Referee Report to the Discussion Paper No. 2011-38 entitled "Trade Policy versus Trade Facilitation: An Application Using 'Good Old' OLS" submitted to *Economics - The Open-Access, Open-Assessment E-Journal*

This paper disentangles the effects of various elements of trade costs (transport costs, time, bureaucratic costs and tariffs) using a gravity equation specification as proposed by Baier and Bergstrand (2009) for cross-section industry-level data of the year 2000. The main finding is that time to trade, technological innovation and number of documents as proxy for trade facilitation are more important determinants in relative terms for trade flows than trade policy measures.

Overall I think disentangling the various barriers to trade is a very fruitful area of research. I think the present discussion paper is well motivated.

However, I have a couple of comments to the methodologies used and the empirical specification applied.

First, on page 9 the authors motivate that they use the first-order Taylor expansion of Baier and Bergstrand (2009) because this methodology is theoretically founded, simpler than the non-linear methodology suggested by Anderson and van Wincoop (2003) and accounts for asymmetric bilateral trade costs. I have a couple of problems with this motivation:

i) On page 9 the authors state: "The most commonly applied approach to estimate potentially unbiased gravity equation coefficients since Anderson and van Wincoop (2003) is to use region-specific fixed effects, as already suggested by the authors and by Feenstra (2004). Although this method is very simple and avoids the measurement error associated with measuring regions 'internal distances' (as in CNLS), it does not allow direct estimation of the comparative static effects of trade costs." I think this is wrong. You can do comparative statics after estimating the parameters with fixed effects. One has clearly to distinguish the estimation stage (which can include fixed effects) from the counterfactual analysis. One can use the parameter estimates from the fixed effects specification in a counterfactual analysis employing the underlying theoretical structure of the model.

ii) Again on page 9 the authors state: "Moreover, the Anderson and van Wincoop (2003) approach is only valid in a world with symmetrical bilateral trade costs ($t_{ij} = t_{ji}$), whereas the MR approximation terms also work under asymmetrical bilateral trade costs and, in reality, many trade costs are bilaterally asymmetric, such as tariff rates and transport costs." This statement highly overstates the importance of the symmetric bilateral trade costs in the original contribution of Anderson and van Wincoop (2003). Actually, the Anderson and van Wincoop (2003) methodology can easily be adapted to account for asymmetries (see for example Anderson and Yotov (2010) or Bergstrand, Egger and Larch (2011)).

iii) Another statement on page 9 reads as follows: "Baier and Bergstrand (2009) suggest applying a first-order Taylor expansion to the explanatory variables and then using OLS to estimate the gravity model specified with the transformed variables. The focus in this paper is on estimation (not in comparative statics) and therefore the simple average weights (1/N) are used in the MR construction, instead of the GDP shares used as weights in Baier and Bergstrand (2009)." Given this statement, I am a little bit puzzled why the authors use the Baier and Bergstrand (2009) first-order Taylor expansion. My understanding of the paper of Baier and Bergstrand (2009) is that they wanted to suggest a simple methodology that allows for counterfactual analysis. If you do not focus on the counterfactual analysis, the easiest thing you may want to do is to use a fixed effects specification.

iv) Related to point iii): On page 10 the authors write: "Estimating equation (2) by OLS would yield identical coefficients to other estimates used to obtain unbiased gravity equation coefficients (fixed effects and CNLS), although as with any linear approximation, an approximation error is introduced." Actually, the Baier and Bergstrand (2009) methodology is numerically identical to fixed effects estimation if you use a balanced sample, i.e. one that has the same importers and exporters and includes the x_{ii}-observations. But the CNLS is not. Coefficients will in general not be identical between CNLS and fixed effects.

Second, the specification given in equation (2) raises the following questions:

i) Why does $\ln(Y_i Y_j)$ appear on the right-hand side? Baier and Bergstrand (2009) include $\ln(Y_i)$ and $\ln(Y_j)$ as two distinct regressors. Including $\ln(Y_i Y_j)$ restricts the parameters for income of the both countries to be the same. Additionally, from an econometric point of view including GDPs as regressors maybe problematic as they are likely to be endogenous. This also follows from the methodology of Anderson and van Wincoop (2003), where (changes) in GDPs capture general equilibrium effects. This would be a further argument to use importer- and exporter-specific fixed effects, as they would control for GDPs.

ii) Equation (2) includes $\left(\ln Tariffs_{ijk} - \ln Tariffs_{P_iP_j}\right)$ in exactly the same way as bilateral distance is included. However, tariffs generate revenues and do not consume any resources. This is different to standard iceberg trade costs which are by definition resource consuming. Hence, I am not sure whether taking the model in Baier and Bergstrand (2009) and extending it for tariffs (which they did not consider) and doing the log-linear first-order Taylor approximation would lead to exactly this term for tariffs. This is something that remains to be shown.

iii) The authors given no motivation why the easy-to-trade variables (technological innovation, internal transport costs, time and the number of documents required to trade) for the exporting and importing country ET_i and ET_j, respectively, enter the specification as $\ln(ET_iET_j)$. Why not additively? Or in some other functional form. Or as separate regressors? This issue pops up again on page 11 where the authors explain that "due to the complementarity of the ET variables considered, models 1-4 include each trade facilitation variable separately, namely technological innovation, transport costs, number of days and number of documents required to trade, respectively. In order to improve the measure of ET, we also computed an average ET that is calculated as the simple average of the variables: $\sum_{m=1}^{3} \ln(x_ix_j)/3$, where x denotes time, internal transport costs and number of documents." Why not add them as separate regressors? According to the bottom of Table 2 they are correlated, but not perfectly. Another question here is why you exclude technological innovation in the constructed average? And is there not the subscript m missing in $\sum_{m=1}^{3} \ln(x_ix_j)/3$, i.e. should it not be $\sum_{m=1}^{3} \ln(x_{im}x_{jm})/3$?

iv) You follow the methodology of Baier and Bergstrand (2009) which was developed for aggregate trade flows while using sector-level data. You just add subscripts k for the sector-level, but do not discuss whether the one-sector approach suggested by Baier and Bergstrand (2009) is still applicable if extended to multiple sectors. Actually, work by Anderson and Yotov (2010) suggests that extending the gravity equation to multiple sectors leads to a different specification, where bilateral trade at the sector level no longer depends on overall GDPs and overall multilateral resistance terms but rather on sectoral spendings and sector-specific multilateral resistance terms.

Third, in the robustness analysis in section 6 the authors account for zero-trade flows using the two-stage estimation procedure suggested by Helpman, Melitz and Rubinstein (2008). Helpman, Melitz and Rubinstein (2008) use additional exclusion restrictions in their estimation as the non-linearity of the heterogeneity-bias term is insufficient to identify the model. I guess this is the motivation of the authors to write in footnote 17: "The selection equation should contain at least one variable that is not in the outcome equation, therefore colony; language and contiguity are not included in equation (4)." However, one should keep in mind what Helpman, Melitz and Rubinstein also mention in their paper on page 466: "The key is for this variable to be correlated with the z_{ij}'s but not be correlated with the residual of the second stage equation that has been estimated with the reliable excluded variables (the reliable excluded variables are believed to satisfy the exclusion restrictions on theoretical grounds). In our case this means that the residuals from the trade flow equation should be uncorrelated with this variable." However, colony, language and contiguity all turned out to be highly significant in the results given up to section 6. Hence, they are not candidates for excluded variables. I would suggest that either the authors come up with other, more plausible excluded variables or skip section 6 at all.

Fourth, section 7 entitled "Simulations" presents some numbers based on the estimated elasticities. It is not clear to me why these numbers should be helpful. The authors tried to convince us that they base their empirical specification on a sound theoretical basis. But if this is the case, then the estimated elasticities on their own are not telling. Only if you undertake a proper counterfactual analysis the magnitudes can be meaningfully interpreted. For me this is the basic contribution of the Anderson and van Wincoop (2003) paper. Hence, I would suggest dropping section 7 and really stick to the estimation stage as claimed on page 9 by the authors.

References

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

Anderson, J.E. and Y. Yotov (2010), "The changing incidence of geography," *American Economic Review* 100(5), 2157-2186.

Baier, S.L. and J.H. Bergstrand (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.

Bergstrand, J.H., P. Egger and M. Larch (2011), "Gravity Redux: Estimation of gravity-equation coefficients, elasticities of substitution, and general equilibrium comparative statics under asymmetric bilateral trade costs, unpublished working paper," available at http://www.nd.edu/~jbergstr/working_papers.html.

Helpman, E., M. Melitz and Y. Rubinstein (2008), "Estimating trade flows: Trading partners and trading volume," *Quarterly Journal of Economics* 123(2), 441-487.