate.

A test of homogeneity of the θ_i is provided by referring the statistic

$$Q = \sum_{i=1}^{n} w\left(\hat{\theta}_i - \hat{\theta}_F\right)^2 \tag{8}.$$

to a χ^2 distribution with $n -1$ degrees of freedom. If Q exceeds the upper-tail critical value, the observed variance in estimated effect sizes is greater than what we would expect by chance if all studies shared the same 'true' parameter (Higgins and Thompson, 2002).[6]

The Q test should be used cautiously, among other things because its power is low (Sutton 2000): when we have a large sample of observations, for example, Q is likely to be rejected even

[6] A moment-based estimate of $\hat{\tau}^2$ may be obtained by (8) equating the observed value of Q with its expectation $E[Q] = \hat{\tau}^2\left(\sum_{i=1}^{n} w_i - (\sum_{i=1}^{n} w_i^2 \Big/ \sum_{i=1}^{n} w_i)\right) - n + 1$ yielding $\hat{\tau}^2 = \dfrac{Q - n + 1}{\sum_{i=1}^{n} w_i - (\sum_{i=1}^{n} w_i^2 \Big/ \sum_{i=1}^{n} w_i)}$.

when the individual effect sizes do not differ much. Anyway, its computation is an intermediate step to compute the preferred tests – H^2 and I^2 – that we are going to use in our analysis.

The statistic H^2 provides a possible measure of the amount of heterogeneity:

$$H^2 = \frac{Q}{n-1} \tag{9}$$

through the ratio of Q over its degrees of freedom. In absence of heterogeneity

$$E[Q] = n-1 \tag{10}$$

so that $H^2 = 1$ indicates homogeneity in effect sizes.

The I^2 statistic, on the other hand, measures the proportion inconsistency in individual studies, that cannot be explained by change:

$$I^2 = \frac{H^2 - 1}{H^2} = \frac{Q - n + 1}{Q} \tag{11}$$

Values close to 1 denote a high degree of heterogeneity. Multiplied by 100, the I^2 statistic describes the percentage of variability in point estimates that is due to heterogeneity rather than sampling error. In the following, we will assign adjectives of low, moderate, and high to values of I^2 lower or equal to 25%, 50%, and 75%. respectively.

The simple mean of estimates could be misleading in presence of more than one mode or outliers in the sample of estimates, because a large part of the estimates may lie to one side of the mean value. If the distribution is multimodal or there are outliers (as extreme data points) the mean could be biased. Skewness is usually tested by comparing mode, median and mean of the distribution. However, this would not be true in the case of symmetrically distributed outliers, since they tend to cancel out each other, or when outliers have smaller statistical weights than other data points so that they contribute less to the mean. In any case, some authors prefer to remove the outliers, since they compress the variation of the rest of the sample and are likely to lead to fragile findings (Disdier and Head, 2004); while others claim that removing outliers and extreme results at an early stage of the meta-analysis could introduce (substantial) bias into the meta-results, and the influence of removing outliers should be explored in a sensitivity analysis (Stanley 2001, Florax 2002).

Finally, there is a general belief that publication bias exists because research with statistically significant, or interesting, results is potentially more likely to be submitted and published than work with null or nonstatistically significant, or uninteresting results. The presence of publication bias may greatly affect the magnitude of the estimated effect. Publication bias occurs when researchers, referees, or editors have a preference for statistically significant results. Some meta-regression and graphical methods can differentiate genuine empirical effect from publication bias (Stanley, 2005).

The simplest and conventional method to detect publication bias is by inspection of a funnel graph diagram. The funnel graph is a scatter diagram presenting a measure of sample size or precision of the estimate on the vertical axis, and the measured effect size on the horizontal axis. The most common way to measure precision is the inverse of the standard error (1/Se). Asymmetry is the mark of publication bias: in the absence of such a bias, the estimates will vary randomly and symmetrically around the true effect. The diagram, then, should resemble an inverted funnel, wide at the bottom for small-sample studies, narrowing as it rises.

A Meta-regression Analysis (MRA) model can be used to investigate and correct publication bias. The model regresses estimated coefficients (γ_i) on their standard errors (Card and Krueger, 1995; Ashenfelter et al 1999):

$$\gamma_i = \beta_1 + \beta_0 Se_i + \varepsilon_i \tag{12}$$

In the absence of publication selection, the magnitude of the reported effect will vary randomly around the 'true' value, β_1, independently of its standard error. Then, when the standard error of the effect of RTA is not significantly different from 0 at any conventional level, the publication bias is not a major issue.[7]

Since the studies in the literature may differ greatly in data sets, sample sizes, independent variables, variances of these estimated coefficients may not be equal. As a consequence, meta-regression errors are likely to be heteroscedastic, but the OLS estimates of the MRA coefficients remain unbiased and consistent.

A weighted least squares (WLS) corrects the MRA for heteroscedasticity, and permits to obtain efficient estimates of equation (12) with correct standard errors. The WLS version of equation (12) is obtained dividing regression equation by the individual estimated standard errors:

$$\frac{\gamma_i}{Se_i} = t_i = \beta_0 + \beta_1 (1/Se_i) + e_i \tag{13}$$

where t_i is the conventional t-value for γ_i, the intercept and slope coefficients are reversed and the independent variable becomes the inverse of Se_i.[8] Thus, the potential for heteroscedasticity causes the meta-analyst to direct his attention towards the reported t-statistics (Stanley and Jarrell, 2005). Equation (13) is the basis for the funnel asymmetry test (FAT), and it may now be estimated by OLS. In the absence of publication selection the magnitude of the reported effect will be independent of its standard error, then β_0 will be zero.

Stanley (2001) proposes a method to remove or circumvent publication selection by using the relationship between a study's standardized effect (its t-value) and its degrees of freedom or sample size n as a means of identifying genuine empirical effect rather than the artefact of publication selection:

$$\ln|t_i| = \alpha_0 + \alpha_1 \ln n_i + v_i \tag{14}$$

When there is some genuine overall empirical effect, statistical power will cause the observed magnitude of the standardized test statistic to vary with n: this method is known as meta-significance testing (MST).

Information on interpretation of meta-regression tests is summarized in Table 1. In the next section we will use these approaches in order to assess genuine empirical effects beyond random and selected misspecification biases.

Table 1: ***MR tests for publication bias and empirical significance***

Test	MRA model	H_1	Implications
Funnel asymmetry *Precision-effect*	$t_i = \beta_0 + \beta_1 (1/Se_i) + e_i$	$\beta_0 \neq 0$ $\beta_1 \neq 0$	*Publication bias* *Genuine empirical effect*
Meta-significance	$\ln|t_i| = \alpha_0 + \alpha_1 \ln n_i + v_i$	$\alpha_1 > 0$	*Genuine empirical effect*
Joint precision-effect/ meta-significance	*Both of the above MRA tests*	$\beta_1 \neq 0$ $\alpha_1 > 0$	*Genuine empirical effect*

Source: Stanley, 2005

4. Meta-analysis regression

The standard meta regression model includes a set of explanatory variables (X) to integrate and explain diverse findings:

$$\gamma_{ji} = \beta_1 + \beta_0 Se_{ji} + \sum_{k=1}^{K} \alpha_k X_{jik} + \varepsilon_{ji} \tag{15}$$

[7] In such a case, the standard error can be omitted from the regression.

[8] Longhi et al. (2006) weight each effect size by the square root of the sample size from which it is estimated. Since there is no relationship between the standard errors of the estimated effect sizes and the sample sizes from which they are estimated, standard errors can still be used as an explanatory variable in the meta-regression in order to correct for publication bias.

where γ_{ji} is the reported estimate i of the jth study in literature, β expresses the true value of the parameter of interest, X_{jik} is the independent variable which measures relevant characteristics of an empirical study and explains its systematic variation from other results in the literature, α_k is the regression coefficient which reflects the biasing effect of particular study characteristics, and ε_{ji} is the disturbance term.

As it was mentioned in the previous section, meta-regression errors are likely to be heteroscedastic. Accordingly, a common practice in meta-regression analysis is to weigh each effect by some measure of precision of the estimated effect and then explain the heterogeneity in study results by means of a linear regression model estimated with weighted least squares (WLS). Dividing (15) by the standard error of the estimates we get:

$$\frac{\gamma_{ji}}{Se_{ji}} = t_{ji} = \beta_0 + \beta_1(1/Se_{ji}) + \sum_{k=1}^{K}(\alpha_k X_{jik}/Se_{ji}) + e_{ji} \qquad (16).$$

The previous regression may still lead to inefficient, though consistent, estimators, since it does not take into account the dependence among estimates obtained in the same study. In order to get correct standard errors, we adopt a "robust with cluster" procedure, adjusting standard errors for intra-study correlation.[9] Each cluster identifies the study each estimate belongs to: this changes the variance-covariance matrix and the standard errors of the estimators, but not the estimated coefficients themselves.

Finally, we adopt a specification that investigates on factors that determine whether the estimated effects are positive and significantly different from zero. The estimated model is given by:

$$s_{ji} = a + \sum_{k=1}^{K} b_k X_{jik} + e_{ji} \qquad (17)$$

where the dependent variable is a dummy (s) that takes the value 1 if the estimated effect size is positive and statistically significant The probability that an estimated effect size is positive and significant is explained by a set explanatory variables (X) and estimated running a probit regression.

4.1 Explanatory variables

The set of variables X includes dummy variables and specification variables that account for differences in functional forms, types of regression, data definitions or sources, sample size, etc. Accordingly, we add two groups of explanatory variables: the first includes dummies explaining the diversity in the results from a methodological point of view; the second includes dummies regarding structural features of the studies considered.

From a methodological point of view, we distinguish estimates obtained by cross-sectional regression models and by panel-data regression techniques; estimates obtained by pooled data; and estimates obtained with ordinary least squares (OLS) estimator or with other estimators. Point estimates in our sample are obtained by studies based on cross-section data, or by studies using pooled cross-section time series or panel data. Panel data regression techniques are preferable to cross-sectional and pooled regression models, since the latter may be affected by an omitted variables bias, or they may be biased by the exclusion or mismeasurement of trading pair–specific variables (Baldwin, 2006). The dummy *cross-section* is equal to "1" if cross-sectional data are used and "0" otherwise; the dummy *pooled* is equal to "1" if panel data are pooled and treated in a cross-sectional regression model and "0" otherwise.[10] Regarding the estimation method the dummy *ols* is equal to "1" if estimates are obtained through simple OLS and "0" whether estimates are obtained with other models (instrumental variable, Hausman-Taylor, etc.).

[9] The "robust" specification adopts the Huber/White/sandwich estimator of variance in place of the traditional one. Some authors (Jeppensen et al, 2002; Disdier and Head, 2004) adopt a panel specification, but such a choice seems questionable: since any ordering of estimates is arbitrary, the data do not form a proper panel.

[10] To avoid collinearity problems we do not include an additional dummy variable for panel studies.

As far as the structural characteristics are concerned, some studies do not specify the type of agreement, while others include different dummies for each agreement. Accordingly, the dummy *agreement* takes the value "1" if the original paper used a variable for each type of agreement, and "0" otherwise. An important aspect to be considered is that several agreements in our dataset are more than simple free trade areas, for example EU, CARICOM, MERCOSUR, CACM, CISCU are customs unions,[11] so that the integration between partners is deeper and, intuitively, their impact on trade should be higher. In order to account for that, we add a dummy *cu* taking values "1" for customs unions and "0" otherwise.

Studies use different measures of bilateral trade flows. Even though the most frequently used measure is the average of bilateral trade, namely the average of the two-way exports,[12] some studies focus on directional trade using only data on bilateral import or export flows. These differences are accounted for adding the dummy *trade flow* equal to "1" when the it refers to average bilateral trade flows, and equal to "0" when trade flows refer to imports or exports.

We expect to find different effect sizes in studies that use sectoral data, since they may refer to sectors that are politically sensitive or more open to international trade. Accordingly, a dummy *sector* is added for estimates that assess the RTA effect on international trade using sectoral data. Since our database includes both published and unpublished studies, we use a dummy called *published* equal to "1" if papers are published in academic journals and 0 otherwise. We also distinguish papers that are primarily interested in estimating the RTA effect, from those including it as a simple control variable. The dummy used for estimates of these papers is called *interested.* We expect higher RTA effect both for published paper, since the choice of the reported estimates may be greatly affected by the preference of researchers for positive and larger results, and for paper that are primarily interested in estimating the RTA effect. Finally, an important aspect relates to the time period analyzed, so we use dummies for four periods: before 1970, the 70s, the 80s, and the 90s: these dummies collect the studies using data explicitly referring to different decades.

4.2 Econometric results

- Sample of 75 estimates

Using a single observation raises the question of which one. Some meta-analyses identify a "preferred" estimate: while they try to choose the preferred/representative estimates to match the intentions, they do choose them (Card and Krueger, 1995; Rose and Stanley, 2005). Others use averages or medians of the estimates from each paper, or select a single measure either randomly or using a more objective statistical procedure, such as the highest R^2 for the corresponding regression (Disdier and Head, 2004).

Bijmolt and Pieters (2001) show that the procedures using a single value for each study generate misleading results. Indeed, if we look at the fixed and random effects estimates based on study's minimum, median and maximum estimate of γ, we obtain very different results (Table 2).

[11] EU (European Union), CARICOM (Caribbean Community and Common Market), MERCOSUR (Southern Cone Common Market), CACM (Central American Common Market), CISCU (Commonwealth of Independent States Customs Union).

[12] Gravity models are usually estimated in log form, but most researchers commit a simple, but grave, error taking the log of the average with very unbalanced trade: the sum of the logs, as a metter of fact, is approximately the log of the sums, but the approximation gets worse as the two flows summed become increasingly different (Baldwin, 2006).

Table 2: Sensitivity of the choice of "preferred" estimate

		Pooled Estimate	Lower Bound of 95% CI	Upper Bound of 95% CI	p-value for H_0: no effect	test Q (p-value)	H^2	I^2
Min	**Fixed-effects**	0.00	-0.01	-0.00	0.55	0.00	54.73	98%
	Random-effects	0.16	0.08	0.23	0.00			
Median	**Fixed-effects**	0.21	0.19	0.22	0.00	0.00	28.22	96%
	Random-effects	0.51	0.43	0.59	0.00			
Max	**Fixed-effects**	0.34	0.33	0.36	0.00	0.00	67.66	98%
	Random-effects	1.41	1.21	1.59	0.00			

In all cases we reject the null hypothesis of homogeneity among estimates (Q-test: p=0.00) and the tests H^2 and I^2 confirm the results of Q test. Apparently, pooled estimates are decreasing as one moves towards the lower percentiles within studies.

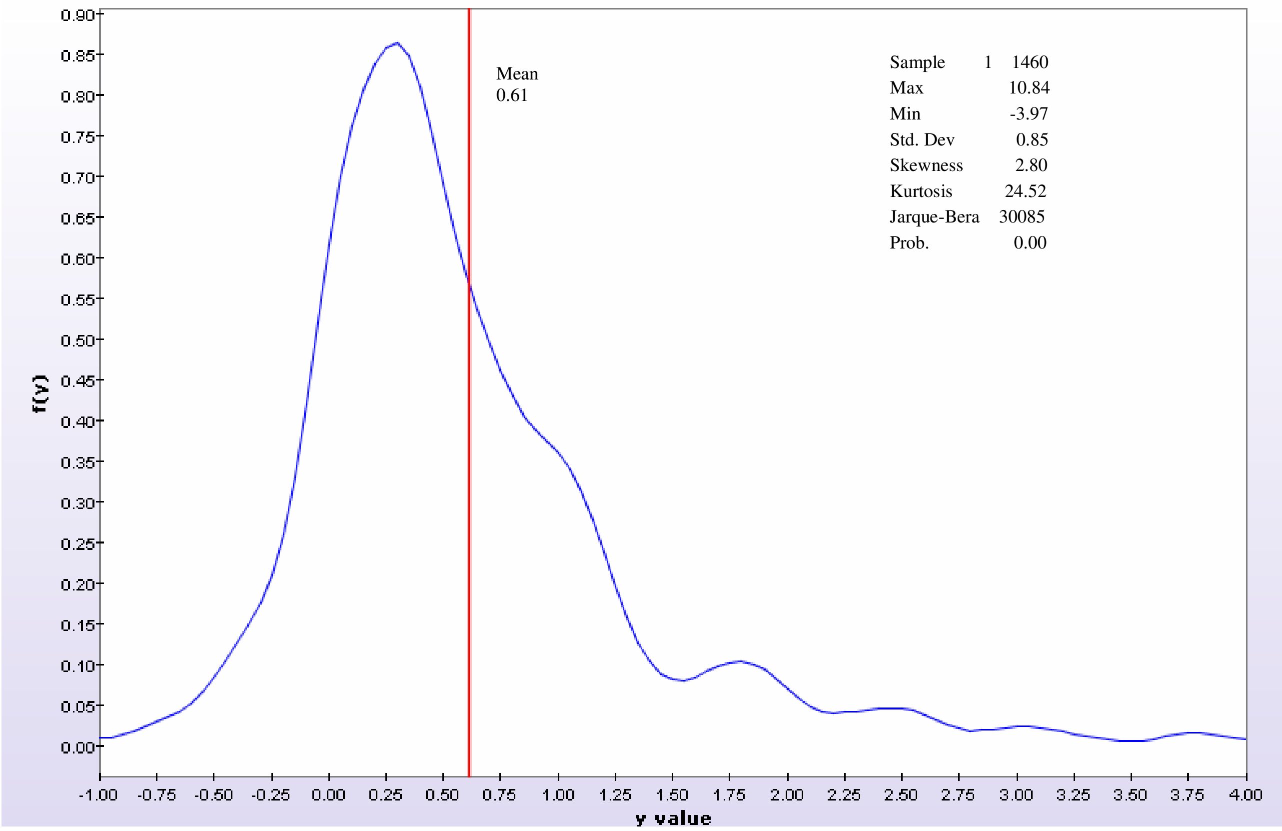
Moreover, 28 out of 75 minimum estimates are negative, and the pooled fixed effects estimate does not reject the hypothesis of no effect. All other estimates and their confidence bounds are positive and strongly reject the null hypothesis. The lowest estimate (minimum estimates – random effects) implies an increase in trade of 17%, while the highest estimate (maximum estimates – random effects) would be larger than 300%!

Given these results, and considering that we would lose valuable information from studies that estimate gravity equations for multiple years, in the following we include in the meta-regression the largest possible number of available observations.

- Sample of 1460 individual estimates

Our database consists of 1460 effect sizes collected from 75 papers estimating the effect of RTAs on international trade. Figure 1 provides the kernel density estimates of the effect sizes. The mean RTA effect (vertical line) is 0.61 and the median is 0.42. These simple statistics do not make use of any information on the precision of each estimate. However, if we combine these effect sizes to test the null hypothesis that $\gamma = 0$, the F-test shows that this hypothesis is rejected at any standard significance level (prob. F-statistic = 0.00).

Figure 1: ***Distribution of RTA effects (γ).***



The estimated trade coefficients range from -3.97 to 10.84, though the majority of coefficients are clustered between zero and one. We employ the Grubbs test in order to detect the existence of outliers (Disdier and Head, 2004), finding 19 extreme values. However, the removal of these values does not significantly affect the results: in the following, then, we present the results obtained from the largest sample.

The distribution in the Figure 1 is clearly lopsided, because few estimates (195 out of 1460) report negative RTAs effects. The values are not symmetrically distributed, with a longer tail to the right than to the left, and the distribution appears to be positively skewed. This is certainly not surprising, since economic theory predicts a positive impact of RTAs on trade.

Table 3 shows combined meta-estimates of γ along with the p-values associated with the test of H_0 of no effect and the test statistic Q, the tests H^2 and I^2. Differences between fixed and random effects estimates confirm a considerable heterogeneity among estimates.

Table 3: ***Meta-Analysis of 1460 estimates of RTAs effect on trade***

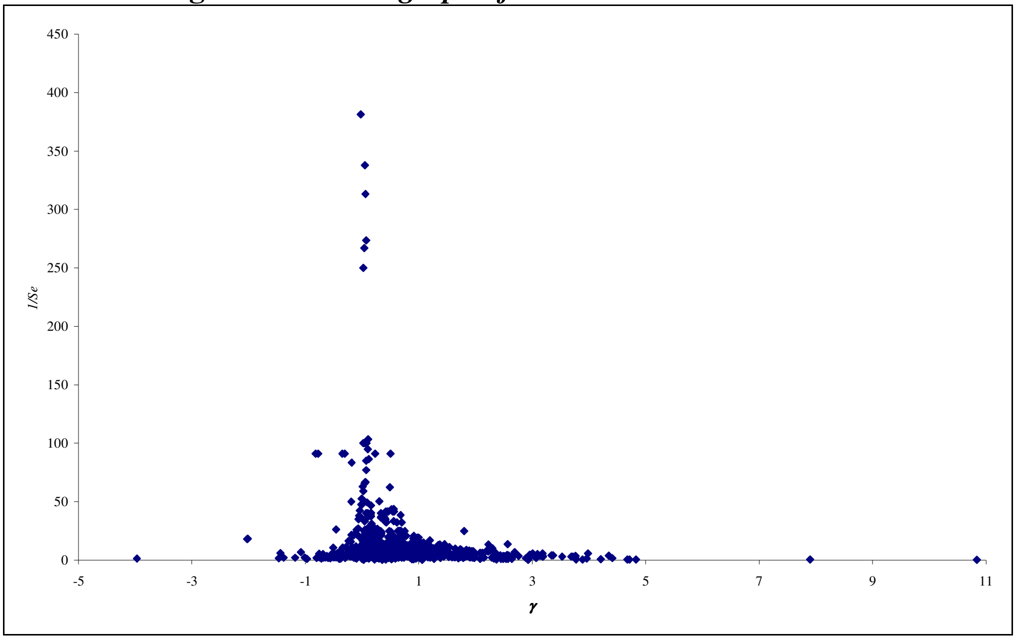
	Pooled Estimate	Lower Bound of 95% CI	Upper Bound of 95% CI	p-value for H_0: no effect	test statistic Q (p-value)	H^2	I^2
Fixed-effects	0.10	0.094	0.098	0.00	0.000	51.41	98%
Random-effects	0.54	0.526	0.562	0.00			

The hypothesis of no common RTA effect is easily rejected (p-value < 0.01), signalling the existence of a RTA effect on bilateral trade. The smaller fixed-effects estimate indicates that RTAs raise trade by 10% ($e^{0.1}-1=0.1$), while the random-effects estimate indicates an increase by 72% ($e^{0.54}-1=0.72$). Differences between fixed and random effects estimates show a considerable heterogeneity among estimates, also confirmed by the tests Q, H^2 and I^2.

In Appendix 4 are reported both fixed and random effects estimates of γ for each of the 75 studies along with the p-values associated with the significance test, as well as the tests for homogeneity within each study. For most of the studies null hypothesis of no effect is easily rejected at any standard significance level. The fixed and random effects estimators do not differ greatly in magnitude, but heterogeneity seems to be present within most of the individual studies. Since studies such a large heterogeneity cannot be summarized by the fixed-effects estimate, the random-effects estimate is to be considered more appropriate.

Following Stanley (2005), we look for publication bias in our sample of disparate effects sizes. A simple way of identifying publication bias is to plot a funnel graph.

Figure 2*: Funnel graph of 1460 individual estimates*



Even though the graph in Figure 2 slightly resembles a funnel, it does not present the symmetry that is crucial to exclude publication bias. Estimates of RTAs effects seem to indicate a positive effect on trade, but Figure 2 clearly shows that the plot is overweighted on the right side. Then publication selection assumes a particular direction.

The six different estimates with the smallest standard errors do not differ significantly from each other. The average of the top six points on the graph, that is the estimates associated with the smallest standard errors, is equal to 0.04, implying a 4.1% increase in trade. Consequently, if research reporting was unbiased, estimates should vary randomly and symmetrically around the value 0.04, whereas the simple average of all 1460 estimates is 0.61, implying a 84% increase in trade.

Table 4 reports the result of the MRA tests. Robust ordinary least squares estimation is used and standard errors are recorded in parenthesis. Both of these MRA test find evidence of a publication bias and an genuine empirical effect ($\beta_0 \neq 0$, $\beta_1 \neq 0$ and $\alpha_1 > 0$). The estimate of β_0 reported in the first column of Table 4 is significant and confirms the apparent asymmetry of the funnel graph (Figure 2). Table 4 shows signs of a positive RTA effect (β_1, $\alpha_1 > 0$ and statistically significant): thus, we can say that MST finds evidence of a genuine empirical effect.

Table 4: ***MRA tests of Effect and Publication Bias***

Variables		Dependent Variables	
		1: t	*2:ln |t|*
β_0:	*intercept*	3.92* (0.18)	-1.22* (0.15)
β_1:	$1/Se$	0.02* (0.01)	-
α_1:	$Ln(n)$	-	0.27* (0.02)
Obs		1460	1289
R-squared		0.01	0.15
S.E. of regression		6.29	1.20

Column 2: studies not reporting the number of observations are excluded
*Standard errors are reported in parenthesis – *: significant at 1 percent.*

Table 5 reports the meta-regression results for equation (16). Column 1 presents the estimated coefficients, while in columns 2 and 3 we report the standard errors and the standard errors adjusted for 75 studies/clusters, in order to show how the significance of some explanatory variables changes when we account for the dependence of estimates within each study. Results show an authentic common RTA effect on trade with a positive and significant impact on trade exceeding 30%.

We find a significant and negative coefficients on the dummies for period ranges (except for the *1970s*). The effect size is much smaller before 1970, while the most recent studies seem to get higher estimates. Such a result is consistent with the often noted evolution from ‘shallow’ to ‘deep’ regional integration agreements, where the latter reduce trade costs through behind-the-border reforms.

Regarding the variables introduced to explain the diversity of estimates from a methodological point of view, coefficients for the *ols*, *cross-section* and *pooled* dummies are not significant using “robust standard errors”. The fact that these estimates are not significantly different from the others seems to suggest that some of the “medal” mistakes of gravity models (Baldwin, 2006) – such as the omitted terms correlated with the trade-cost term – may not be so serious at the end of the day, or they are (partially, at least) handled through the inclusion of origin-nation and destination-nation dummies in a cross-section framework.

Regarding the variables related to study characteristics, we find a negative and highly significant coefficient for the *agreement* dummy, suggesting that studies focusing on specific RTAs find lower values for the estimated γ. This means that estimation problems do not cancel out when all the RTAs are lumped together, rather they make the overestimation bias larger.

The use of average bilateral trade flows leads to a significantly lower RTAs effect. The interpretation of this result is not straightforward. On the one hand, the negative sign could be explained by the smoothing effect due to the computation of averages. On the other hand, Baldwin (2006) pointed out a systematic mistake in the literature, namely the confusion between the log of the average and the average of the logs: such an error always make the bilateral trade look bigger. Since the consequences of the error are large only with very unbalanced trade, it may be the case that it is not particularly relevant for members of RTAs that often have bilaterally balanced trade.

The coefficients regarding customs union and sectoral data are not statistically significant, so it is not confirmed the expectation that custom unions should imply a deeper integration. On the other hand, the dummies *published* and *interested* are positive and, especially in the case of the latter, strongly significant. This suggests the existence of another form of publication bias, since authors primarily interested in estimating the RTA effect tend to report larger effect sizes.

Table 5: ***Multivariate Meta-Regression Analysis (MRA) of Common RTAs Effects***

Variables	*Coefficient*	*Standard Errors*	*Robust with Cluster Standard Errors*
Intercept	3.44	0.21***	0.47***
1/Se$_i$	0.27	0.03***	0.13**
Before 1970	-0.39	0.08***	0.12***
1970s	-0.16	0.10	0.14
1980s	-0.31	0.08***	0.10***
After 1990	-0.25	0.03***	0.04***
Cross-section	0.08	0.03***	0.09
Pooled	-0.14	0.02***	0.14
Ols	0.10	0.02***	0.09
Agreement	-0.25	0.03***	0.10***
Cu	-0.01	0.03	0.13
Trade flow	-0.19	0.03***	0.08***
Sector	0.07	0.08	0.17
Published	0.06	0.03**	0.07
Interested	0.19	0.03***	0.06***
Obs		1460	1460
No of Clusters		-	75
R-squared		0.17	0.17
Prob > F		0.00	0.00
S.E. of regression		5.77	5.77

:significant at 10%; **: significant at 5%; *: significant at 1%; Moderator variables are divided by Se.*

- Focus on single RTAs.

38 studies out of 75 estimate the RTAs impact on trade introducing different dummies for each trade agreement, yielding 1015 estimates. Table 6 summarizes the main results obtained for each RTAs.

Table 6: ***Descriptive statistics of estimates of single RTAs***

Variable |	Obs	Mean	Std. Dev.	Min	Max
RTAs	445	0.69	0.65	-3.97	4.83
ADEAN	2	0.26	0.05	0.22	0.29
AIFTA	10	0.06	0.04	-0.01	0.10
ANZER	10	0.83	1.23	-0.16	3.98
AseanFTA	34	0.97	0.72	-0.07	2.35
BFTA	24	2.96	0.43	2.37	3.77
CACM	20	0.94	0.55	0.12	2.22
CAN	2	1.19	0.00	1.19	1.19
CARICOM	16	1.64	1.25	-0.35	4.41
CEFTA	57	0.41	0.36	-0.51	1.52
CISCU	6	2.66	0.60	1.98	3.37
CUFTA	5	0.39	1.09	-0.35	2.29
EFTA	274	0.25	0.50	-1.38	1.93
EU	401	0.41	0.65	-2.03	4.71
LAFTA	5	0.98	0.92	0.30	2.57
LAIA	9	0.53	0.12	0.39	0.82
MERCOSUR	40	0.64	0.70	0.12	4.35
NAFTA	93	1.16	1.63	-1.47	10.84
US-ISRAEL	7	0.46	0.61	-0.08	1.64

The largest number of observations refers to generic RTAs, since the original studies do not specify the name of the trade agreement. Also for EU there are a large number of observations. Manifestly, the range between minimum and maximum estimates are very large for the most of agreements, showing the large variety of estimates provided by the literature.

We carry out MAs of the RTAs for which we were able to collect effect sizes (Table 7). The tests show that random effects estimates would be the most appropriate in most of the cases. Only 5 out of the 19 agreements do not show significant differences between fixed and random effects estimates (in bold in the table), and most of these cases are characterized by a fairly low number of observations.

The largest effect is registered for the Baltics-RTA (BFTA): the fixed effects estimate suggests an increase in trade around 2000%! Other agreements presenting exceedingly high estimates are the CISCU (1570%) and the Caribbean Community (376%).

Table 7: ***Meta-Analysis of estimates of specific RTAs***

RTA		Pooled Estimate	Variation in Trade (%)	Lower Bound of 95% CI	Upper Bound of 95% CI	test Q (p-value)	H^2	I^2	No. of Estimates
ADEAN	**Fixed**	**0.23**	**26%**	**-0.01**	**0.47**	**0.84**	**0.04**	**0%**	**2**
	Random	**0.23**	**26%**	**-0.01**	**0.47**				
CAN	**Fixed**	**1.19**	**228%**	**0.98**	**1.40**	**0.98**	**0.00**	**0%**	**2**
	Random	**1.19**	**228%**	**0.98**	**1.40**				
CUFTA	Fixed	-0.35	-29%	-0.37	-0.33	0.00	8.72	89%	5
	Random	0.32	38%	-0.43	1.07				
LAFTA	Fixed	1.14	212%	1.07	1.21	0.00	133.70	99%	5
	Random	0.98	167%	0.16	1.81				
CISCU	Fixed	2.94	1783%	2.69	3.19	0.02	2.61	62%	6
	Random	2.82	1570%	2.38	3.26				
US-ISRAEL	Fixed	0.78	118%	0.65	0.91	0.00	22.39	96%	7
	Random	0.47	59%	-0.24	1.17				
LAIA	**Fixed**	**0.52**	**68%**	**0.47**	**0.57**	**0.13**	**1.58**	**37%**	**9**
	Random	**0.52**	**69%**	**0.45**	**0.60**				
AIFTA	**Fixed**	**0.07**	**7%**	**0.05**	**0.08**	**0.18**	**1.40**	**29%**	**10**
	Random	**0.07**	**7%**	**0.05**	**0.09**				
ANZCER	Fixed	0.26	30%	0.19	0.33	0.00	72.53	99%	10
	Random	0.84	130%	0.00	1.67				
CARICOM	Fixed	0.78	117%	0.69	0.86	0.00	28.47	96%	16
	Random	1.56	376%	1.07	2.06				
CACM	Fixed	0.88	140%	0.81	0.94	0.00	11.33	91%	20
	Random	0.95	157%	0.72	1.18				
BFTA	**Fixed**	**3.03**	**1957%**	**2.92**	**3.14**	**0.41**	**1.57**	**36%**	**24**
	Random	**3.06**	**2011%**	**2.91**	**3.21**				
AseanFTA	Fixed	0.79	119%	0.75	0.82	0.00	30.12	97%	34
	Random	0.95	158%	0.74	1.17				
MERCOSUR	Fixed	0.35	42%	0.33	0.38	0.00	13.56	93%	40
	Random	0.58	77%	0.49	0.67				
CEFTA	Fixed	0.26	30%	0.24	0.28	0.00	13.95	93%	57
	Random	0.40	49%	0.30	0.50				
NAFTA	Fixed	0.87	138%	0.83	0.87	0.00	18.54	95%	93
	Random	0.96	159%	0.75	0.96				
EFTA	Fixed	0.05	5%	0.04	0.05	0.00	15.36	93%	274
	Random	0.25	28%	0.21	0.29				
EU	Fixed	0.05	5%	0.05	0.05	0.00	67.50	99%	401
	Random	0.34	41%	0.31	0.37				

Looking at the most widely studied agreements – EU, EFTA and NAFTA –, the largest impact is for NAFTA (159%), while the European agreements register much lower, but possibly more realistic values: 28% in the case of EFTA, 41% for the EU. It is also worth noting that custom unions – EU, CARICOM, MERCOSUR, CACM, CISCU – does not seem to consistently outperform the free trade areas in terms of trade impact. Indeed, in the meta-analysis regression the coefficient of the CU variable was never significant.

4.3 Probit Significance Equation

In our dataset of 1460 effect sizes, 942 are significantly different from zero at the level of 5%, and 887 of these estimates are positive. This is the sample used in the probit estimate (equation 17) with the same set of explanatory variables presented in section 4.1, adjusting the standard errors through the robust with cluster procedure. The computation of the marginal effects at the sample means is shown in Table 8. The value at the mean of the linear combination of the explanatory variables (Z) is -0.7, while the marginal probability of finding a positive and significant impact on trade is 0.3. The estimates indicate that papers studying older agreements (or first stages of implementation) are less likely to detect a positive impact on trade: using data *before 1970*, for instance, reduces the probability by almost 25 percent. This confirms the evidence about the 'deepening' of the liberalization implied by recent RTAs.

Regarding the methodological variables, the probit regression confirms that cross-section estimates do not differ significantly from the other results. On the other hand, the coefficients for pooled data and OLS estimates present opposite signs. Using Pooled data, notwithstanding the "smoothing effect" mentioned above, increases the probability to find a positive impact by more than 11 percent. On the contrary, OLS estimates are less likely to detect an impact. As it was mentioned in section 2, the OLS-estimator may yield biased and inconsistent estimates due to omitted variables and selection bias. Trade between any pair of countries is likely to be influenced by certain unobserved individual effects, if the unobserved effects are correlated with the explanatory variables, coefficients of the latter may be higher because they incorporate these unobserved effects. Indeed, it is very likely that the omitted pro-bilateral trade variables are positively correlated with the 'variable of interest', i.e. the RTA effect. The point is that the decision to join a trade agreement is not random, but rather driven by many factors, including many of the factors omitted from the gravity regression. Concerning the variables describing study characteristics, the use of data on specific agreements or the average of bilateral trade reduces the probability of estimating a positive impact on trade by 37 and 16 percent, respectively. Conversely, but not unexpectedly, published papers and studies explicitly focused on the estimation of the RTA impact on trade are much more likely to register significant and positive results.

Table 8: Probit Analysis

Probit Estimation	Mean	β	Mean* β	$f(Z)$	$\beta f(Z)$
Before 1970	0.077	-0.770**	-0.059	0.321	-0.247
1970s	0.085	-0.360	-0.031	0.321	-0.116
1980s	0.124	-0.630**	-0.078	0.321	-0.202
After 1990	0.455	-0.230	-0.105	0.321	-0.074
Cross-section	0.636	0.030	0.019	0.321	0.010
Pooled	0.177	0.360**	0.064	0.321	0.116
Ols	0.699	-0.190*	-0.133	0.321	-0.061
Agreement	0.697	-1.160***	-0.809	0.321	-0.373
Cu	0.520	0.110	0.057	0.321	0.035
Trade flow	0.478	-0.500***	-0.239	0.321	-0.161
Sector	0.147	-0.040	-0.006	0.321	-0.013
Published	0.538	0.530***	0.285	0.321	0.170
Interested	0.490	0.770***	0.377	0.321	0.247
Intercept	-0.001	1.000***	-0.001		
Total			-0.658		
No. of Obs	1460				
No. of Clusters	75				
Wald $\chi^2_{(13)}$	108				
(p-value)	(0.000)				
Pseudo R^2	0.16				

: significant at 10%; **: significant at 5%; *:significant at 1%.*

5. Conclusion.

RTAs have been widely studied, and the interest on this type of trade liberalization is likely to increase in the next future due to the crisis of the multilateral liberalization process. One way to carry out a comparative study of the empirical results is to simply tabulate authors, country, methodology, and results. However, for policy analysis and a better understanding of the consequences of RTAs, it is useful to complement broad qualitative conclusions with a more precise quantitative research synthesis. This is the purpose of the present paper with respect to one core issue: the impact of these agreement on member countries' bilateral trade flows. In particular, we decided to overcome the main limitations of qualitative reviews, summarizing statistically the whole body of work through meta-regression analysis.

In this paper, we have investigated the result of previous studies analysing the effect of RTAs: the estimated effect varies widely from study to study and sometimes even within the same study. From the methodological point of view, this suggests the opportunity to retain all the available observations in most of our statistical analysis, though considering estimates from the same study as possibly correlated observations. Accordingly, by means of meta-analysis techniques, we statistically summarized 1460 estimates collected from a set of 75 studies.

All combined estimates imply a substantial increase in trade, but they vary a lot depending on the estimation method. In particular, the 'random-effects' estimate entails an increase of 72%. The more modest 'fixed-effects' estimate (10%) cannot be trusted because its basis is undermined by obvious heterogeneity in this literature. However, there is also strong statistical evidence of publication selection, favoring the reporting of significantly positive trade effects: such publication bias causes all simple combined estimates of trade effects, whether fixed- or random-effects, to be exaggerated.

Our analysis also provides a range of additional results helping to explain the wide variation in reported estimates. In this respect, meta-analysis statistical techniques are something more than mere weighted averages of all point estimates. Even if we avoid the temptation to assign "weights" (or "medals", as in Baldwin, 2006) according to which of the studies we deem as good or bad, we can attempt to filter out biases due to the publication selection or possibly questionable methodological choices. For example, estimates obtained from pooled data are more likely to find a positive and significant impact, though they report smaller values. On the other hand, studies reporting OLS estimates are less likely to get (statistically speaking) "good results" and provide results that may be upward biased due to misspecifications and omitted variables. Several studies lump different trade agreements together and employ average trade flows: this has a negative impact on the likelihood of finding significant results, and lead to an underestimation of the impact on trade. Conversely, published papers and studies mainly interested in studying the RTAs' impact are more likely to report significant results that tend to be overestimated.

After filtering out the publication selection and other biases, the meta-analysis confirms a robust, positive RTA effect, equivalent to an increase in trade exceeding 30%. The estimates tend to get larger in recent years, and this could be a consequence of the evolution from 'shallow' to 'deep' trade agreements. Surprisingly, we do not find any evidence of a differentiated impact according to the type of trade agreement. The custom union dummy, as a matter of fact, is never significant, but this may be due to the fact that some custom unions are only putative, since they have not been fully and consistently implemented. Indeed, in the case of a true custom union such as the EU, the MA estimate of the impact on trade (41%) largely exceeds the estimate for all the agreements combined.

The meta-analysis of the trade effects of RTAs provide a combined estimate more plausible than some extreme values reported in the literature. Moreover, our results shed some light on the role played by some research characteristics in explaining the variation in reported estimates. However, our findings should still be considered as provisional, since there remains excess unexplained variation in our meta-regression models.

References

Adam A., Kosma T. S., McHugh J. (2003), *Trade-Liberalization Strategies: What Could Southeastern Europe Learn from the CEFTA and BFTA?*, IMF Working Paper, WP/03/239.

Aikeen (1973), *The effect of the EEC and EFTA on European trade: A temporal Cross-section Analysis*, The America Economic Review, vol. 63, no.5.

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

Ashenfelter O., Harmon C., Oosterbeek H. (1999), *A review of estimates of the schooling/earnings relationship, with tests for publication bias*, Labour Economics, No. 6, pp. 453-470.

Baier, Bergstrand (2005), *Do Free Trade Agreements Actually Increase Members' International Trade?* Federal Reserve Bank of Atlanta, WP 2005-3.

Baldwin R.E. (1994), *Towards an Integrated Europe*, London: Centre for Economic Policy Research.

Baldwin R.E. (2006), *The Euro's Trade Effects*, European Central Bank Working Paper Series, 594.

Baxter M., Kouparitsas M. A . (2006), *What determines bilateral trade flows?*, NBER Working Paper No. 12188.

Bayoumi, Eichengreen (1995), *Is regionalism simply a diversion? Evidence from the evolution of the EC and EFTA*, NBER WP 5283.

Bergstrand J. (1985), *The gravity equation in international trade: some microeconomic foundation and empirical evidence*, The Review of Economics and Statistics 67, 474– 481.

Bergstrand J. H. (1989), *The Generalized Gravity Equation, Monopolistic Competition, and the Factor Proportions Theory in International Trade*, Review of Economic and Statistics 71: 143-53.

Bijmolt, T.H.A. and R.G.M. Pieters, 2001, *Meta-Analysis in Marketing when Studies Contain Multiple Measurements*", Marketing Letters, 12(2): 157-169.

Brada J. C., Mendez J. A. (1985), *Economic integration among developed, developimg and centrally planned economics: a comparative analysis*, Review of Economics and Statistics, 67, pp. 549-56.

Breuss F., Egger P. (1999), *How reliable are estimations of East-West Trade potentials based on Cross-Section Gravity Analyses?*, Empirica, 26, pp. 81-94.

Bun M. J., Klassen F. J.G.M. (2002), *Has the Euro increase Trade?*, Tibergen Institute Discussion Paper 2002-108/2.

Card D., Krueger A.B. (1995), *Time-Series Minimum-Wage Studies: a Meta-analysis*, The American Economic Review, 85(2): 238-243.

Carrère C. (2006), *Revisiting the Effects of Regional Trading Agreements on Trade Flows with Proper Specification of the Gravity Model*, European Economic Review, 50, pp. 223-247.

Cernat L. (2001), *Assensing regional trade arrangements: are south-south RTAs more trade diverting?*, Policy Issues in International Trade and Commodities Study Series No. 16, United Nations Conference on Trade and Development.

Cheng I-H., Tsai Y-Y. (2005), *Estimating the Staged Effects of Regional Economic Integration on Trade Volumes*, Department of Applied Economics, Taiwan, unpublished.

Cheng I-H., Wall H. J. (2004), *Controlling for Heterogeneity in Gravity Models of Trade and Integration*, Working Paper Series 1999-010E, The Federal Reserve Bank of St. Louis.

Clarete R., Edmonds C., Wallack J.S. (2003), *Asian regionalism and its effects on trade in the 1980s and 1990s*, Journal of Asian Economics 14 91–129

Colin J. R. (2005), *Issues in Meta-Regression Analysis: an overview*, Journal of Economic Surveys 19: 295-298

De Benedictis L., De Santis R.,Vicarelli C. (2005), *Hub-and Spoke or else? Free trade agreements in the enlarged EU. A gravity model estimate*, European Network of Economic Policy Research Istitutes, Working Paper No. 37.

Disdier A. C., Head K. (2004), *The Puzzling Persistence of the Distance Effect on Bilateral Trade*, Centro Studi Luca D'Agliano, Development Studies Working Papers No. 186.

Eaton J., Kortum S. (1997), *Technology and Bilateral Trade*, NBER Working Paper No. 6253.

Egger P. (2005), *Alternative Techniques for Estimation of Cross-Section Gravity Models*, Review of International Economics, 13(5), pp. 881–891.

Eichengreen, B., Irwin, D., 1996. The Role of History in Bilateral Trade Flows. NBER Working Paper No. 5565.

Elliot R. J., Ikemoto K. (2004), *AFTA and the Asian Crisis: Help or Hindrance to ASEAN intra-regional trade*?, Asian Economic Journal, vol. 18(1):1-10.

Endoh M. (2000), *The transition of post-war Asia-Pacific trade relations*, Journal of Asian Economics 10, pp. 571–589.

Faruquee H. (2004), *Measuring the Trade Effects of EMU*, IMF Working Paper WP/04/154.

Fazio G., MacDonald R., Mélitz J. (2005), *Trade costs, trade balances and current accounts: an application of gravity to multilateral trade*, CEPR Discussion Paper No. 5137.

Feenstra R. C. (1998), *Integration of Trade and Disintegration of Production in the Global Economy*, Journal of Economic Perspectives, 12, 31-50.

Feenstra, Markusen, Rose (2001), *Using the gravity equation to differentiate among alternative theories of trade*, Canadian Journal of Economics, vol. 34, no.2.

Fidrmuc J., Fidrmuc J. (2003), *Disintegration and trade*, Review of International Economics 11(5): 811–829.

Fisher R.A. (1932), *Statistical Methods for Research Workers*, 4th edn, Oliver andBoyd, London.

Florax, R., de Groot H., de Mooij R. (2002), *Meta-analysis in policy-oriented economic research*, CPB, Report # 1: 21-24.

Frankel J. A., Rose A. K. (2000), *An estimate of the effect of currency unions on trade and output*, CEPR, Discussion Paper No. 2631.

Frankel J. A., Stein E., Wei S-J. (1995), *Trading blocs and the Americas: The natural, the unnatural, and the super-natural*, Journal of Development Economics Vol. 47, pp. 61-95

Frankel J. A., Stein E., Wei S-J. (1997), *Regional Trading Blocs*. Institute for International Economics.

Frankel J. A., Wei (1995), *Open regionalism in a world of continental trade blocs*, NBER Working Paper No.5272.

Frankel J. A., Wei S-J. (1997), *ASEAN in a Regional Perspective*, in J. Hicklin, D. Robinson and A. Singh, eds., Macroeconomic Issues Facing ASEAN Countries. Washington, DC: International Monetary Fund.

Fratianni M., Kang H. (2006), *Heterogeneous distance–elasticities in trade gravity models*, Economics Letters 90, pp. 68–71

Freund C. L., Weinhold D. (2004), *The effect of the Internet on international trade*, Journal of International Economics 62, pp. 171– 189

Ghosh, Yamarik, (2004), *Are regional trading arrangements trade creating? An application of extreme bounds analysis*, Journal of International Economics, 63.

Glass G.V., McGaw B., Lee Smith M. (1981), *Meta-analysis in Social Research*, Sage Publications, Beverly Hills.

Glick R., Rose A. K. (2002), *Does a currency union affect trade? The time series evidence*, European Economic Review 46(6): 1125–1151.

Greenaway D., Milner C. (2002), *Regionalism and Gravity*, Scottish Journal of Political Economy, Vol. 49, No. 5

Grünfeld L. A., Moxnes A. (2003), *The Intangible Globalization. Explaining the Patterns of International Trade in Services*, NUPI Paper No. 657.

Hassan (2001), *Is SAARC a viable economic block? Evidence from gravity model*, Journal of Asian Economics, 12.

Heckman J. J. (1997), *Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations*, Journal of Human Resources 32, No. 3, pp. 441-462.

Hejazi ., Safarian A.E. (2005), *NAFTA effects and the level of development*, Journal of Business Research 58, pp. 1741– 1749

Higgins J. P. T., Thompson S. G. (2002), *Quantifying heterogeneity in a meta-analysis*, Statistics in Medicine 21: 1539–1558.

Jakab Z. M., Kovacs M. A., Oszlay A. (2001), *How Far Has Trade Integration Advanced?: An Analysis of the Actual and Potential Trade of Three Central and Eastern European Countries*, Journal of Comparative Economics 29, pp. 276–292

Jarrell S. B., Stanley T. D. (1990), *A meta-analysis of the union wage gap*, Industrial and Labour Relations Review 44, pp. 54–67

Jayasinghe, Sarker (2004), *Effects of Regional Trade Agreements on Trade in Agrifood Products: Evidence from Gravity Modeling Using Disaggregated Data*, Center for Agricultural and Rural Development Iowa State University, Working Paper 04-WP 374.

Jeppesen, T., List J.A., Folmer H., (2002), *Environmental Regulations and New Plant Location Decisions: Evidence from a Meta-Analysis*, Journal of Regional Science, 42(1): 19-49.

Kein N. T., Hashimoto Y. (2005), *Economic analisys of ASEAN Free Trade Area; by a country panel data*, Discussion Papers in Economics and Business, No. 05-12.

Kenen P. B. (2002), *Currency unions and trade: Variations on themes by Rose and Persson*, RBNZ DP/2002/08.

Kimura F., Lee H-H. (2004), *The Gravity Equation in International trade in Services*, Kangwon National University, Korea, unpublished.

Klein M. W. (2005), *Dollarization and trade*, Journal of International Money and Finance, 24, pp. 935-943.

Knell M., Stix, H. (2005), *The income elasticity of money demand: a meta-analysis of empirical results*, Journal of Economic Surveys 19, pp. 513–533.

Krueger A. (1999), *Trade creation and trade diversion under Nafta*, NBER WP.4729.

Lee J-W., Park I., Shin K. (2004), *Proliferating Regional Trade Arrangements: Why and Whither?*, Korea University, unpublished.

Lee J-W., Park I. (2005), *Free Trade Areas in East Asia: Discriminatory or Non-discriminatory?*, The World Economy, Vol. 28, No.1, pp 21-48.

Levine R., Renelt D. (1992), *A Sensitivity Analysis of Cross-Country Growth Regressions*, American Economic Review, 82 (4), pp. 942–963.

Longhi S., Nijkamp P., Poot J. (2005), *A meta-analytic assessment of the effect of immigration on wages*, Journal of Economic Surveys, Vol. 19, pp. 451–478.

Longhi S., Nijkamp P., Poot J. (2006), *The fallancy of "Job Robbing"*, Tinbergen Istitute Discussion Paper, TI 2006-050/3.

Martìnez-Galàn E., Fontoura M. P., Proenca I. (2005), *Trade Potential in an Enlarged European Union: a recent approach*, ISEG, unpublished.

Martinez-Zarsoso I., Nowak-Lehman F. (2003), *Augmented Gravity Model: An Empirical Application to Mercosur-European Union Trade Flows*, Journal of Applied Economics, Vol. 6, No. 2, pp. 291-316.

Martìnez-Zarzoso I., Horsewood N. (2005), *Regional trading agreements: Dynamic panel data evidence from the gravity model*, unpublished.

Mayer T, Zignago S. (2005), *Market Access in Global and Regional Trade*, CEPII Working Paper No. 2005-02.

Meade J. (1955), *The theory of customs unions* (North-Holland, Amsterdam).

Melitz J. (2001), *Geography, trade and currency union*, CEPR Discussion Paper, No. 2987.

Melitz J. (2002), *Language and foreign trade*, University of Strathclyde working paper, unpublished.

Micco A., Stein E., Ordon˜ ez G. (2003), *The currency union effect on trade: Early evidence from EMU*, Economic Policy, Vol. 37, pp. 316–356.

Nijkamp P., Poot J. (2005), *The last word on the wage curve?*, Journal of Economic Surveys, Vol.19, pp. 421–450.

Nilsson L. (2005), *Comparative effects of EU and US trade policies on developing country exports*, European Commission, TRADE Directorate General, unpublished.

Nitsch V. (2002), *Honey, I shrunk the currency union effect on trade*, World Economy, Vol. 25, No. 4, pp. 457–474.

Paiva C (2005), *Assessing Protectionism and Subsidies in Agriculture: A gravity Approach*, IMF Working Paper, WP/05/21.

Pakko M. R., Wall H. J. (2001), *Reconsidering the trade-creating effects of a currency union*, FRB St. Louis Review, Vol. 83, No. 5, pp. 37–45.

Rauch J. (1996), *Networks versus markets in International Trade*, NBER Working Paper No. 5617.

Rauch J., Trindade V. (1999), *Ethnic Chinese Networks in International Trade*, NBER Working Paper No. 7189.

Rose A. K. (2000), *One money, one market: Estimating the effect of common currencies on trade*, Economic Policy, Vol. 30, pp. 9–45.

Rose A. K. (2001), *Currency unions and trade: The effect is large*, Economic Policy, Vol. 33, pp. 449–461.

Rose A. K. (2003), *Does the WTO Make Trade More Stable?*, NBER Working Paper No. 10207.

Rose A. K. (2004), *Do We Really Know that the WTO Increases Trade?*, American Economic Review, Vol. 94(1).

Rose A. K. (2005), *One reason countries pay their debts: renegotiation and international trade*, Journal of Development Economics, Vol. 77, pp. 189– 206

Rose A. K., Engel C. (2002), *Currency unions and international integration*, Journal of Money, Credit, and Banking, Vol. 34, No. 4, pp. 1067–1089.

Rose A. K., Stanley T. D. (2005), *Meta-analysis of the effect of common currencies on international trade*, Journal of Economic Surveys, Vol. 19, pp. 347–366.

Rose A. K., van Wincoop E. (2001), *National money as a barrier to trade: The real case for monetary union*, American Economic Review, Vol. 91, No. 2, pp. 386–390.

Sanso M., Cuairan R., Sanz F. (1993), *Bilateral Trade Flows, the Gravity Equation, and Functional Form*, The Review of Economics and Statistics, Vol. 75, No. 2, pp. 266-275

Sapir A. (2001), *Domino effects in Western European regional trade, 1960–1992*, European Journal of Political Economy, Vol. 17, pp. 377-388

Silva J. S., Tenreyro S. (2003), *Gravity-Defying Trade*, unpublished.

Silva J. S., Tenreyro S. (2005), *The Log of Gravity*, CEPR Discussion Paper No.5311.

Sissoko A. A. (2004), *Measuring Trade Intensity within the European Zone*, unpublished.

Soloaga, Winters (2001), *Regionalism in the nineties: What effect on trade?*, North American Journal of Economics and Finance, Vol. 12, pp. 1-29

Stanley T.D. (2001), *Wheat From Chaff: Meta-Analysis as Quantitative Literature Review*, Journal of Economic Perspectives, Vol. 15, pp. 131–150

Stanley T. D. (2004), *Meta-regression methods for detecting and estimating empirical effect in the presence of publication selection*, Center for Entrepreneurial Studies Discussion Paper 2004–2, Hendrix College.

Stanley T. D. (2005), *Beyond publication bias*, Journal of Economic Surveys, Vol. 19, pp. 309–346.

Stanley T. D., Jarrell S. B. (1989), *Meta-regression analysis: a quantitative method of literature surveys*, Journal of Economic Surveys Vol. 3, pp. 54–67.

Stanley T. D., Jarrell S. B. (1998), *Gender wage discrimination bias? A meta-regression analysis*, Journal of Human Resources, Vol. 33, pp. 947–973.

Stanley T. D., Jarrell S. B. (2005), *Meta-regression analysis: a quantitative method of literature surveys*, Journal of Economic Surveys, Vol. 19, pp. 299–308.

StataCorp. (2005), Stata Statistical Software: Release 9. College Station, TX: Stata-Corp LP.

Subramanian A., Wei S.-J. (2003), *The WTO promotes trade, strongly but unevenly*, NBER Working Paper No.10024.

Subramanian A., Wei S.-J. (2005), *The WTO promotes trade, strongly but unevenly*, CEPR Discussion Paper No 5122.

Sutton A.J., Abrams K.R., Jones D.R., Sheldon T.A., Song F. (2000), *Methods for Meta-Analysis in medical research*, Chichester: John Wiley & Sons, 2000.

Tang D. (2005), *Effects of the Regional Trading Arrangements on Trade: Evidence from the NAFTA, ANZCER and ASEAN Countries, 1989 – 2000*, Journal of International Trade & Economic Development, Vol. 14, No. 2, pp. 241 – 265.

Tenreyro S. (2001), *On the causes and consequences of currency unions*, Harvard University, unpublished.

Thom R., Walsh B. (2002), *The effect of a common currency on trade: Ireland before and after the sterling link*, European Economic Review, Vol.46, No. 6, pp. 1111–1124.

van den Bergh J., Button K.J. (1997), *Meta-analysis of Environmental Issues in Regional Urban and Transport Economics*, Urban Studies, 34(5-6), pp. 927-944.

Verdeja L. (2005), *EU's Preferential Trade Agreements With Developing Countries Revisited*, unpublished.

Viner J. (1950), *The Customs Union Issue*, Carnegie Endowment for International Peace, New York.

World Bank, (2005), *Global Economic Prespectives 2005, Regional Trade Agreements: Effects on Trade*, Chapter 3.

Weichselbaumer D., Winter-Ebmer R. (2005), *A meta-analysis of the international gender wage gap, Journal of Economic Surveys*, Vol. 19, pp. 479–512.

Yeyati L. E. (2003), *On the impact of a common currency on bilateral trade*, Economics Letters Vol. 79, No. 1, pp. 125–129.

APPENDIX 1

Reciprocal Trade Agreements (in chronological order of date of entry into force)

Trade Agreements	Date	Trade Agreements	Date
European Union (EU):	**1958**	**NAFTA:**	**1994**
Belgium-Luxembourg, France, Germany, Italy,		Canada, Mexico, United States	
Netherlands		**Bolivia-Mexico**	**1995**
Denmark, Ireland, United Kingdom	1973	**Caribbean Community-Dominican**	
Greece	1981	**Republic**	**1995**
Portugal, Spain	1986	**Costa Rica-Mexico**	**1995**
Austria, Finland, Sweden	1995	**EU-Bulgaria**	**1995**
Cyprus, Czech Rep., Estonia, Hungary, Latvia,		**EU-Romania**	**1995**
Lithuania, Malta, Poland, Slovak Rep., Slovenia	2004	**Group of Three:**	**1995**
The Customs Union of West African States*:	**1959**	Columbia, Mexico, Venezuela	
Burkina Faso, Mali, Mauritania, Niger, Senegal		**Mercado Comun del Sur (Mercosur):**	**1991**
European Free Trade Association (EFTA):	**1960**	Argentina, Brazil, Paraguay, Uruguay	formed in 1991 and
Switzerland, United Kingdom	Until 1973	**Caribbean Community-Dominican**	FTA in 1995
Norway, Portugal	Until 1986	**Republic**	**1998**
Sweden	Until 1995	**Common Market for Eastern and**	**effective 1998**
Denmark	1973	**Southern Africa*:**	
Finland	1986-1995	Egypt, Kenya, Madagascar, Malawi,	
Austria	1995	Mauritius, Sudan, Zimbabwe, Zambia	
Latin American Free Trade Agreement/Latin		**Andean Community*:**	**1995**
American Integration Agreement, (LAFTA/		Columbia, Ecuador, Peru Venezuela.	
LAIA):	**1961-1979, 1993**	Customs Union: Columbia,	
Argentina, Bolivia, Brazil, Chile, Ecuador,	ineffective 1980-1990	Ecuador,Venezuela.	
Mexico, Paraguay, Peru, Uruguay, Venezuela	reinitiated 1993	Free Trade Area: Bolivia	1995
African Common Market*:	**1963**	Peru	1997
Algeria, Egypt, Ghana, Morocco		**Mercosur-Chile**	**1996**
Central American Common Market*:	**1961-1975, 1993**	**Mercosur-Bolivia**	**1996**
El Salvador, Guatemala, Honduras, Nicaragua,		**Canada-Chile**	**1997**
Costa Rica	1965	**Canada-Israel**	**1997**
Economic Customs Union of the Central		**Association of Southeast Nations**	
African States*:	**1966**	**ASEAN:**	**1998**
Cameroon, Congo, Gabon		Indonesia, Philippines, Singapore,	
Caribbean Community (CARICOM):	**1968**	Thailand	
Jamaica, Trinidad and Tobago, Guyana	1995	**Hungary-Turkey**	**1998**
EU-EFTA/European Economic Area*	**1973/1994**	**Hungary-Israel**	**1998**
Australia -New Zealand Closer Economic		**India-Sri Lanka**	**1998**
Relations (CER)	**1983**	**Israel-Turkey**	**1998**
US-Israel	**1985**	**Mexico-Nicaragua**	**1998**
US-Canada	**1989**	**Romania-Turkey**	**1998**
EFTA-Israel	**1993**	**Poland-Israel**	**1998**
Central Europe Free Trade Agreement		**Romania-Turkey**	**1998**
(CEFTA):	**1993**	**Mexico-Chile**	**1999**
Hungary, Poland, Romania	1997	**EU-Israel Agreement**	**2000**
Bulgaria	1998	**EU-Mexico**	**2000**
EFTA-Bulgaria	**1993**	**Poland-Turkey**	**2000**
EFTA-Hungary	**1993**	**Mexico-Guatemala**	**2000**
EFTA-Poland	**1993**	**Mexico-Honduras**	**2000**
EFTA-Romania	**1993**	**Mexico-Israel**	**2000**
EU-Hungary	**1994**	**Mexico-El Salvador**	**2000**
EU-Poland	**1994**	**New Zealand-Singapore**	**2000**

Source: Baier and Bergstrand, 2005

*** Free Trade Areas that a also Customs Unions.**

APPENDIX 2

Econometrics results from the literature

	Selected Articles	Trade Agreements*	Sample	Number of Estimates	
				Positive (significant**)	Negative (significant**)
1	*Adam et al. 2003*	BFTA, CEFTA, EU	Cross-section data on Central and Eastern Europe total exports from 1992-2003	9 (8)	-
2	*Aitken, 1973*	EEC, EFTA	Cross-sectional trade flow model considering European trade relations over the period 1951-67	17 (8)	17 (0)
3	*Babetskaia-Kukharchuk, Maurel, 2004*	EU	Panel of data on trade of 14 EU countries over the period 1994-2001.	6 (3)	-
4	*Baier, Bergstrand, 2005*	RTAs	Panel of cross-section time-series data at 5 year intervals from 1960 to 2000 for 96 countries	17 (12)	8 (4)
5	*Bayoumi, Eichengreen, 1995*	EEC, EFTA	Panel of cross-section time-series data at 2 year intervals from 1956 to 1973	12 (4)	-
6	*Bergstrand, 1985*	EEC, EFTA	Cross-section using data for years 1965, 1966, 1975, 1976 for 15 OECD countries	8 (5)	-
7	*Bergstrand,1989*	EEC, EFTA	Cross-section using data for years 1965, 1966, 1975, 1976 for 15 OECD countries	70 (44)	2 (0)
8	*Brada, Mendez, 1985*	RTAs	Cross-section using data for over period 1990-94 for all OECD countries	3 (1)	-
9	*Breuss, Egger, 1999*	EU	Cross-section using data for years 1970, 1973, 1976 for member countries of EEC, EFTA CACM, LAFTA and Adean Community.	3 (3)	-
10	*Bun, Klaassen, 2002*	RTAs	Panel data on bilateral exports between the 15 European Union countries and the G7 countries outside Europe (Canada, Japan and the U.S.) from 1965 through 2001	3 (3)	-
11	*Carrère, 2006*	CACM, EU, LAIA	Panel data set including observations from over 130 countries from 1962 to 1996	25 (25)	-
12	*Cernat, 2001*	AFTA, CARICOM, EU, MERCOSUR, NAFTA	Cross-section dataset of more than 100 countries for three individual years: 1994, 1996, and 1998.	16 (12)	4 (0)
13	*Cheng, Tsai, 2005*	CUSFTA, EEC, EFTA, EU, LAFTA, MERCOSUR, NAFTA.	Pooled cross-section over the period 1981-97 for 44 exporting and 57 importing countries	28 (24)	8 (5)
14	*Cheng, Wall, 2004*	ANZCER, EU, MERCOSUR, NAFTA, US-ISRAEL	Panel data including 797 unidirectional country-pairs in each of four years: 1982, 1987, 1992, and 1997	20 (11)	5 (1)
15	*De Benedictis et al., 2005*	EU, EU-CEECs	Panel data on bilateral trade flows between eight CEECs and the EU-23 over period 1993-2003	2 (1)	-
16	*Eaton, Kortum, 1997*	EFTA, EU	Cross-section using 1990 data on trade in manufactures of 19 OECD countries	7 (1)	1 (0)
17	*Egger, 2005*	RTAs	Cross-section data on average 1990–97 bilateral exports of a sample of countries including OECD and non-OECD economies	3 (3)	-
18	*Elliot, Ikemoto, 2004*	ASEAN, EEC, EU, NAFTA	Panel on data of APEC, ASEAN, EU, NAFTA and other 16 countries over period 1982 to 1999	28 (23)	-
19	*Endoh, 2000*	ASEAN, EAEC	Cross-section analysis using a data set of 80 countries for every five-year term from 1960 to 1995.	50 (48)	2 (0)
20	*Faruquee, 2004*	RTAs	Panel data for 22 industrial countries, sample period from 1992-2002	12 (0)	2 (0)
21	*Fazio et al, 2005*	RTAs	Cross-section analysis using a data set of annual observations for 134 countries over 1980-2000	8 (1)	-
22	*Feenstra et al 2001*	RTAs	Five different cross-sections: 1970,1075, 1980, 1985, 1990. Sample of 110 countries considering differentiated and homogeneous goods.	34 (26)	-
23	*Fidrmuc, Fidrmuc, 2003*	BFTA, CEFTA, EC, EC+EFTA, EFTA,	Cross-section analysis using data on bilateral trade from 1990 to 1998 for OECD countries and Central and Eastern European countries.	130 (116)	5 (5)
24	*Frankel, Wei, 1997*	EEC, EFTA	Cross-sectional trade flow model considering 63 countries and data at 5 year intervals from 1960 to 1990	35 (14)	11 (1)
25	*Frankel, Rose, 2000*	RTAs	Panel data set including observations from over 180 countries at 5-year intervals from 1970 through 1995	4 (4)	-
26	*Frankel, Stein, Wei, 1995*	EEC, EFTA	Cross-sectional trade flow model considering 63 countries and data at 5 year intervals from 1965 to 1990	25 (10)	11 (0)
27	*Frankel, Stein, Wei, 1997*	EC	Cross-sectional trade flow model considering 63 countries and data at 5 year intervals from 1970 to 1990	2 (2)	
28	*Fratianni, Kang, 2006*	RTAs	Cross-section analysis using data on bilateral trade from 1970 to 1999, at five-years intervals, for 175 countries.	1 (1)	-
29	*Freund, Weinhold, 2004*	RTAs	Cross-section analysis using data on bilateral trade from 1995 to 1999 for 56 countries.	5 (1)	-
30	*Ghosh, Yamarik, 2004*	CACM, CARICOM, EEA, EFTA, EU, LAIA	The data set consists of six annual observations for 186 developing and developed countries. The annual observations are for 1970, 1975, 1980, 1985, 1990, and	5 (4)	1 (0)

			1995		
31	*Glick, Rose, 2002*	RTAs	Pooled panel, data set of 186 countries from 1948 to 1997.	1 (1)	-
32	*Grünfeld, Moxnes, 2003*	RTAs	Cross-section analysis using data on service exports for 1999 and 2000 of 22 OECD countries.	3 (0)	1 (0)
33	*Hassan, 2001*	EEC, NAFTA	Cross-section analysis using annual data on bilateral trade flow of 27 countries in years 1996 and 1997.	8 (2)	-
34	*Hejazi, Safarian, 2005*	NAFTA	Panel of data on U.S. trade with 52 countries over the period 1970 to 2002.	2 (2)	-
35	*Jakab et al., 2001*	CEE, CEFTA, EC+EFTA, MERCOSUR, NAFTA	Cross-section of trade data from 1990 to 1997 for 53 developed and non-developed countries.	8 (8)	10 (10)
36	*Jayasinghe, Sarker, 2004*	NAFTA	pooled cross-sectional time-series regression for trade of six selected agrifood products from 1985 to 2000 for NAFTA	24 (12)	6 (1)
37	*Kenen, 2002*	RTAs	Rose's (2000) data set (113 countries for 1990)	6 (6)	-
38	*Kien, Hashimoto, 2005*	AFTA, EU, MERCOSUR,NAFTA	Panel data on exports flows of 39 countries for the period 1988-2002.	12 (11)	4 (3)
39	*Kimura , Lee, 2004*	RTAs	Cross-section data on bilateral services trade and goods trade between 10OECD countries and other countries (OECD members and non OECD members) for the years 1999 and 2000.	42 (30)	4 (0)
40	*Klein, 2002*	RTAs	Cross-section of annual observations on 165 countries (27 industrial countries and 138 non-industrial countries) from 1948 to 1997	14 (7)	-
41	*Klein, 2005*	RTAs	Cross-section of annual observations on 165 countries (27 industrial countries and 138 non-industrial countries) from 1948 to 1997	6 (6)	1 (0)
42	*Krueger, 1999*	EU, ANZCER	pooled time-series-cross-section regression using data from 1987 to 1997 for members of various PTAs.	12 (1)	4 (0)
43	*Lee et al., 2004*	RTAs	Panel data set of 175 countries from 1948 to 1999.	15 (14)	-
44	*Lee, Park, 2005*	RTAs, AFTA, ANZCER, CACM, CAN, CARICOM, EC/EU, EFTA, MERCOSUR, NAFTA, and US-Israel FTA.	Panel data set of 175 countries from 1948 to 1999.	18 (12)	4 (1)
45	*Martìnez-Galàn et al, 2005*	EU	Cross-section analysis using data on trade of manufactured products between EU25 and Eastern European countries from 1999 to 2002	1 (0)	3 (0)
46	*Martìnez-Zarzoso, Horsewood, 2005*	CACM, CARICOM, EU, NAFTA	Sample of 47 countries from 1980 to 1999.	30 (28)	-
47	*Martìnez-Zarzoso, Nowak-Lehmann, 2003*	EU, MERCOSUR	Panel data of a sample of 20 countries, 15 EU countries and 5 Mercosur countries, from 1988 to 1996	28 (27)	-
48	*Mayer, Zignago, 2005*	RTAs, ADEAN, ASEAN, EU, MERCOSUR, NAFTA	Cross-section analysis data for 67 developing and developed countries over the period 1976-1999.	26 (21)	-
49	*Meliz, 2001*	RTAs	Frankel and Rose (2000) database	4 (4)	-
50	*Meliz, 2002*	RTAs	Frankel and Rose (2000) database	3 (3)	-
51	*Micco et at, 2003*	RTAs	Panel data set including information on bilateral trade for 22 developed countries from 1992 to 2002.	13 (5)	5 (1)
52	*Nilsson, 2005*	EEA, NAFTA	Cross-section analysis for the 2001-2003 period of EU and US imports from developing countries.	4 (2)	-
53	*Nitsch, 2002*	RTAs	Rose's (2000) data set.	15 (15)	-
54	*Paiva, 2005*	RTAs	Data set covers bilateral trade in agricultural goods for 152 countries over the periods 1990–93 and 1999–2002.	5 (5)	-
55	*Pakko, Wall, 2001*	RTAs	Rose's (2000) data set	6 (3)	-
56	*Rauch, 1996*	EEC, EFTA	Cross-section, data on 63 countries for the years 1970, 1980, 1990	42 (6)	30 (11)
57	*Rauch, Trindade, 1999*	EEC,EFTA	Cross-section, data on 63 countries for the years 1970, 1980, 1990	29 (4)	19 (2)
58	*Rose, 2000*	RTAs	Panel data, bilateral observations for five during 1970-90 covering 186 countries	50 (50)	3 (1)
59	*Rose, 2003*	RTAs	Panel data, bilateral observations for five during 1950-2000 covering 175 countries.	12 (12)	-
60	*Rose, 2004*	RTAs	Panel data, bilateral observations for five during 1950-2000 covering 175 countries.	6 (6)	-
61	*Rose, 2005*	RTAs	Panel data, bilateral observations for five during 1950-2000 covering 150 countries.	17 (16)	-
62	*Rose, Engel, 2002*	RTAs	Cross-section analysis using a data set of annual	4	-

		observations for 210 countries between 1960 and 1996	(4)	
63 *Rose, van Wincoop, 2001*	RTAs	Panel data on bilateral observations for five during 1970-95 covering 200 countries.	2 (2)	-
64 *Sanso et al., 1993*	EEC, EFTA	Cross-section of annual observations on trade in 16 OECD countries from 1964 to 1987	27 (11)	1 (0)
65 *Sapir, 2001*	EFTA	Cross-section, annually over the period 1960–1992 on the 240 bilateral trade flows	18 (0)	15 (4)
66 *Silva, Tenreyro, 2003*	RTAs	Cross-section analysis of 137 countries in 1990.	6 (6)	-
67 *Silva, Tenreyro, 2005*	RTAs	Cross-section analysis of 137 countries in 1990.	12 (10)	-
68 *Sissoko, 2004*	BFTA, CEFTA	Panel of 36 countries of the European zone with annual data during the period 1988 – 2000.	27 (19)	-
69 *Subramanian, Wei 2003*	RTAs	Panel data set of annual data over the period 1960–1992 on the 240 bilateral trade flows	34 (32)	-
70 *Subramanian, Wei 2005*	RTAs	Panel data set of annual data over the period 1960–1992 on the 240 bilateral trade flows	32 (31)	1 (0)
71 *Tang, 2005*	ANZCER, ASEAN, NAFTA	The data set covers the bilateral trade flows for 21 countries from 1989 to 2000.	10 (8)	-
72 *Tenreyro, 2001*	RTAs	Panel data set of annual observations for over 200 countries from 1978 to 1997.	4 (4)	-
73 *Thom, Walsh, 2002*	AIFTA, EEC	Panel and cross-section analysis for Anglo–Irish trade over the period 1950–1998	13 (7)	7 (4)
74 *Verdeja, 2005*	EFTA	Cross-section data covering 137 countries for the period 1973-2000.	11 (6)	4 (1)
75 *Yeyati, 2003*	RTAs	Rose's (2000) data set (186 countries for 1995)	10 (10)	-

** RTAs indicates estimates that do not specify the type of agreement.*
*** Statistically significant at the 5% level.*

APPENDIX 3

Papers included in the database

Descriptive Statistics

References	Nb. of Estimates	Ranges		Simple Mean	Standard Error
		Min	*Max*		
Adam et al. 2003	9	0.48	3.69	1.70	1.16
Aitken, 1973	34	-0.21	0.89	0.16	0.33
Babetskaia-Kukharchuk, Maurel, 2004	18	-0.51	3.37	0.98	1.28
Baier, Bergstrand, 2005	25	-3.97	2.51	0.12	1.19
Bayoumi, Eichengreen, 1995	12	0.01	0.21	0.07	0.07
Bergstrand, 1985	8	0.18	0.73	0.47	0.24
Bergstrand,1989	72	-0.11	1.93	0.73	0.48
Brada, Mendez, 1985	3	3.77	4.83	4.43	0.57
Breuss, Egger, 1999	3	0.29	0.42	0.38	0.08
Bun, Klaassen, 2002	3	0.02	0.08	0.05	0.03
Carrère, 2006	25	0.22	0.99	0.58	0.21
Cernat, 2001	20	-0.72	4.41	1.05	1.42
Cheng, Tsai, 2005	35	-0.35	4.35	0.74	1.00
Cheng, Wall, 2004	25	-0.16	3.98	0.59	0.97
De Benedictis et al., 2005	2	0.11	0.14	0.12	0.02
Eaton, Kortum, 1997	8	-0.12	0.54	0.23	0.23
Egger, 2005	3	0.52	1.29	0.78	0.44
Elliot, Ikemoto, 2004	28	0.10	2.35	0.55	0.56
Endoh, 2000	52	-0.07	1.93	0.85	0.50
Faruquee, 2004	14	-0.01	0.01	0.01	0.01
Fazio et al, 2005	8	0.10	0.27	0.17	0.05
Feenstra et al 2001	34	0.18	2.20	1.09	0.53
Fidrmuc, Fidrmuc, 2003	135	-0.15	3.96	0.68	0.98
Frankel, Wei, 1997	46	-0.41	1.15	0.21	0.33
Frankel, Rose, 2000	4	1.16	1.31	1.25	0.06
Frankel, Stein, Wei, 1995	36	-0.32	1.51	0.16	0.36
Frankel, Stein, Wei, 1997	2	0.24	0.31	0.28	0.05
Fratianni, Kang, 2006	2	1.04	1.37	1.21	0.23
Freund, Weinhold, 2004	5	0.02	0.28	0.16	0.09
Ghosh, Yamarik, 2004	6	-0.11	2.22	0.79	1.02
Glick, Rose, 2002	1	0.99	0.99	0.99	.
Grünfeld, Moxnes, 2003	4	-0.14	0.17	0.02	0.13
Hassan, 2001	8	0.22	4.71	2.45	1.78
Hejazi, Safarian, 2005	2	7.90	10.84	9.37	2.08
Jakab et al., 2001	18	-2.03	0.69	-0.17	0.84
Jayasinghe, Sarker, 2004	30	-1.47	3.76	0.83	1.13
Kenen, 2002	6	0.59	2.32	1.12	0.64
Kien, Hashimoto, 2005	16	-0.46	2.23	0.41	0.62
Kimura , Lee, 2004	48	0.19	0.65	0.36	0.12
Klein, 2002	14	0.07	2.35	0.85	0.59
Klein, 2005	7	-0.48	2.52	0.99	0.88
Krueger, 1999	2	0.07	0.50	0.29	0.30
Lee 2004	15	0.09	2.57	0.64	0.58
Lee, Park, 2005	22	-0.35	1.97	0.65	0.61
Martìnez-Galàn et al, 2005	4	-0.06	0.00	-0.02	0.02
Martìnez-Zarzoso , Horsewood, 2005	30	0.12	2.63	1.16	0.66
Martinez-Zarzoso, Nowak-Lehmann, 2003	28	0.04	0.65	0.24	0.16
Mayer, Zignago, 2005	26	0.08	2.44	1.46	0.77

Meliz, 2001	4	1.03	1.24	1.16	0.10
Meliz, 2002	3	1.00	1.02	1.01	0.01
Micco et at, 2003	18	-0.19	0.10	0.03	0.06
Nilsson, 2005	4	0.27	0.28	0.28	0.01
Nitsch, 2002	15	0.68	1.28	1.07	0.19
Paiva, 2005	5	1.01	1.15	1.10	0.05
Pakko, Wall, 2001	6	0.05	0.91	0.43	0.40
Rauch, 1996	72	-1.18	1.11	0.03	0.47
Rauch, Trindade, 1999	48	-0.64	0.46	0.07	0.31
Rose, 2000	54	-0.97	1.54	0.78	0.45
Rose, 2003	12	0.18	1.50	0.92	0.38
Rose, 2004	6	0.94	1.50	1.19	0.18
Rose, 2005	17	0.07	0.75	0.53	0.21
Rose, Engel, 2002	4	0.75	0.95	0.88	0.09
Rose, van Wincoop, 2001	2	0.46	1.09	0.78	0.45
Sanso et al., 1993	28	-0.05	1.32	0.34	0.35
Sapir, 2001	33	-0.54	0.34	0.01	0.26
Silva, Tenreyro, 2003	6	0.26	0.79	0.44	0.26
Silva, Tenreyro, 2005	12	0.14	1.29	0.48	0.36
Sissoko, 2004	27	0.21	2.64	1.57	0.77
Subramanian, Wei 2003	34	0.18	1.99	0.90	0.39
Subramanian, Wei 2005	33	-0.13	1.99	0.93	0.45
Tang, 2005	10	0.26	1.83	0.80	0.60
Tenreyro, 2001	4	0.29	0.70	0.53	0.19
Thom, Walsh, 2002	20	-0.10	0.74	0.11	0.25
Verdeja, 2005	15	-1.38	1.90	0.39	0.90
Yeyati, 2003	10	0.47	1.00	0.58	0.16

APPENDIX 4

Within-Study Meta-Analysis of RTAs effect on trade

Study		Coefficient	$H_0 : \gamma = 0$ (p-value)	test Q (p-value)	H^2	I^2	Heterogeneity
Adam et al. 2003	Fixed Random	0.91 1.61	0.00 0.00	0.00	45.26	98%	High
Aitken, 1973	Fixed Random	0.20 0.19	0.00 0.00	0.00	1.78	44%	Moderate
Babetskaia-Kukharchuk, Maurel, 2004	Fixed Random	0.51 0.94	0.00 0.00	0.00	25.55	96%	High
Baier, Bergstrand, 2005	Fixed Random	0.14 0.16	0.00 0.03	0.00	25.78	96%	High
Bayoumi, Eichengreen, 1995	Fixed Random	0.09 0.08	0.00 0.00	0.02	2.08	52%	Moderate
Bergstrand, 1985	Fixed Random	0.45 0.46	0.00 0.00	0.01	2.71	63%	Moderate
Bergstrand,1989	Fixed Random	0.80 0.76	0.00 0.00	0.00	3.92	74%	High
Brada, Mendez, 1985	Fixed Random	4.34 4.34	0.00 0.00	0.93	0.07	0%	Low
Breuss, Egger, 1999	Fixed Random	0.38 0.38	0.00 0.00	0.43	0.86	0%	Low
Bun, Klaassen, 2002	Fixed Random	0.05 0.05	0.00 0.00	0.03	9	89%	High
Carrère, 2006	Fixed Random	0.51 0.57	0.00 0.00	0.00	13.06	92%	High
Cernat, 2001	Fixed Random	0.47 1.04	0.00 0.00	0.58	17.18	94%	High
Cheng, Tsai, 2005	Fixed Random	0.18 0.70	0.00 0.00	0.00	168.97	99%	High
Cheng, Wall, 2004	Fixed Random	0.30 0.60	0.00 0.00	0.46	54.08	98%	High
De Benedictis et al., 2005	Fixed Random	0.12 0.12	0.01 0.01	0.74	0.11	0%	Low
Eaton, Kortum, 1997	Fixed Random	0.20 0.20	0.01 0.01	0.43	1.00	0%	Low
Egger, 2005	Fixed Random	0.87 0.80	0.00 0.00	0.00	6.22	84%	High
Elliot, Ikemoto, 2004	Fixed Random	0.33 0.49	0.00 0.00	0.00	13.51	93%	High
Endoh, 2000	Fixed Random	0.73 0.83	0.00 0.00	0.00	12.01	92%	High
Faruquee, 2004	Fixed Random	0.00 0.00	0.83 0.83	1.00	0.07	0%	Low
Fazio et al, 2005	Fixed Random	0.17 0.17	0.00 0.00	0.98	0.21	0%	Low
Feenstra et al 2001	Fixed Random	1.23 1.14	0.00 0.00	0.00	12.44	92%	High
Fidrmuc, Fidrmuc, 2003	Fixed Random	0.05 0.49	0.00 0.00	0.00	34.58	97%	High
Frankel, Wei, 1997	Fixed Random	0.25 0.23	0.00 0.00	0.00	2.13	53%	Moderate
Frankel, Rose, 2000	Fixed Random	1.25 1.25	0.00 0.00	0.79	0.35	0%	Low
Frankel, Stein, Wei, 1995	Fixed Random	0.45 0.23	0.00 0.00	0.00	3.43	71%	High
Frankel, Stein, Wei, 1997	Fixed Random	0.29 0.29	0.00 0.00	0.52	0.42	0%	Low
Fratianni, Kang, 2006[*]	Fixed Random	1.20 1.20	0.00 0.00	0.00	9.55	90%	High
Freund, Weinhold, 2004	Fixed Random	0.11 0.12	0.01 0.01	0.29	1.24	20%	Low
Ghosh, Yamarik, 2004	Fixed Random	0.38 0.74	0.00 0.03	0.00	47.12	98%	High
Glick, Rose, 2002[*]	Fixed Random	0.99 0.99	0.00 0.00	-	-	-	-
Grünfeld, Moxnes, 2003	Fixed Random	0.03 0.03	0.70 0.70	0.56	0.70	0%	Low
Hassan, 2001	Fixed Random	2.61 2.61	0.00 0.00	0.44	0.98	0%	Low
Hejazi, Safarian, 2005	Fixed Random	8.68 8.68	0.00 0.00	0.42	0.64	0%	Low

Jakab et al., 2001	Fixed	-0.44	0.00	0.00	484.32	100%	High
	Random	-0.18	0.15				
Jayasinghe, Sarker, 2004	Fixed	0.69	0.00	0.00	4.71	79%	High
	Random	0.81	0.00				
Kenen, 2002	Fixed	0.77	0.00	0.07	2.04	51%	Moderate
	Random	0.86	0.00				
Kien, Hashimoto, 2005	Fixed	0.19	0.00	0.00	92.80	99%	High
	Random	0.41	0.01				
Kimura , Lee, 2004	Fixed	0.37	0.00	0.00	1.02	0%	Low
	Random	0.37	0.00				
Klein, 2002	Fixed	0.87	0.00	0.01	2.15	54%	Moderate
	Random	0.87	0.00				
Klein, 2005	Fixed	1.06	0.00	0.00	6.12	84%	High
	Random	1.04	0.00				
Krueger, 1999	Fixed	0.10	0.13	0.10	2.64	62%	Moderate
	Random	0.21	0.29				
Lee et al., 2004	Fixed	0.49	0.00	0.00	32.78	97%	High
	Random	0.55	0.00				
Lee, Park, 2005	Fixed	0.51	0.00	0.00	20.54	95%	High
	Random	0.62	0.00				
Martìnez-Galàn et al, 2005	Fixed	-0.02	0.82	0.99	0.03	0%	Low
	Random	-0.02	0.82				
Martinez-Zarzoso, Horsewood, 2005	Fixed	0.95	0.00	0.00	10.49	90%	High
	Random	1.11	0.00				
Martinez-Zarzoso, Nowak-Lehmann, 2003	Fixed	0.06	0.00	0.00	23.05	96%	High
	Random	0.18	0.00				
Mayer, Zignago, 2005	Fixed	1.53	0.00	0.00	23.74	96%	High
	Random	1.49	0.00				
Meliz, 2001	Fixed	1.16	0.00	0.45	0.88	0%	Low
	Random	1.16	0.00				
Meliz, 2002	Fixed	1.01	0.00	0.99	0.02	0%	Low
	Random	1.01	0.00				
Micco et at, 2003	Fixed	0.02	0.00	0.00	8.51	88%	High
	Random	0.02	0.19				
Nilsson, 2005	Fixed	0.27	0.00	1.00	0.00	0%	Low
	Random	0.27	0.00				
Nitsch, 2002	Fixed	1.06	0.00	0.08	1.57	36%	Moderate
	Random	1.07	0.00				
Paiva, 2005	Fixed	1.10	0.00	0.85	0.34	0%	Low
	Random	1.10	0.00				
Pakko, Wall, 2001	Fixed	0.53	0.00	0.00	17.23	94%	High
	Random	0.43	0.01				
Rauch, 1996	Fixed	0.03	0.35	0.00	3.16	68%	Moderate
	Random	0.04	0.47				
Rauch, Trindade, 1999	Fixed	0.10	0.00	0.00	1.54	35%	Moderate
	Random	0.10	0.01				
Rose, 2000	Fixed	0.10	0.00	0.01	85.92	99%	High
	Random	0.83	0.00				
Rose, 2003	Fixed	0.82	0.00	0.00	12.86	92%	High
	Random	0.92	0.00				
Rose, 2004	Fixed	1.18	0.00	0.00	1.69	41%	Moderate
	Random	1.18	0.00				
Rose, 2005	Fixed	0.60	0.00	0.00	9.53	90%	High
	Random	0.54	0.00				
Rose, Engel, 2002	Fixed	0.88	0.00	0.00	0.35	0%	Low
	Random	0.88	0.00				
Rose, van Wincoop, 2001	Fixed	0.83	0.00	0.13	16.27	94%	High
	Random	0.78	0.01				
Sanso et al., 1993	Fixed	0.32	0.00	0.79	7.19	86%	High
	Random	0.35	0.00				
Sapir, 2001	Fixed	-0.02	0.68	0.00	1.55	36%	Moderate
	Random	-0.01	0.90				
Silva, Tenreyro, 2003	Fixed	0.35	0.00	0.00	5.45	82%	High
	Random	0.42	0.00				
Silva, Tenreyro, 2005	Fixed	0.41	0.00	0.02	9.01	89%	High
	Random	0.46	0.00				
Sissoko, 2004	Fixed	1.48	0.00	0.00	1.99	50%	Moderate
	Random	1.46	0.00				
Subramanian, Wei 2003	Fixed	0.86	0.00	0.00	6.80	85%	High
	Random	0.88	0.00				
Subramanian, Wei 2005	Fixed	0.90	0.00	0.00	9.00	89%	High
	Random	0.92	0.00				
Tang, 2005	Fixed	0.76	0.00	0.00	18.02	94%	High

	Random	0.80	0.00				
Tenreyro, 2001	Fixed Random	0.53 0.53	0.00 0.00	0.00	5.16	81%	High
Thom, Walsh, 2002	Fixed Random	0.04 0.10	0.00 0.00	0.00	18.54	95%	High
Verdeja, 2005	Fixed Random	0.07 0.40	0.00 0.01	0.00	6.06	84%	High
Yeyati, 2003	Fixed Random	0.58 0.58	0.00 0.00	0.05	1.91	48%	Moderate

** Studies with only one estimate.*