Comments on Referee Reports - Distance and Border Effects in International Trade: A Comparison of Estimation Methods

Glenn Magerman,* Zuzanna Studnicka† Jan Van Hove‡

May 2, 2016

1 Introduction

The authors would like to thank the Editor and three anonymous referees for the various comments and suggestions on an earlier draft of this paper, which help to significantly improve the quality of this paper. We have carefully considered the remarks by the referees, and we would like to address those in this document.

2 Main comments

2.1 Discussion on dispersion, strengths and weaknesses of each estimation method

Several comments of the referees center around the weighting schemes and the lack of a discussion on the strengths and weaknesses of each estimation method and its impact on the estimated coefficients.

Comments of referees:

1. The authors should be able to go further than the saccharine assessment that all methods have merits and shortcomings. Even if they are unwilling to identify a preferred approach, based on their understanding as a result of doing this research the authors could provide an assessment of weaknesses, biases, etc. of the five estimation techniques. Their order of exposition implies a preference for least squares with dummy variable, but is this really a preference or just an expositional convenience based on most common practice?

2. A discussion of reasons for the fairly large range of estimates (especially PPML and GPML) would be useful.

*KU Leuven - Center for Economic Studies (CES) & Leuven Centre for Irish Studies Address: KU Leuven - University of Leuven, Faculty of Business and Economics, Centre for Economic Studies, Naamsestraat 69, 3000 Leuven, Belgium. E-mail: glenn.magerman@kuleuven.be;

†KU Leuven - Center for Economic Studies (CES) & Leuven Centre for Irish Studies; zuzanna.studnicka@kuleuven.be;

‡KU Leuven - Center for Economic Studies (CES), Leuven Centre for Irish Studies & INFER; jan.vanhove@kuleuven.be

Reply: A summary of differences between the estimation methods is given in Table 1 in the paper. This table discusses the inclusion/exclusion of zero trade flows, the impact of heteroscedasticity on the estimated coefficients and the different weighting schemes. Additional evaluation of small-sample bias and non-linearities are discussed in Section 5. However, we agree that the discussion of the weighting schemes and their impact on the estimates can be more elaborate. We propose a more extended discussion along the following lines.

A large source of discrepancy between the estimation methods is due to differences in the set of first-order conditions of the estimators. The pseudo maximum likelihood estimators PPML and GPML approximate the first-order conditions of the original objective in the NLS method, driven by the goal for easier computation of that objective.

We can nest the non-linear models as follows. Let $\hat{\mu}(x) \equiv exp(X\hat{\beta})$, where X is a vector of regressors and $\hat{\beta}$ is the estimate of β that minimizes the objective (i.e. minimizes the error). We can then write the first-order conditions of NLS, PPML and GPML in vector form as

$$\sum(\mu - \hat{\mu})\hat{\mu}^k X = 0 \tag{1}$$

where for NLS: $k = 1$, PPML: $k = 0$ and GPML: $k = -1$ respectively, and $\hat{\mu}^k X$ represents the weighting scheme of the errors to be minimized. Hence, from eq. 1, it is immediately visible that NLS gives a weight to deviations from the mean proportional to that mean. PPML puts equal weights on all deviations, and GPML allocates weights that are inversely proportional to the mean. Depending on the location of the outliers, this can lead to an upward or downward adjustment of the estimated coefficient.

For instance, consider a toy example with one single outlier in trade-distance space (e.g. one very distant island that trades with only one other country in the world). It is easy to imagine the downward sloping regression line in this space. Suppose first that this island trades very little with its partner (so that the observation lies in the North-West of the trade-distance space). Then this outlier will pull the regression line towards itself, leading to a more negative coefficient on distance, and more so for the NLS scheme than the other weighting schemes. However, if this island trades a lot with its partner (now being in the North-East of the trade-distance space), again it will "pull" the coefficient towards itself, but now leading to a less negative coefficient. Considering then that various observations can be in various locations, makes it clear that it is impossible to predict *a priori* the net impact on the size of the estimated coefficient.

Depending on the setting, researchers might then consider which scheme best represents the observed data. For instance, if one assumes that large GDP countries better report GDP values (i.e. with less measurement error), one can opt for the NLS method, so that these countries get more weight in estimating the GDP coefficient of the gravity model. At the same time, the researcher has to be aware that this choice skews the point estimate towards the impact of large GDP countries on bilateral trade flows. Furthermore, this choice simultaneously affects the estimation of all variables, which could involve contrasting interests in the direction of the weighting scheme.

2.2 Estimation of border effect across and within continents

Comments of referees:

1. The "border effect" is universally seen in terms of national borders, with a useful hypothesis being whether the concept should be adjusted for the European Union (or other customs unions/ economic unions), but nowhere (even in the paper's own literature survey) is the discussion of continental "borders". In the literature on global value chains there is evidence of regional value chains, suggesting some diminution of national border effects within continents, but the RVCs do not apply to the Americas or Asia as a whole.

2. Another point that is not clear is why the authors choose border and distance at continental level, no explanation is provided for this. One might expect large heterogeneities between USA and Europe, but also Sub-Saharian or between countries with different level of income –I.e. developing and developed countries. Therefore, I suggest the authors to offer additional evidence on how border and distance impact trade flow by considering additional countries heterogeneities – i.e. democracy – since considering continents is hardly enough.

3. I would also looked at border and distance effects within a country and contrast with its trade with outside of country trade.

Reply: This is a great remark, and we acknowledge that the mainstream literature centers around the within-country effect of regional borders (e.g. Anderson and van Wincoop (2003)). Using the BACI/CEPII data, there is unfortunately no information on within-country trade, which is the main reason why we opted for intra-continental trade. A useful observation is that intra-EU trade might be comparable to intra-US trade: there is a functional internal market without tariffs, leading to relatively higher intra-EU trade, comparable to intra-US trade in Mc Callum (1995).

At the same time, there is a large set of papers that indicate that the border effect exists on several levels, with similar implications: within-country across states/provinces (Mc Callum (1995), Anderson and van Wincoop (2003)), within states across counties, and even within cities across districts. We add to the literature by considering within-continental trade and its border effect, but acknowledge that there are several sources of heterogeneity to be accounted for, such as global value chains, North versus South America, Sub-Saharian vs other African countries etc.

We propose the following additions:

1. Re-estimate using North versus South America, Sub-Saharian vs other African countries and ASEAN vs. other Asian countries to allow for more heterogeneity to be captured by the model.

2. Regarding other country heterogeneities, we consider these captured in the various control variables and/or in the fixed effects if they are unobserved. Note that the estimated coefficient of the continental border effect is then conditional on these other covariates. With respect to the democracy index proposed in particular, we have looked at the publicly available Freedom

House Index. The quality of the data is however fairly poor and does not allow us to construct a balanced panel.

3. Combined, we will reconsider and extend our discussion of the border effect to include: (i) adjacency/common borders, and (ii) intra-continental borders.

3 Other comments

3.1 Preference of estimation method

Comments of referees: The authors should be able to go further than the saccharine assessment that all methods have merits and shortcomings. Even if they are unwilling to identify a preferred approach, based on their understanding as a result of doing this research the authors could provide an assessment of weaknesses, biases, etc. of the five estimation techniques. Their order of exposition implies a preference for least squares with dummy variable, but is this really a preference or just an expositional convenience based on most common practice?

Reply: We absolutely agree with the referee. We will add to the paper the following statements.

First, the use of fixed effects is preferable. If the identification of the variable of interest is in the bilateral dimension (e.g. tariffs, non-tariff trade barriers etc.), it is suggested to use OLS with exporter and importer fixed effects. These fixed effects or dummies capture unobserved heterogeneity, including multilateral resistance terms as pointed out by e.g. Anderson and van Wincoop (2003) and Behar and Nelson (2014). If the variable of interest is in the countrylevel dimension, identification of the variable is impossible with exporter and importer fixed effects. Then we suggest using the Baier and Bergstrand (2009) Taylor approximation approach, which approximates the trade cost variables, corrected for multilateral resistance, while leaving open the country dimension for identification.

Second, in the presence of structural zeros (such as in Helpman, Melitz and Rubinstein (2008)), log-linearizing the gravity equation leads to a biased sample to perform the estimation on: all zero flows drop out as $ln(0)$ is undefined. Previous suggestions such as tobit estimation or adding a constant to all trade flows, have been shown to generate biased coefficients (e.g. Santos Silva and Tenreyro (2006)). Therefore, non-linear estimation methods such as PPML are preferred in the presence of these structural zeros. Note also that the two-stage Heckman-type selection model as proposed by Helpman et al. (2008) faces two practical drawbacks: (i) finding a convincing exclusion variable for the first stage, and (ii) the procedure only converges under severe sample size restrictions in Stata. This is probably the reason why this procedure is not widely implemented in the empirical gravity literature.

Third, pseudo-maximum likelihood methods are linked to the gravity model through the resemblance of the first-order conditions. No other distributional assumptions on the data are required. Hence, alternative methods such as zero-inflated Poisson estimation are not at all related to the gravity model, and generate biased results. See for instance http://personal.lse.ac.uk/tenreyro/LGW.html.

In the end, we suggest researchers to provide a side-by-side presentation of the fixed effects (or Baier and Bergstrand (2009)) and PPML approaches. Note also that we follow the theoretical gravity literature and estimate a cross-sectional gravity equation. If a researcher deems panel estimation to be appropriate, we suggest estimating the model using LSDV including exporter-time and importer-time fixed effects. If the bilateral variable of interest is also time-varying, one can include country-pair fixed effects. We are aware that this dimensionality causes potential convergence problems using panel PPML in Stata (e.g. the xtpqml command).

3.2 Specification of the loss function

Comments of referees: One of estimation issues is the loss function. However, it seems to me that there is more needs to be discussed on how the assumed function cause the estimation problem.

Reply: We agree that the discussion of the loss function is somewhat underrepresented in this paper. We propose an addition/clarification along the following lines:

First, the two main candidates for the loss function of the argmin objective are the quadratic loss function (for least squares regressions) and the absolute value loss function (for least absolute deviations regression). The standard choice is the least squares setting, as it generates a unique and stable solution. However, the least absolute deviations regression is more robust to outliers in the data when optimizing the objective. At the same time, the latter can generate multiple and unstable solutions. Moreover, the chosen quadratic loss function directly leads to the specified first-order conditions in the paper. From there on, we compare the other first-order conditions of PPML and GPML directly, without considering the original objective function. Hence also the name "pseudo" estimators, and these are exactly chosen for their resemblance to the original first-order conditions which derive from a quadratic loss function.

Second, as we are looking to estimate a conditional mean, the loss function in all estimation methods is quadratic. Alternative loss functions result in the estimation of other moments of the data. For instance, a linear loss function results in an estimator equal to a quantile (e.g. median) of the posterior distribution, and an "all-or-nothing" loss function results in the estimation of the mode of the posterior distribution. Moreover, the quadratic loss function is symmetric, so that positive and negative deviations from the estimated parameter get the same weight.

We will replace "3. Loss function" on p.7 with a discussion on the weighting schemes, and discuss the loss function separately.

3.3 Estimations at product-level data

Comments of referees: One might have expected to see how those methods differ when using aggregate and disaggregated product data. Indeed, at disaggregated level other methods might perform relatively better that the proposed ones. That is, using zero-inflated Poisson Quasi

Likelihood (PQL) or Gamma Two Parts model for gravity at product level might suggest how distance and border effects impact trade flows.

Reply: We thank the referee for the suggestion on estimating the model at the product level. We have considered this as well, and we believe there are a few issues related to our setup.

First and most importantly, there is no theoretical support for the existence of the gravity equation at the product level. This makes the estimation of the product-level gravity model at best a reduced form empirical exercise.

Second, if we would estimate a country-level gravity model, accounting for products or sectors, this exponentially increases the dimensionality of the fixed effects. The non-linear estimation techniques are already confronted with potential convergence problems given our current set of dummies. The proposed extension would severely increase this problem.

Third, if we would opt for a particular product gravity estimation, the results are naturally very dependent on the chosen products. At the 2-digit level, this would imply around 80 gravity models to be estimated per estimation scheme, with the estimated coefficient highly dependent on the chosen product flows.

Fourth, we have considered zero-inflated (and negative binomial) models as well, but as stated in footnote 8, these are known to be biased and/or not scale invariant. That is the main reason we did not include this type of models in the paper. Moreover, the non-linear estimators naturally deal with zeros, and recently Martin and Pham (2015) have shown that PPML generates consistent estimates in the presence of structural zeros, including at a highly disaggregated level.

In conclusion, we agree that the product-level estimation is an interesting exercise on its own, worthy to devote a separate paper to in its own right.

3.4 Linearity of relationships

Comments of referees: Eq (8) presents a linear effect of border and distance on trade flow. However, a squared distance term might be relevant. In fact, the term in level might capture a negative effect but non-linear effect of distance on trade. But with a squared term we might see a positive effect that captures the high concentration of trade over the shortest distance. Similar argument might apply for border effects.

Reply: This is a very interesting comment, which we would like to address from both a theoretical and an agnostic/empirical point of view.

From a theoretical point of view, we follow the mainstream structural gravity literature (surveyed in Head and Mayer (2014)) that stipulates a linear relationship between distance and trade. Almost all quantitative models of international trade (e.g. Eaton and Kortum (2002), Anderson and van Wincoop (2003), Helpman, Melitz and Rubinstein (2008), Arkolakis et al. (2008), Chaney (2008), Costinot and Rodriguez-Clare (2013)) predict a linear relationship between (the logs of) distance and trade. There is some discussion on empirical non-linearities in distance and trade flows and non-parametric approaches (e.g. Lai and Trefler (2002), Leamer (2007)), even within countries (e.g. Hillberry and Hummels (2008), Crozet and Koenig (2010)). However, despite the added flexibility of semi-parametric and non-parametric (reduced form) estimation methods, predictions are not more consistent. Furthermore, formal statistical tests fail to reject the theoretically consistent parametric functional form (Henderson and Millimet (2008)). Finally, the positive effect of high concentration of trade over the shortest distances is mostly captured by the Adjacency dummy, indicating whether two countries share a common border or not. The estimated coefficient on this variable is always positive.

From a completely agnostic empirical point of view, we indeed address the issue of non-linear correlations between distance (and all other variables) and trade in the second robustness test in Section 5.

3.5 Agglomeration effects

Comments of referees: Agglomeration is barely touched by the authors. Again this might have important implication on why economic activities cluster in some regions. More, emphasis on this should be put in place with border and distance effect on trade flows.

Reply: The interaction between international trade and agglomeration effects is elegantly presented in Krugman (1991). It is possible to generate an economic geography model that derives a gravity equation, taking into account spatial concentration. This paper however, compares different estimation methods of the mainstream gravity model, and does not have the ambition to extend the gravity model to include spatial econometrics. Furthermore, agglomeration effects are likely to happen across sectors within countries. We estimate the gravity model at the country level and not at the product level for reasons explained above. We absolutely agree that there is room for more appropriate methods to estimate agglomeration effects, drawing from the economic geography literature, in a separate paper.

3.6 Exclusion of GDP variables with fixed-effects specifications

Comments of referees: GDPs aren't present or reported in the table, except OLS and BB cases. However, in the text, I do not find discussion on the logic of exclusion of economic sizes from the regressions, nor I find explanation on elimination of impact of monadic terms. I normally include GDPs even when I use country level fixed effects.

Reply: The estimates represented in Table 2 - "Borders by continents" are performed on the pooled cross-section 1998-2011 (not a panel). Hence, fixed effects at the country level capture all monadic terms, including GDP. Identification of the coefficients of GDP is hence impossible.

3.7 Robustness tests

Comments of referees: A further robustness test would have been to vary the control variables. Adjacency, Common Language, Colonial Ties and RTA membership feature in all runs. These are commonly used controls, but in the gravity literature the choice of controls does vary; the selection obviously affects results, and the interesting question is the size of this effect.

Reply: We fully agree that there is a wide variety of controls used in estimating the gravity equation. However, the inclusion/exclusion of variables in particular dimensions (country, country-pair, country-time, country-pair-time) is most likely subject to omitted variable bias, hence generating biased estimated coefficients. For instance, if the coefficient of interest is at the country-time level (e.g. WTO membership of the exporter), it is best practice to account for all country, country-pair, and country-pair-time observed and unobserved variance by using fixed effects in these dimensions (including time-invariant bilateral variables such as adjacency, common language, colonial ties etc.). This leaves the dimension of interest open to be estimated by particular variables. If there are well-known covariates in this "open" dimension, which correlate with the variable of interest, the best next thing is to take these into the estimation regression as well, assuming that there are no other observables that can bias the coefficients of the variables of interest. Similar reasoning holds for the other dimensions.

We are also fully aware that the dimensionality of the estimation (e.g. with time-varying fixed effects and/or product interactions) can lead to convergence problems, especially with non-linear estimation methods such as the PPML and GPML methods.

3.8 Local polynomials

Comments of referees: The authors propose a local polynomial regression, but whether there exist potential problems with bandwidth or smoothing selection is not argued. Since, the paper compare econometrics methods to potentially solve problems that arise with these models, one might expect to see if the proposed models might have some limitation compared to previous approaches.

Reply: We are not sure whether the referee refers to the estimation of the main gravity specification, or the estimation of the residual distribution. The local polynomial regression is used to get a grip on the distribution of the residuals from the estimated model. It is only used as an alternative to a residual scatter plot, and not related to the estimation of the main model. Moreover, the polynomial regression clearly resembles a third degree polynomial. Hence, bandwidth tuning, kernel smoothing etc. would not alter the perceived fit of this polynomial. We do not relate the local polynomial regression to the main estimation method and/or comparison to other estimation techniques.

3.9 Typos

Comments of referees: Various typos in text.

Reply: All corrected, thank you for pointing these out.

References

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

[2] Arkolakis, C., Demidova, S., Klenow, P.J. and Rodríguez-Clare, A. (2008). Firm Heterogeneity and International Trade, *American Economic Review: Papers and Proceedings,* 98(2): 444-450.

[3] Baier, S. L. and Bergstrand, J. H. (2009). Bonus Vetus OLS: A Simple Method for Approximating International Trade-Cost Effects Using the Gravity Equation, *Journal of International Economics*, 77(1): 77–85.

[4] Behar, A. and Nelson, B. (2014). Trade Flows, Multilateral Resistance and Firm Heterogeneity, *The Review of Economics and Statistics*, 96(3): 538–549.

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

[6] Costinot, A. and Rodriguez-Clare, A. (2013). Trade Theory with Numbers: Quantifying the Consequences of Globalization, in Gopinath, G, E. Helpman and K. Rogoff (eds), vol. 4 of the Handbook of International Economics.

[7] Crozet and Koenig (2010)). Structural gravity equations with intensive and extensive margins, *Canadian Journal of Economics,* 43(1): 41-62.

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

[9] Head, K., Mayer, T. (2014). Gravity Equations: Workhorse,Toolkit, and Cookbook, in Gopinath, G, E. Helpman and K. Rogoff (eds), vol. 4 of the Handbook of International Economics.+

[10] Helpman, E., Melitz, M. and Rubinstein, Y. (2008). Estimating Trade Flows: Trading Partners and Trading Volumes, *Quarterly Journal of Economics,* 123: 441-487.

[11] Henderson, D.J. and Millimet, D.L. (2008)). Is Gravity Linear?, *Journal of Applied Econometrics,* 23(2): 137–172.

[12] Hillberry, R. and Hummels, D. (2008). Trade Responses to Geographic Frictions: A Decomposition Using Micro-Data, *European Economic Review*, 52(3): 527-550.

[13] Krugman, P. (1991). Increasing Returns and Economic Geography, *Journal of Political Economy*, 99(3): 483–499.

[14] Lai, H. and Trefler, D. (2002). The Gains from Trade with Monopolistic Competition: Specification, Estimation, and Mis-Specification, Working Paper No. 9169, National Bureau of Economic Research.

[15] Leamer, E.E. (2007). A Flat World, a Level Playing Field, a Small World After All, or None of the Above? A Review of Thomas L Friedman's The World is Flat, *Journal of Economic Literature*, 45(1): 83-126.

[16] Martin, W. and Pham C.S. (2015). Estimating the Gravity Model When Zero Trade Flows are Frequent and Economically Determined, World Bank Policy research working paper 7308.

[17] McCallum, J. (1995). National Borders Matter: Canada-U.S. Regional Trade Patterns, *American Economic Review,* 85(3): 615-23.

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