Policy Measurement and Multilateral Resistance in Gravity Models

Cipollina, Maria Pina and De Benedictis, Luca and Salvatici, Luca and Vicarelli, Claudio

Università del Molise, Università di Macerata, Università Roma3, ISTAT

November 2016

POLICY MEASUREMENT AND MULTILATERAL RESISTANCE IN GRAVITY MODELS

Maria Cipollina[a], Luca De Benedictis[b], Luca Salvatici[c], Claudio Vicarelli[d]

(NOVEMBER 2016)

ABSTRACT

Over the past decade, the gravity equation has emerged as the empirical workhorse in international trade to study the ex-post effects of trade policies on bilateral trade. In this paper we are concerned with the issue of how the econometric specification and the policy measurement choices can affect the goal to obtain accurate estimates of the coefficient associated with bilateral trade policies within a theoretically-consistent model. The problem is even more serious when the policy treatment is approximated through dummies as it is still often the case in the literature. Using a Monte Carlo simulation analysis, this paper shows that the use of fixed effects to control for unobserved heterogeneity leads to biased estimates of the policy impact even when the policy is measured through a continuous variable. The bias highlighted by our results is the combination of measurement error about bilateral trade costs (or preferences) and the specification used to proxy multilateral resistance terms.

KEYWORDS: Gravity model; Multilateral trade resistance; Policy evaluation; Monte Carlo Analysis.

JEL CLASSIFICATION: C10; C50; F14

[a] University of Molise, Via Francesco De Sancits 1, 86100 Campobasso, Italy. Cipollina@unimol.it

[b] University Roma Tre, Italy. Dipartimento di Economia - Via Silvio D'Amico, 77 Roma 00145. debene@unimc.it

[c] University of Macerata, Via Crescimbeni 20, 62100 Macerata, Italy. Luca.salvatici@uniroma3.it

[d] ISTAT, Italian National Insitute of Statistics, Via Cesare Balbo 16, 00184 Rome, Italy. cvicarelli@istat.it

1. Introduction

The gravity model has been used as a workhorse for analysing the determinants of bilateral trade flows for 50 years since being popularized by Tinbergen (1962) and it represents an important tool for the analysis of the effects of trade policies. In this paper, we are concerned with two fundamental problems related to gravity-based estimates of the trade impact of policies: multilateral trade resistance and policy measurement.

The first issue was emphasized by Anderson and van Wincoop (2003), whose work presented what would become one of the most widely cited gravity models, showing that the typical gravity equation should account for both "bilateral resistance," e.g. the barriers to trade between a pair of countries, and the so-called "multilateral resistance" term (MRT), capturing the general equilibrium effect associated with the barriers to trade that each country faces with all its trading partners. In its essence, what matters in explaining bilateral trade flows are relative bilateral trade costs, and an omission of the MRT may lead to bias estimates of the relevant elasticities.

The treatment of the multilateral resistance terms in gravity estimations has evolved over the years and researchers have proposed various solutions (Piermartini and Yotov, 2016).
In order to overcome the computational difficulties of a structural estimation explicitly expressing the exporter- and importer-specific terms as a function of the economic model's variables, as suggested by Anderson and van Wincoop (2003),[1] many researchers have approximated the MR by the so-called 'remoteness indexes' that are constructed as weighted averages of bilateral distance, with Gross Domestic Products (GDPs) used as weights (see for example Wei, 1996; and Baier and Bergstrand, 2009). Head and Mayer (2014) criticize such reduced- form approaches as they "bear little resemblance to [their] theoretical counterpart." (p.150).[2]

Hummels (2001) and Feenstra (2004) advocate the use of directional (exporter and importer) fixed effects in cross-section estimations, and the recommendation has been extended from the cross-sectional setting to a longitudinal one in a static (Baldwin and Taglioni, 2006; De Benedictis and Taglioni, 2011) or a dynamic context (De Benedictis and

[1] Fally (2015) demonstrates that gravity regressions with fixed effects and Poisson PML can be used as a simple tool to solve the estimation problem raised by Anderson and van Wincoop (2003). Indeed, in the Poisson PML specification the estimated fixed effects are consistent with the definition of outward and inward multilateral resistance indexes and the equilibrium constraints that they need to satisfy.

[2] An alternative approach to handle the multilateral resistances is to simply eliminate these terms by using appropriate ratios based on the structural gravity equation. Notable examples include Head and Ries (2001), Head, Mayer and Ries (2010), and Novy (2013).

Vicarelli, 2008; Olivero and Yotov, 2012): the MR terms should be accounted for by exporter-time and importer-time fixed effects in a gravity estimation framework with panel data. Structural estimators fully exploit the information on the data-generating process and are thus potentially more efficient than other approaches. However, they require the underlying economic model to be correctly specified.

In contrast, the fixed effects approach is to leave the fixed effects unspecified, being consistent regardless of the nature of their dependence (Egger and Staub, 2016). Fixed effects also avoid misspecification problems that could lead to spurious regressions. In particular, fixed effects include all the possible time-invariant unobserved factors specifically affecting bilateral trade flows in each country-pair: geographical, political, cultural, institutional factors. Fixed effect method produces consistent estimates of average border effects across countries; moreover, by replacing bilateral distances with importer-exporter fixed effects, it avoids the shortcomings of distance as a measure of transport and information costs.[3]

In order to control for time-varying unobserved trade costs in gravity equations, Baltagi et al. (2003) suggested a full interaction effects design to analyse bilateral trade flows controlling for all sorts of unobserved heterogeneity. However, the exporter-time and importer-time fixed effects will also absorb the size variables from the structural gravity model as well as all other observable and unobservable country-specific characteristics which vary across these dimensions, including various national policies. Finally, as for nonlinear models concern, an important drawback of fixed effects estimators is that they suffer from the incidental parameters problem. In the context of the gravity equation, this means that as the number of observations grows, the number of parameters that has to be estimated grows as well. In general, this implies that fixed effects coefficients cannot be estimated consistently and, worse, the inconsistency passes over to the estimate of the main variables of interest (Egger and Staub, 2016).

The second issue refers to the trade policy measurement. Although it is well known that gravity type models using policy dummies are misspecified and lead to erroneous economic inferences (Màtyàs, 1997; Stack, 2009; Hornok, 2011), the approximation of trade policy

[3] Geographic distance is usually proxied by kilometers between capital cities of exporting and importing countries. This measure implicitly assumes that transport costs do not vary depending on transport mode, i.e. overland and overseas transport costs are equal. Indeed, the straight-line distance assumes only one economic center per country, but in fact a large country may have several economic centers.

through dummies – even when the policy measure is inherently continuous - is still quite common.[4]

The empirical literature emphasizes the importance of expressing trade policies through continuous variables since they vary widely across products, importers and exporters, and at least in the case of tariffs detailed data are currently available (De Benedictis and Salvatici, 2011). In this paper we consider preferential trade policies because differently by other policies (e. g. trade creating effect of a custom union, of a free trade agreement or of a currency unions that are measured by dichotomous dummy) can be measured through a continuous variable.

The use of a simple dummy to capture the impact of preferential policies on trade is inadequate because: (i) it confuses the policy implementation with all other factors that are specific to the country-pair and are contemporaneous to the policy; (ii) it does not discriminate among different policy instruments; (iii) it does not match the level of trade barriers (Cardamone, 2011). Moreover, since country-time dummies tend to absorb too much of the variation in the data, the heterogeneous policy effects are hardly identified.

The two issues, MRT and policy measurement, are related since trade policies have both direct and indirect effects on trade: a reduction or increase in a tariff in a particular sector has an effect on the price of goods in that sector but also in the rest of the economy, as a consequence it affects the general equilibrium prices at home and abroad (Caliendo and Parro, 2015). As a matter of fact, Egger and Staub (2016) show that predictions and average marginal effects are functions not only of the direct coefficient but also of the fixed effects estimates). Cipollina et al (2016) include in the preference margin definition policy the price index consistent with the CES demand functional form. Moreover, they show that the policy coefficient represents the elasticity of substitution: a crucial parameter in the debate about impact of trade liberalization, which our results show how much is influenced by the modelling choices in terms of MRT and policy measurement.

In this paper, we provide a quantitative assessment of the following proposition: the more we control for the MTR and unobserved heterogeneity through a full set of fixed effects, the less we are able to estimate the policy effect, and this is particularly true when the latter is proxied by a dummy variable. As a consequence, the estimates of trade impact is biased.

[4] As a proxy for the existence of preferential trade agreements (PTA), many contributions to the gravity literature use a "PTA dummy". The quantitative survey by Cipollina and Pietrovito (2011) shows that the results of the empirical analyses using a "PTA dummy" specification tend to overestimate the effect of the trade policy, relatively to the results obtained by studies using some explicit measures of the preference margins.

Our goal is to assess to what extent estimates of the impact of the policy treatment are affected when all the cross sectional and time series heterogeneity is controlled for. Such an assessment shows the possible consequences of compounding the error in measuring the treatment (i.e., the policy variable) with a more or less complete specification of the fixed effects structure.

To do this, we implement a data-generating-process and set up a Monte Carlo simulation analysis that allows to control for the consequences of various fixed-effects specifications on the estimated value of the parameter of interest (i.e., the policy impact).[5] Our data-generating process is built on an Heckman selection model. We show that the MRT fixed effects strategy may not be appropriate when it is coupled with a measurement error in the policy variable. Furthermore, through the simulation exercise we are able to precisely assess the consequences of different possible fixed effects specifications of the gravity equation in the presence of zero trade flows.

2. Gravity for Policies

The gravity model has been defined as the workhorse of international trade and its ability to correctly approximate bilateral trade flows makes it one of the most stable empirical relationships in economics (Leamer and Levinsohn 1995). Over the years there has been dramatic progress both in understanding the theoretical basis for the equation and in improving its empirical estimation (De Benedictis and Taglioni, 2011). Although for several years these models has been thought of as a purely "empirical fact" reminiscent of Newton's law of universal gravitation, trade theorists have highlighted the fact that gravity-like equations can be derived from several trade modelling frameworks (see Head and Mayer (2014) for a survey, and Baltagi, Egger and Pfaffermayer (2015) for the empirical counterparts of different theoretical models): CES product differentiation with imperfect competition (Anderson, 1979; Bergstrand, 1985), Ricardian models (Eaton and Kortum, 2002), heterogeneous firms (Chaney, 2008; Helpman et al., 2008).

One of the traditional purposes of gravity equations has been to estimate the elasticity of bilateral trade to various policies promoting trade, or to covariates capturing the "cost of

[5] In an influential article, Santos Silva and Tenreyro (2006) argued that gravity equations should be estimated in their multiplicative form and use Monte Carlo simulations to compare the performance of nonlinear estimators with that of ordinary least squares (in the log linear specification). Since then, the use of Monte Carlo analysis has become increasingly common in gravity equation experiments (Santos Silva and Tenreyro, 2011; Martinez Zarzoso, 2013; Bergstrand et al., 2013; Head and Mayer, 2014; Martin and Pham, 2014).

distance". In this paper, we are interested in the wider use of the gravity equation: the ex-post evaluation of the trade-enhancing effect of preferential trade policy. The trade cost is a negative function of the preferential margin since trade preferences reduce border costs as a consequence of tariff reduction. As in Cipollina et al (2010) we define the preferential margin ($pref_{ij,t}^{k}$) as the ratio between the reference tariff factor $(1+T_{j,t}^{k})$ and the applied tariff factors faced by each exporter $(1+\tau_{ij,t}^{k})$:

$$pref_{ij,t}^{k} = \frac{(1+T_{j,t}^{k})}{(1+\tau_{ij,t}^{k})} \quad (1)$$

In particular, the measure preference margin is expressed in relative rather than absolute terms and this definition focuses on actual preferences with respect to possible competitors, rather than measuring theoretical margins with respect to "bound" Most Favoured Nation (MFN) tariffs (i.e., the ceiling set by the World Trade Organization (WTO) commitments). The critical issue is the measurement of the reference tariff with respect to whom the preference margin is determined. To account the advantage with respect to actual not potential exporters, thus to avoid possible overestimation of the competitive advantages enjoyed by exporting countries, we use as reference tariff the highest applied tariff factor.

The continuous policy variable quantifies the advantage granted with respect to other importers: higher preferences decrease trade costs and, thus, reduce the negative trade impact of the bilateral distance.

One of the most widely used specifications of a theory-based gravity equation is provided by Anderson and van Wincoop (2003). In their formulation, bilateral trade flows between two regions depend on the output of both regions and the bilateral trade costs relative to the trade costs faced by the other regions, the so-called "multilateral resistance" term (MRT). The latter term is defined as the sum of three components: (1) the bilateral trade barrier between region i and region j, (2) i's resistance to trade with all regions, and (3) j's resistance to trade with all regions. We argue that the most common way to handle the MRT in gravity models – namely through the introduction of fixed effects – can be highly misleading since country-time-product dummies are likely to absorb the variation in the variable representing the relevant policy effect. This is confirmed by the results of Monte Carlo simulations, and the problem turns out to be particularly serious when the policy treatment, which cannot be correctly measured, is proxied through dummy variables.

3. The Monte Carlo Simulation

To explore the performance of different modelling choices in finite samples, we set up a Monte Carlo experiment. The questions that we seek to explore are how the estimate of the policy impact is affected by the policy measurement, and by an increase in the number of fixed effects.

3.1 The data-generating process

The data generating process consists of a structural part, including the bilateral-, exporter-, and importer-specific determinants of exports; and a stochastic part, in which random errors are drawn and joined to the structural part of exports.

Following the most recent literature highlighting the need to account for the variability of trade policy at the tariff-line level (Baldwin and Taglioni, 2006; De Benedictis and Salvatici, 2011; Head and Mayer, 2014), this work is based on disaggregated data. The setup of the Monte Carlo simulation includes 23 industrial sectors (ISIC Revision 2 Classification), 14 countries importing from 17 exporters, over the period 1998-2001. The dataset is built on information provided by the TradeProd, TradePrices and the GeoDist Cepii databases (http://www.cepii.fr/) on existing trade flows, production, expenditure, applied tariffs, import and export price indices, distances between countries and dummies for contiguity, common language, and former colonial links. Table A1 in the Appendix presents summary statistics.

The main weakness of the use of published data on price indexes is that existing price indexes may not reflect true border effects accurately (Feenstra, 2002). However, this is not a problem in our analysis, since trade flows are generated through the Monte Carlo simulation.

Starting from the initial dataset of 21,896 observations, we implement a data-generating-process, replicated 10,000 times, using a standard bootstrap methodology, to draw a random sample containing 5 importer countries, 10 products and 5 exporter countries over a period of 4 years. In each replication, the final dataset contains 1,000 observations that are manageable in regressions with the full set of dummies.

We address the issue of zero flows by adopting the Heckman (1979) approach, assuming that the selection equation contains at least one variable that is excluded from the behavioral equation. We use equation (2) for the behavioral equation:

$$y_i = exp(\beta x_i) + \varepsilon_i \tag{2}$$

where x_i is a vector of explanatory variables; β is a vector of coefficient to be estimated; and ε_i is the error term for observation *i*.

With a constant-elasticity functional form, almost always used, model estimated in levels will have heavily skewed and heteroskedastic errors. We follow Santos Silva and Tenreyro (SST)[6] (2006) to design heteroskedastic error, and rewrite equation (2) in their formulation as follows:

$$y_{1i} = exp(\beta x_i) * \eta_i \tag{3}$$

with $\eta_i = 1 + \varepsilon_i / exp(\beta x_i)$,[7] drawn as log normal random variable with mean 1 and variance σ_i^2.

We add the following sample selection equation:

$$y_{2i} = exp(\alpha z_i) * \eta_i^* \tag{4}$$

where z_i is the vector of explanatory variable in x_i plus the regressor (ξ_i) excluded from the behavioural equation[8]. We follow Santos Silva and Tenreyro (SST) (2006) to design heteroskedastic error, $\eta_i = 1 + \varepsilon_i / exp(\beta x_i)$,[9] drawn as log normal random variable with mean 1 and variance σ_i^2. The error term, η_i^*, in the sample selection equation is drawn as log normal random variable with a correlation of 0.5 with η_i.

We apply the selection rule to generate zeros of y_{1i} as follows: for y_{2i} with values less than zero, replace y_{1i} with zero. For y_{2i} with values greater than zero, y_{1i} remains the same.

Coherently with the gravity equation, determinants of trade flows are:

$$exp(\beta x_i) = (\beta_0 + \beta_1 \ln E_{j,t}^k + \beta_2 \ln Y_{i,t}^k - \beta_3 \ln Y_t^k + \gamma_1 lndist_{ij} + \gamma_2 border_{i,j} + \gamma_3 language_{ij\ ij}$$
$$(\sigma - 1) \ln pref_{ij,t}^k + (\sigma - 1) \ln P_{j,t}^k + (\sigma - 1) \ln \Pi_{i,t}^k + \gamma_4 D_t) \tag{5}$$

---

[6] Santos Silva and Tenreyro (2006) page 647.

[7] Rewriting $\varepsilon_i = exp(\beta x_i) v_i$, where v_i is a random variable statistically independent of x_i, then the term η_i (equal to $1 + v_i$) is statistically independent of x_i, implying that the conditional variance of y_i (and ε_i) is proportional to $exp(2\beta x_i)$.

[8] The excluded restriction variable is generated as: $\xi_i = y_{2i} - [\, exp(\beta x_i) * \eta_i^*]$.

[9] With $\varepsilon_i = exp(\beta x_i) v_i$, where v_i is a random variable statistically independent of x_i, then the term η_i (equal to $1 + v_i$) is statistically independent of x_i, implying that the conditional variance of y_i (and ε_i) is proportional to $exp(2\beta x_i)$.

where $E_j^k\,/Y^k$ is the importing country *j*'s share of the world spending on *k*; $Y_i^k\,/Y^k$ is the exporting country *i*'s share of the world's sales of goods class *k*; *Border, Language* and *Colony* are dummy variables taking the value of 1 for a pair of countries showing, respectively, common border, language and colonial ties, and zero otherwise; $\sigma > 1$ is the elasticity of substitution across goods in *k*; $pref_{ij,t}^k$ is the preferential margin as eq.(1); P_j^k and Π_i^k are the MRTs; and D_t are time dummies.

All coefficients are set to 1 but the elasticity of substitution, σ, is set equal to 8, a value within the range of the estimates included in the survey by Anderson and van Wincoop (2004).

3.2 Simulations

After generating *Trade flow* ($X_{ij,t}^k$) we use a Heckman two-step estimator with times dummies, we estimate the following model:

$$\ln(X_{ij,t}^k) = \beta_0 + \beta_1 \ln E_{j,t}^k + \beta_2 \ln Y_{i,t}^k - \beta_3 \ln Y_t^k + \gamma_1 lndist_{ij} + \gamma_2 border_{i,j} + \gamma_3 language_{ij}$$
$$(\sigma - 1)\ln pref_{ij,t}^k + (\sigma - 1)\ln P_{j,t}^k + (\sigma - 1)\ln \Pi_{i,t}^k + \gamma_4 D_t + e_{ij,t}^k\,,\ select\ (E_{j,t}^k,$$
$$Y_{i,t}^k,\ Y_t^k,\ distance_{ij,}\ border_{i,j},\ language_{ij}\,,\ colony_{ij}\,,\ pref_{ij,t}^k,\ P_{j,t}^k,\ \Pi_{i,t}^k,\ \xi_i,\ Dt) \qquad (6),$$

and replicate each regression 10,000 times to get the set of $\hat{\beta} = (\hat{\sigma} - 1)$ and its standard errors. In the first column of Table 1 we show the statistics of the distribution of estimated coefficients ($\hat{\beta}$) obtained using the continuous policy variable, $\ln pref_{ij,t}^k$, assuming that all variables in equation (6) are known (Model 1).

In Model 2 we drop the price indices, P_j^k and Π_i^k, to confirm that the omission of the MRT leads to biased estimates, estimating the following regression:

$$\ln(X_{ij,t}^k) = \beta_0 + \beta_1 \ln E_{j,t}^k + \beta_2 \ln Y_{i,t}^k - \beta_3 \ln Y_t^k + \gamma_1 lndist_{ij} + \gamma_2 border_{i,j} + \gamma_3 language_{ij} +$$
$$(\sigma - 1)\ln pref_{ij,t}^k + \gamma_4 D_t + e_{ij,t}^k\,,\ select\ (E_{j,t}^k,\ Y_{i,t}^k,\ Y_t^k,\ distance_{ij,}\ border_{i,j},$$
$$language_{ij}\,,\ colony_{ij}\,,\ pref_{ij,t}^k,\ \xi_i, Dt) \qquad (7).$$

Then, to proxy for the MRT we add different sets of fixed effects (FE):

$$\ln(X_{ij,t}^{k}) = \beta_0 + \beta_1 \ln E_{j,t}^{k} + \beta_2 \ln Y_{i,t}^{k} - \beta_3 \ln Y_t^{k} + \gamma_1 lndist_{ij} + \gamma_2 border_{i,j} + \gamma_3 language_{ij} + (\sigma - 1) \ln pref_{ij,t}^{k} + \gamma_4 D_t + FE + e_{ij,t}^{k}\ , \ select\ (E_{j,t}^{k},\ Y_{i,t}^{k},\ Y_t^{k},\ distance_{ij,} border_{i,j},\ language_{ij}\ , colony_{ij}\ , pref_{ij,t}^{k},\ \xi_i,\ Dt, FE) \quad (8).$$

Models from 3 to 7 differ in terms of the fixed effects structure, that is:

- Model 3 includes importer, exporter, and product dummies, i.e. $FE = \varphi_i + \theta_j + \omega_k$;
- Model 4 includes importer-, and exporter-, product-specific dummies $FE = \varphi_i^k + \theta_j^k$;
- Model 5 adds to model 3 time invariant exporter-by-importer (bilateral) interaction effects (Egger and Pfaffermayr, 2003), i.e., $FE = \delta_{ij} + \omega_k$;
- Model 6 includes importer and exporter time variant dummies in addition to the product-specific dummies, i.e., $FE = \varphi_{i,t} + \theta_{j,t} + \omega_k$.
- Model 7 introduces the full-blown specifications including importer-, and exporter-, product-specific time variant dummies, i.e. $FE = \varphi_{i,t}^k + \theta_{j,t}^k$.

Independently from the presence of time-varying dummies, time dummies are included in all models.

3.3 Results for preference margin

As expected, in model 1 of the estimated coefficient is $\hat{\beta} = 7$ that is the one assumed in the data-generating process (Table 1). The estimated coefficient of the variable ln($pref_{ij,t}^{k}$) implies an elasticity of substitution equal to 8 ($\sigma = 7 + 1$). The marked homogeneity of the estimates is confirmed by the low value of the coefficient of variation (the ratio between the standard deviation and the mean is around 0.8). Model 2 confirms that the omission of MRT leads to biased estimates. The estimated coefficients are higher on average (7.67) and statistically different from coefficients in Model 1. We use results of Model 1 as the benchmark to assess the consequences of compounding the error in measuring the policy variable with a more or less complete specification of the fixed effects structure (model 3-7).

Table 1: ***Estimated coefficients of the policy variable: preference margin. The omission of MRT***

Policy Variable: $ln\ (Pref^{k}_{ijt})$	MODEL 1	MODEL 2
Simple mean	7.00	7.67
Std. Deviation	0.58	1.24
Median	7.03	7.67
5^{th} percentile	6.06	5.78
95^{th} percentile	7.86	9.52

Notes: Statistics of the distribution of estimated coefficients obtained using Heckman two-step Model. Dependent variable: ln ($X^{k}_{i,j,t}$). All independent variables are specified in eq. (6). In model 2 multilateral price indices are excluded (eq.(7)). Time Dummies included in all models. Simulations with 10.000 replications.

Table 2 shows the consequences of modelling the MRT with various structures of fixed effects. In Model 3 we introduce time invariant exporter-, importer- and product-fixed effects: the average of estimated coefficients of the policy variable is 7.09; in Model 4 we interact the exporter and importer fixed effects with the product fixed effects and we get the average of the estimated coefficients equal to 7.24; finally, specification of Model 5 adds the time invariant bilateral fixed effects, as suggested by Egger and Pfaffermayr (2003), and leads to the average result 7.14. These results differ significantly from Model 1[10], and the higher coefficients of variation imply a potentially large under/overestimation of the trade policy impact. Surprisingly, the more complete structure of fixed effects does not improve the performance of Model 2, since the addition of fixed effects does not seem to correct the bias due to the omission of MRT, only reducing the overestimation on average. However, the difference with respect to the benchmark is still significant, and this implies that working with disaggregated data is better to control for the (unobserved) product heterogeneity than for the bilateral one.

[10] For each model we employ a t-test to test the null hypothesis that the difference between the means of parameter of interest is equal to zero. We reject the null hypothesis in all experiments.

Table 2: ***Estimated coefficients of the policy variable: preference margin with FE***

Policy Variable: *ln* $(Pref^{k}_{ijt})$	*Time- invariant FE*			*Time- varying FE*	
	MODEL 3	MODEL 4	MODEL 5	MODEL 6	MODEL 7
Simple mean	7.09	7.24	7.14	6.71	-3.53
Std. Deviation	0.70	1.40	0.90	0.75	932.95
Median	7.10	7.22	7.13	6.72	8.72
5^{th} percentile	5.95	5.14	5.70	5.47	-112.77
95^{th} percentile	8.22	9.45	8.58	7.90	119.63
Importer	Yes				
Exporter	Yes				
Product	Yes		Yes	Yes	Yes
Importer*Product		Yes			
Exporter*Product		Yes			
Importer*Exporter			Yes		
Importer*time				Yes	
Exporter*time				Yes	
Importer*product*time					Yes
Exporter*product*time					Yes

Notes: Statistics of the distribution of estimated coefficients obtained using Heckman two-step Model. Dependent variable: ln $(X^{k}_{i,j,t})$. All independent variables are specified in eq. (8). Time Dummies included in all models. Simulations with 10.000 replications.

Model 6 and Model 7 represent the (in principle) more accurate approximation of the MRT based on time-varying fixed effects specifications. However, the introduction of time varying fixed effects in models 6 and 7 does not improve the accuracy of the estimate. In Model 6, the average impact is significantly lower (6.71) than our benchmark, and the variability of the estimates increases. Since the trade of the preferential policy is slightly underestimated, it appears that the time-varying fixed effects in addition to the MRT also tend to absorb part of the trade policy effect.

Things get even worse in the case of Model 7 which provides the most complete specification of the fixed effects structure. The mean of the estimated impacts is negative but the distribution presents a very large standard deviation (around 933), However, the central value is 8.72, much higher respect to results of previous models. Apparently, the country-product-time dummies lead to very unrealistic estimates, likely due to high multicollinearity.

3.4 Results for preference dummy

In order to shed some light on the capability of the dummy variable to proxy the policy effect, we use the bilateral trade flows generated by equation (5) and re-run the regression including a dichotomous policy variable, $dummypref_{ij,t}^{k}$ instead of $\ln(pref_{ij,t}^{k})$. In Table 4 we show the descriptive statistics of the estimated impact of the preference dummy.

Under the assumption that all variables are known, Model 1 generates an estimated coefficient of the dummy variable equal to 0.220 on average. The results for Model 1 in Table 1 and Table 3 are not comparable: the estimated coefficient of the policy dummy represents the total effect, while the estimated effect size of the preference margin represents the marginal one (elasticity). However, the policy dummy cannot provide an accurate assessment of policies that, by definition, often discriminate among countries and products, and this is likely to lead to a large overestimation of the impact of preferential schemes. Moreover, results are characterized by a high coefficient of variation implying a considerable risk of under/overestimating the trade policy impact.

Table 3: ***Estimated coefficients of the policy variable: preference dummy. The omission of MRT***

Policy Variable: dummyPref$^{k}_{ijt}$	MODEL 1	MODEL 2
Simple mean	0.22	0.50
Std. Deviation	0.22	0.38
Median	0.22	0.49
5^{th} percentile	-0.09	-0.04
95^{th} percentile	0.55	1.06

Notes: Statistics of the distribution of estimated coefficients obtained using Heckman two-step Model. Dependent variable: ln $(X^{k}_{i,j,t})$. All independent variables are specified in eq. (6), dummypref$^{k}_{ij,t}$ instead of ln(pref$^{k}_{ij,t}$). In models 2 multilateral price indices are excluded (eq.(7)), dummypref$^{k}_{ij,t}$ instead of ln(pref$^{k}_{ij,t}$). Time Dummies included in all models. Simulations with 10.000 replications.

When we include $dummypref_{ij,t}^{k}$ instead of $\ln(pref_{ij,t}^{k})$ in equation (7), it is still the case that the omission of the MRT leads to an overestimation of the policy impact (Model 2 in

Table 4). On the other hand, in this case the introduction of the time invariant fixed effects does not lead to more accurate estimates. Table 4 shows that the specifications provided by models 3, 4 and 5 lead to results that are not significantly different from those of Model 2 (with a statistically significant level of 1%)

Table 4: ***Estimated coefficients of the policy variable: preference dummy with FE***

Policy Variable: dummy$Pref^{k}_{ijt}$	MODEL 3	MODEL 4	MODEL 5	MODEL 6	MODEL 7
Simple mean	0.45	0.51	0.49	0.42	1.86
Std. Deviation	0.27	0.61	0.42	0.30	135.65
Median	0.45	0.51	0.49	0.42	0.50
5th percentile	0.01	-0.44	-0.17	-0.07	-10.90
95th percentile	0.90	1.50	1.18	0.90	12.57
Importer	Yes				
Exporter	Yes				
Product	Yes		Yes	Yes	Yes
Importer*Product		Yes			
Exporter*Product		Yes			
Importer*Exporter			Yes		
Importer*time				Yes	
Exporter*time				Yes	
Importer*product*time					Yes
Exporter*product*time					Yes

Notes: Statistics of the distribution of estimated coefficients obtained using Heckman two-step Model. Dependent variable: ln $(X^{k}_{i,j,t})$. All independent variables are specified in eq. (8), $dummypref^{k}_{ij,t}$ instead of ln$(pref^{k}_{ij,t})$. Time Dummies included in all models. Simulations with 10.000 replications.

The situation does not improve even when we introduce the time-varying dimension: the average impact is still significantly overestimated in Model 6, while it is confirmed that the full structure of time varying fixed effects in Model 7 is marred by high collinearity and leads to an unrealistic overestimation of the average impact.

4. Conclusions

The assessment of the impact of trade policies on trade flows is at the hearth of a large literature using the gravity model. Given the large volume of data available to gravity modellers, many different specifications of the gravity equation can be “made to work" in the sense that they generate statistical tests that tend to corroborate the specification’s validity. More specifically, the majority of the recent significant contributions on the matter employs exporter and importer fixed sector-specific effects. We use Monte Carlo experiments to show that such a choice, together with measurement errors about the policy treatment, leads to bias in the estimates of the policy impact.

This paper uses hypothetical but plausible data from a known data-generating process to evaluate which models most closely and frequently draw correct conclusions from the underlying structure of the data. In the gravity model literature, it has become increasingly common to include fixed effects to account for the MRT. Our main contribution is to point out that the identification of trade policy effects with a gravity equation that includes fixed effects to control for the multilateral trade resistance terms has limitations. This is especially true if the policy under investigation is not accurately measured: in such a case, the usual fixed effects fix for the omitted (MRT) variable problem could make the consequences of the measurement error even worse.

On the one hand, our results confirm that omission of a crucial variable such as multilateral resistance causes an inaccurate assessment of the policy impact In particular (in Model 2), the trade effect of preferences is on average overestimated if a continuous variable is used, as well as if the policy is measured by a dummy. On the other hand, when we control for country-product and time unobserved heterogeneity (Model 4), the estimated parameter approaches the true (assumed) value of the policy treatment. As a matter of fact, this is true only if the policy treatment is correctly measured by a continuous variable: fixed effects do not solve the problem of the lack of the MRT when we have an error in the measurement of the policy variable (Table 4).

The most striking result is that introducing time varying fixed effects (i.e. the complete specification to proxy for the MRT) worsens the accuracy of the estimate of the policy variable. Apparently, the more we control for time varying fixed effects, the more the estimated parameter is biased (models 6 and 7) since the range of the estimates is greatly inflated.

The bottom line is that a full-blown fixed effects specification risks to severely limit the assessment of trade policy effects.

The situation is even worse if we proxy a policy variable with a dummy instead of a continuous one. This amounts to combine a (policy) measurement problem with a (MRT) omitted variable problem: in such a case any attempts to control for heterogeneity through a fixed effect design would prevent us from evaluating trade policy effects due to the presence of high collinearity.

References

Anderson, J.E. (1979) A theoretical foundation for the gravity equation. *The American Economic Review* 69 (1): 106–116.

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

Anderson, J.E., van Wincoop, E. (2004) Trade Costs. *Journal of Economic Literature*, 42(3): 691-751.

Baier, S., Bergstrand, J. (2009) Estimating the Effects of Free Trade Agreements on International Trade Flows using Matching Econometrics. *Journal of International Economics*, 77(1): 63–76.

Baldwin, R., Taglioni, D. (2006) Gravity for Dummies and Dummies for Gravity Equations. NBER Working Paper 12516.

Baltagi, B.H., Egger, P., Pfaffermayr, M. (2003) A generalized design for bilateral trade flow models. *Economic Letters*, 80: 391-397.

Baltagi, BH, Egger, PH, Pfaffermayr, M. (2015). Panel data gravity models of international trade. In Baltagi BH (ed.) The Oxford Handbook of Panel Data. Oxford, UK: Oxford University Press, 608–641.

Bergstrand, J. (1985) The gravity equation in international trade: some microeconomic foundations and empirical evidence. *The Review of Economics and Statistics* 67 (3): 474–481.

Bergstrand, J. H., Egger, P., Larch, M. (2013) Gravity Redux: Estimation of gravity-equation coefficients, elasticities of substitution, and general equilibrium comparative statics under asymmetric bilateral trade costs. *Journal of International Economics* 89(1): 110-121.

Caliendo, L., Parro, F. (2015) Estimates of the trade and welfare effects of NAFTA. *The Review of Economic Studies* 82 (1): 1-44, Cardamone, P. (2011) Preferential Trade Agreements Granted by the European Union: an Application of the Gravity Model Using Monthly Data. *European Review of Agricultural Economics* 38 (4): 553-586.

Chaney, T. (2008) Distorted Gravity: The Intensive and Extensive Margins of International Trade. *The American Economic Review* 98 (4): 1707–21.

Cipollina, M., Pietrovito, F. (2011) Trade Impact of EU Preferential Policies: a Meta-analysis of the Literature, in Luca De Benedictis and Luca Salvatici (Eds.), The Trade Impact of European Union Preferential policies: an Analysis through gravity models, Berlin/Heidelberg: Springer, 2011.

Cipollina, M., Salvatici, L. (2010) The Impact of European Union Agricultural Preferences. *Journal of Economic Policy Reform*, 13 (1): 87-106.

Cipollina, M., Laborde, D., Salvatici, L. (2016). The tide that does not raise all boats: An assessment of EU preferential trade policies. *Review of World Economics*, pp.1-33, DOI 10.1007/s10290-016-0270-0

De Benedictis, L., Taglioni, D. (2011) The Gravity Model in International Trade, in Luca De Benedictis and Luca Salvatici (Eds.), The Trade Impact of European Union Preferential policies: an Analysis through gravity models, Berlin/Heidelberg: Springer, 2011.

De Benedictis, L., Salvatici, L., eds (2011), The Trade Impact of European Union Preferential policies: an Analysis through gravity models, Berlin/Heidelberg: Springer.

De Benedictis, L., Vicarelli, C. (2005) Trade Potential in Gravity Data Models. *Topics in Economic Analysis and Policy*, 5, (1).

Eaton, J., Kortum, S. (2002) Technology, Geography, and Trade. *Econometrica* 70 (5): 1741–1779.

Egger, P., Pfaffermayr, M. (2003) The proper panel econometric specification of the gravity equation: A three-way model with bilateral interaction effects. *Empirical Economics* 28: 571–580.

Egger, P., Staub, K. E. (2016) GLM Estimation of Trade Gravity Models with Fixed Effects, *Empirical Economics* 50 (1): 137–175

Fally, T. (2015) Structural Gravity and Fixed Effects *Journal of International Economics* 97 78-85

Feenstra, R.C. (2002) Border effects and the gravity equation: Consistent methods for estimation. *Scottish Journal of Political Economy* 49 (5): 491–506

Feenstra, R.C. (2004) Advanced International Trade: Theory and Evidence, Princeton University Press, ch.3.

Head, K.,Mayer, T. (2014) Gravity Equations: Workhorse, Toolkit, and Cookbook, in the Handbook of International Economics, Vol. 4, eds. Gopinath, Helpman, and Rogoff, Elsevier.

Head, K., Mayer, T., Ries, J. (2010) The erosion of colonial trade linkages after independence. Journal of International Economics, 81(1): 1-14

Head, K., Ries, J. (2001) Increasing Returns Versus National Product Differentiation as an Explanation for the Pattern of U.S.-Canada Trade. American Economic Review 91 (4): 858-876

Heckman, J.(1979) Sample selection bias as a specification error. Econometrica 47: 153-61.

Helpman, E., Melitz, M., Rubinstein, Y. (2008) Estimating Trade Flows: Trading Partners and Trading Volumes. Quarterly Journal of Economics 123 (2): 441–487.

Hornok, C. (2011) Gravity or Dummies? The Limits of Identification in Gravity Estimations, CeFiG Working Papers 15.

Hummels, D. (2001).Toward a Geography of Trade Costs. Unpublished manuscript, available at http://www.krannert.purdue.edu/faculty/hummelsd/research/toward/TGTC.pdf.

Leamer, E.E., Levinsohn, J. (1995) International Trade Theory: The Evidence, in G. Grossman and K. Rogoff, (Eds.), Handbook of International Economics, Vol III, 1995, pp. 1339-1394

Martin, W., Pham, C. (2014) Estimating the Gravity Model When Zero Trade Flows Are Frequent and economically determined, Mimeo.

Martinez-Zarzoso, I. (2013) The log of gravity revisited. Applied Economics 45(3): 311-327.

Màtyàs, L. (1997) Proper Econometric Specification of the Gravity Model. The World Economy, 20(3): 363-368.

Novy, D. (2013) Gravity Redux: Measuring International Trade Costs With Panel Data. Economic Inquiry, 51(1): 101–121.

Olivero, M.P., Yotov. V. Y. (2012) Dynamic Gravity: Endogenous Country Size and Asset Accumulation.Canadian Journal of Economics, 45(1): 64–92.

Piermartini, R., Yotov, V.Y. (2016). Estimating Trade Policy Effects with Structural Gravity. LeBow College of Business, Drexel University School of Economics Working Paper Series WP 2016-10.

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

Santos Silva, J.C. M., Tenreyro, S. (2006) The Log of Gravity. The Review of Economics and Statistics 88: 641-58.

Santos Silva, J.C. M., Tenreyro, S. (2011) Further simulation evidence on the performance of the Poisson pseudo-maximum likelihood estimator. Economics Letters 112: 220-222.

Stack, M. M. (2009), Regional Integration and Trade: Controlling for Varying Degrees of Heterogeneity in the Gravity Model. The World Economy, 32(5): 772-789.

Wei, S. J. (1996) Intra-National versus International Trade: How Stubborn are Nations in Global Integration? NBER Working Papers No 5531.

APPENDIX

Table A1. Descriptive statistics: Initial dataset.

Variable	Mean	St. dev.	Min	Max
Bilateral flow of trade (1000$)	42,528	253,218	0	13,700,000
Production (1000$)	3,313	7,065	0.001	75,229
Expenditure (1000$)	32,646	70,724	96	698,130
Distance	6,856	4,890	260	18,310
Language	0.05	0.22	0	1
Border	0.04	0.20	0	1
Exporter Price index	1.04	0.18	0.42	3.17
Importer Price index	1.03	0.09	0.5215	1.55
Preference margin	1.05	0.16	1	4.91
Preference dummy	0.80	0.40	0	1

Notes: Number of Obs. 21,896. Descriptive statistics are calculated on variables in levels.