CESifo Working Papers

www.cesifo.org/wp

The Determinants of Trade Costs: A Random Coefficients Approach

Peter Egger
Jan Průša

CESifo Working Paper No. 4859
Category 8: Trade Policy
June 2014

CESifo *Working Paper No. 4859*

The Determinants of Trade Costs: A Random Coefficients Approach

Abstract

This paper assesses whether the sensitivity of bilateral trade volumes to various trade cost factors is constant or varies across countries. It utilizes a random coeffcients model and analyses a cross-sectional sample of bilateral trade data for 96 countries in 2005. We expect the elasticity of trade to vary particularly with bilateral distance and bilateral tariffs due to measurement error about these factors. Indeed, the variability of coefficients is significant for these trade cost measures. The results indicate that the elasticity of trade with respect to tariffs in different countries varies relatively more than that with respect to distance. This is consistent with there being a host of sources of measurement error about bilateral tariffs (due to strategic or non-strategic misreporting; the potential inappropriateness of the weighting of disaggregated tariffs; etc.).

JEL-Code: C290, F120, F170.

Keywords: bilateral trade flows, gravity equation, random coefficients model, trade costs.

Peter Egger
Department of Management,
Technology, and Economics
ETH Zurich
WEH E6
Weinbergstrasse 35
Switzerland – 8092 Zurich
egger@kof.ethz.ch

Jan Průša
Institute of Economic Studies
Charles University in Prague
Opletalova 26
Czech Republic – 11000 Prague
jan.prusa@ies-prague.org

June 8, 2014
The authors acknowledge financial support from the Czech Science Foundation, grant number P402/12/0982.

1 Introduction

Research on gravity models of international trade is booming (see the surveys by Baltagi, Egger, and Pfaffermayr, 2014; and Head and Mayer, 2014). While earlier empirical work tended to be ad-hoc from a theoretical perspective, Anderson and van Wincoop's (2003, 2004) attempt was the first to fully integrate a structural multi-country gravity model of bilateral trade in general equilibrium with estimation. Since then, gravity models are used to estimate the importance of trade costs and to assess their impact not only on trade flows but also on real consumption (welfare). While such structural models obtain heterogeneous responses of a uniform change in trade costs across country-pairs and countries, it is commonly assumed that the so-called elasticity (or semi-elasticity) of observable trade costs – i.e., the parameter on log distance and a set of binary variables reflecting cultural, economic, institutional, or historical proximity – is invariant across countries and country-pairs.

This paper proposes an alternative approach, where at least observable (and, eventually, even unobservable) measures of trade costs may exert a direct effect on bilateral trade that is variable. It is well known that the measurement of many variables entering gravity-model specifications is imprecise (see Egger and Nigai, 2014): tariff rates may be improperly applied to the wrong product lines due to fraud (Fisman and Wei, 2004) or may be badly recorded; distances between countries pertain to some arbitrary points in space (capital cities or economic centers) which may coincide with origins and destinations of exports and imports or not (see Hillberry and Hummels, 2008); adjacency might matter differently for country-pairs with longer or shorter borders (see Engel and Rogers, 1996); common official language may mis-represent the importance of the actual overlap of (native or non-native) speakers between two countries (see Melitz and Toubal, 2014); domestic as opposed to foreign trade costs might be mis-measured (see Rousslang and To, 1993). These issues of measurement error as well as specific responses per se might lead to heterogeneous direct effects of trade impediments on trade flows, which are typically assumed to be the same. Hence, heterogeneous responses of country-pairs and countries to shocks in trade costs may have two sources: ex-ante asymmetries of countries in their fundamental drivers of trade (levels of trade costs, factor endowments, productivity, and preferences; this represents the *old view*); and (country- or pair-)specific direct responses (representing a *new view* addressed in this paper). The goal of the present paper is precisely to assess the scope of the latter by means of a random effects model.

The remainder of the paper is organized as follows. The subsequent section lays out the theoretical arguments for the approach adopted in the application. Section 3 introduces the econometric model. Section 4 summarizes the findings from an application of the approach to cross-sectional data, and the last section concludes.

2 Normalized trade flows and the trade cost ratio

Let us use indices $i, j \in \{1, ..., J\}$ to denote exporting and importing countries and X_{ij} to denote aggregate exports of country i to j or imports of country j from i. Then, following Baltagi, Egger, and Pfaffermayr (2014) or Head and Mayer (2014), we may write normalized exports for country-pair ij as

$$\frac{X_{ij}X_{ji}}{X_{ii}X_{jj}} \equiv \frac{T_{ij}T_{ji}}{T_{ii}T_{jj}}\frac{U_{ij}U_{ji}}{U_{ii}U_{jj}} \quad \text{for all} \quad i \neq j, \tag{1}$$

$$T_{ij} \equiv \prod_{D\in\mathfrak{T}} D_{ij}^{\beta_D} \quad \text{for all} \quad i, j, \tag{2}$$

where U_{ij} is a non-negative disturbance term which has mean one, D_{ij} are individual trade cost measures which enter with respective (unknown) parameters β_D, and $\mathfrak{T}$ is the set of all trade cost measures. Using the notation $\tilde{v}_{ij} \equiv (v_{ij} + v_{ji} - v_{ii} - v_{jj})$ for any generic variable v_{ij} in logs and the convention that lower-case letters denote logs of variables such that $v_{ij} = \ln(V_{ij})$, the specific set of trade costs in logs employed in this paper is defined as

$$\mathfrak{t} \in \{\underbrace{\tau}_{\text{ln one-plus-tariff-rate}}, \underbrace{g}_{\text{ln geogr. distance}}, \underbrace{a}_{\text{adjacency}}, \underbrace{l}_{\text{common language}}, \underbrace{\alpha}_{\text{domestic sales}}\} \tag{3}$$

Let us assume that the normalized disturbance term exhibits a general error components structure of the form[1]

$$\tilde{u}_{ij} = \mu_i + \lambda_j + \epsilon_{ij}, \tag{4}$$

so that we may write the log-transformed counterpart to (1) as

$$\tilde{x}_{ij} = \beta_\tau \tilde{\tau}_{ij} + \beta_a \tilde{a}_{ij} + \beta_g \tilde{g}_{ij} + \beta_l \tilde{l}_{ij} + \mu_i + \lambda_j + \epsilon_{ij} \tag{5}$$

$$\text{for all} \quad i \neq j,$$

$$t_{ij} \equiv \sum_{D \in \mathfrak{t}} \beta_D \tilde{d}_{ij} + \alpha_i 1(i = j) \quad \text{for all} \quad i, j. \tag{6}$$

Notice that we generally set $\tau_{ij} = 0$, $a_{ij} = 1$, $l_{ij} = 1$, whenever $i = j$ (i.e., countries do not charge tariffs for domestic sales, such sales happen within national borders, and sellers and consumers speak the same official language), while we assume that internal distances are positive, and $g_{ij} > 0$ for $i = j$. Moreover, notice that $\{g, a, l\}$ are symmetric such that $v_{ij} = v_{ji}$, whereas τ is generally not such that $\tau_{ij} \gtreqless \tau_{ji}$. Accordingly,

$$\tilde{\tau}_{ij} \equiv \tau_{ij} + \tau_{ji}, \; \tilde{g}_{ij} \equiv 2g_{ij} - g_{ii} - g_{jj}, \; \tilde{a}_{ij} \equiv 2a_{ij} - 2, \; \tilde{l}_{ij} \equiv 2l_{ij} - 2. \tag{7}$$

[1] This assumption means that the error components pertaining to countries (as exporters or importers) accrue to international trade flows only, while they do not emerge for domestic sales. Certainly, one might adopt alternative sets of assumptions which might lead to even more complicated error components structures. For instance, one could assume a structure, where country-specific error components could differ for a country as an exporter, and importer, and a domestic seller.

With regard to the intercept for trade costs on intranational sales, we obtain

$$\tilde{\alpha}_{ij} \equiv \begin{cases} -\alpha_i - \alpha_j & \text{if } i \neq j \\ 0 & \text{otherwise} \end{cases} . \quad (8)$$

3 A random coefficients gravity model

Rather than assuming the parameters β_D for $D \in \{\tau, a, g, l\}$ to be invariant scalars, as is common in the literature on empirical gravity models, we specify them as

$$\beta_D \equiv \beta_D^0 + \nu_i^D + \xi_j^D, \quad (9)$$

where β_D^0 is a fixed (average, invariant) scalar. Components ν_i and ξ_j represent random fluctuations of the average coefficient β_D^0 with respect to exporting country i and importing country j, respectively. We assume that ν_i and ξ_j are identically and independently distributed (i..i.d.) random variables with zero mean and variances σ_ν^2 and σ_ξ^2, respectively.

In a more general case, β_D might even depend on a random pair-specific term η_{ij}, which could be identified, if more observations on pairs ij existed than parameters η_{ij} had to be estimated (e.g., this would be the case if η_{ij} varied across exporter-continent and importer-continent).[2]

The purpose of the random coefficients model is to identify systematic fluctuations in estimated regression coefficients. Such fluctuations are likely to occur especially in models where regressors do not have direct economic

[2]With as many parameters η_{ij} as country-pairs, η_{ij} is not identified and is captured in the random error term ϵ_{ij}.

meaning, are mis-measured, and are only proxies for true economic variables. Because proxy variables provide only approximate information on the economic process in the background, they are likely to contain noise (measurement error) which can be separated from the systematic variation when imposing a set of suitable assumptions. In general, Egger and Nigai (2014) provide evidence that parameterized trade cost functions appear to generally lead to sizable measurement errors with aggregate gravity models of bilateral international trade. In the introduction, we gave examples of earlier work that pointed to measurement problems with specific observable trade cost variables.

In this paper, we examine the presence of noise or measurement error in the two most important variables in gravity models: tariffs and geographical distance between partner countries. At the aggregate country-pair level, there are numerous reasons for why tariffs may be mis-measured. First, except for intra-union trade within customs unions, there is no such thing as a tariff on aggregate imports, but tariffs are levied at relatively disaggregated trade lines. Some of those trade lines are not recorded for specific country pairs for reasons of data confidentiality (e.g., that is true for trade of some commodities, arms, etc.) so that a properly weighted aggregate bilateral tariff cannot be constructed. Moreover, product-level trade flows may be mismeasured so that even properly-measured tariff rates cannot be weighted suitably. Apart from that, tax fraud leads to applications of the wrong tariffs to certain trade lines as addressed in the introduction. With geographical distance, the problem is that exports do neither originate from a single point in a country nor do they go to a single point. Hence, there is measurement error of distances, in particular, for large, adjacent economies (such as the United States and Canada; the United States and Mexico; Germany and Poland; Brasil and Argentina; China and India; etc.). Moreover, distance should matter differently depending on the transport means used (airplanes, railway, ships, trucks, etc.).

With these arguments in mind, we specify the random coefficients counterpart to equation (5) as

$$\begin{aligned} \tilde{x}_{ij} &= \beta_\tau \tilde{\tau}_{ij} + \beta_g \tilde{g}_{ij} + \beta_a \tilde{a}_{ij} + \beta_l \tilde{l}_{ij} + \epsilon_{ij} + \\ &\quad \mu_i + \tilde{\tau}_{ij} \nu_i^\tau + \tilde{g}_{ij} \nu_i^g + \\ &\quad \lambda_j + \tilde{\tau}_{ij} \xi_j^\tau + \tilde{g}_{ij} \xi_j^g. \end{aligned} \tag{10}$$

The goal is to estimate the parameters β_D for $D \in \{\tau, a, g, l\}$ as well as the variances of μ_i, λ_j, and of ν_i^ℓ as well as ξ_j^ℓ.

4 Empirical analysis

4.1 Data

In order to estimate the model described in the previous section, we utilize cross-sectional data on bilateral exports for the year 2005. Altogether, there are 96 reporting export countries in the data. Due to some missing observations on bilateral exports, the total number of observations is 3,575.[3]

[3] In a perfectly symmetric data-set without any missing observations, the number of observations would be $96 * 95/2 = 4,560$. We assume that missing observations do not induce bias (i.e., there is no sample selection). Indeed, the paper by Egger, Larch, Staub, and Winkelmann (2010) suggests that gravity models which condition on country-specific effects (as we do by tetradic differencing) do unlikely display sample-selection bias, and

The sources for the variables used in the application are as follows:

- bilateral exports are from the United Nations' Comtrade Database,
- bilateral tariffs are from the World Bank's WITS database,
- the other observable trade cost variables are from the geographical database of the Centre d'Études Prospectives et d'Informations Internationales.

These variables are used to estimate equation (10).

4.2 Parameter estimates and model comparison

Implementation of the random coefficients model is straightforward. For instance, numerous statistical and econometric software packages (such as R or Stata) permit using already implemented routines.[4] The results on four models (labeled A-D) are summarized in Table 1.

Models A and B in Table 1 are standard OLS estimates with pooled observations. Model A assumes that neither the coefficients nor the disturbance term exhibit an error component structure. Model B is the same as A but the variance is adjusted for clusters in both the exporting and importing country dimension based on Cameron, Gelbach, and Miller (2006). The latter assumes an error components structure in the disturbances but not the coefficients.

two-part models can be used for the analysis with zeros. In that sense, we focus on the second part of a two-part gravity model.

[4]For the present paper, we employed the `xtmixed` routine of Stata. This package is designed to fit nested models, when observations from one group (e.g., country-pairs) are nested within larger groups (e.g., continents). `xtmixed` can further be adapted to evaluate crossed models as required here, where groups are in fact not hierarchical but horizontal or crossed (e.g., exporting by importing countries).

Model C adds random effects in the disturbances for both exporters and importers. It can be seen directly that the random effects are significant and further improve the significance of parameters on tariffs and distance when compared to the OLS Model A. A likelihood ratio test comparing Model C with Model A rejects the null hypothesis of zero random effects ($H_0 : \sigma(\mu_i), \sigma(\lambda_j) = 0$) at the 0.001 significance level.

Model D shows the full random coefficient model. The random components associated with the coefficients on tariffs and distance (ν_i, ξ_j) are highly significant. The likelihood ratio tests reveal that the null hypothesis $H_0 : \sigma(\nu_i), \sigma(\xi_j) = 0$ (testing Model D versus C) is rejected at an 0.001 significance level. However the random effect λ_j associated with the recipient (importing) country in each ij-pair is not significant. Hence, one might estimate and present a model which sets λ_j to zero.[5] In any case, Model D is preferred among the ones presented in Table 1.

As expected, the parameters β_D^0 in the random coefficients Model D are very similar to their counterparts in the other models. The main insight lies in the decomposition of the country-specific random effects into randomness associated with specific explanatory variables and remainder randomness. The results show that the variance of bilateral export volume across country pairs is more strongly affected by the variability tariff effects than that of distance effects. The reason could be that average distances between countries are large so that the measurement error about domestic distances through the location of exporting firms is relatively small. The measurement error about tariffs, on the other hand, appears to be large and is systematically re-

[5]We have done so but suppress the respective results for the sake of brevity.

lated to the level of tariffs (e.g., for strategic reasons as suggested in Fisman and Wei, 2004).

5 Conclusions

This paper assesses the presence of random coefficients on observable trade cost measures as employed in standard gravity equations of bilateral trade. The paper derives and formulates such a model, utilizes cross-sectional data from 2005 on all available pairs among 96 countries. The random coefficients model provides a remedy to inaccurate measurement of explanatory variables and easily captures the variability of the parameters across exporting and importing countries.

Respecting the random about trade cost coefficients leads to better inference. The results of this paper provide evidence of a large random component in the coefficient on bilateral import tariffs and a somewhat smaller one in the coefficient of geographical distance. This is aligned with earlier evidence of strategic misreporting of tariffs and a lack of enforcement and of systematic misreporting of product-level trade flows (which are used to weight product-specific bilateral tariffs to obtain an aggregate measure).

6 Bibliography

1. Anderson, James E., and van Wincoop, Eric, 2003. Gravity with gravitas: A solution to the border puzzle. American Economic Review 93(1), 170-192.

2. Anderson, James E., and van Wincoop, Eric, 2004. Trade costs. Journal of Economic Literature 42(3), 691-751.

3. Baltagi, Badi H., Egger, Peter H., and Pfaffermayr, Michael, 2014. Panel data gravity models of international trade. In: Baltagi, Badi H. (ed.), The Oxford Handbook of Panel Data, Chapter 20, forthcoming.

4. Cameron, A. Colin, Gelbach, Jonah H., and Miller, Douglas L., 2006. Robust inference with multi-way clustering. NBER technical working paper 327.

5. Egger, Peter H., Larch, Mario, Staub, Kevin E., and Winkelmann, Rainer, 2011. The trade effects of endogenous preferential trade agreements. American Economic Journal: Economic Policy 3(3), 113-143.

6. Engel, Charles, and Rogers, John H., 1996. How wide is the border? American Economic Review 86(5), 1112-1125.

7. Fisman, Raymond, and Wei, Shang-Jin, 2004. Tax rates and tax evasion: Evidence from "missing imports" in China. Journal of Political Economy 112(2), 471-500.

8. Head, Keith, and Mayer, Thierry, 2014. Gravity equations: Workhorse, toolkit, and cookbook. In: Gopinath, Gita, Helpman, Elhanan, and Rogoff, Kenneth R. (eds.), Handbook of International Economics Vol. 4, forthcoming.

9. Hillberry, Russell, and Hummels, David, 2008. Trade responses to geographic frictions: A decomposition using micro-data. European

Economic Review 52(3), 527-550.

10. Melitz, Jacques, and Toubal, Farid, 2014. Native language, spoken language, translation and trade. Journal of International Economics, forthcoming.

11. Rousslang, Donald J., and To, Theodore, 1993. Domestic trade and transportation costs as barriers to international trade. Canadian Journal of Economics 26(1), 208-221.

Table 1: Results

	A Pooled OLS	B Two-way clustered OLS	C Random effects	D Random coeff.
$\tilde{\tau}$	-1.479	-1.479	-1.601	-1.608
	(.133)	(.463)	(.152)	(.182)
$\tilde{g}$	-.720	-.720	-1.050	-1.066
	(.044)	(.097)	(.043)	(.039)
$\tilde{a}$	2.004	2.004	1.608	1.466
	(.228)	(.224)	(.200)	(.161)
$\tilde{l}$	.746	.746	.645	.649
	(.109)	(.204)	(.101)	(.096)
const.	-28.186	-28.186	-22.534	-22.298
	(.781)	(1.689)	(.796)	(.646)
$\sigma(\mu_i)$			2.200	.074
			(.179)	(.824)
$\sigma(\nu_i^\tau)$				.890
				(.176)
$\sigma(\nu_i^g)$				.105
				(.012)
$\sigma(\lambda_j)$			3.461	.016
			(9.039)	(.407)
$\sigma(\xi_j^\tau)$				.816
				(.125)
$\sigma(\xi_j^g)$				.191
				(.004)
$\sigma(\epsilon_{ij})$			.783	.042
			(39.951)	(.472)
Obs.	3575	3575	3575	3575
R^2	0.1832	0.1832		
Wald χ^2			882.25	1678.76