# Controlling For Multilateral Resistance Terms in Linearized Trade Gravity Equations Without Spatial Econometrics


Peter H. Egger*
ETH Zurich
CEPR, CESifo, GEP, WIFO

Kevin E. Staub†
University of Melbourne
IZA


April 2, 2015


**Abstract**

This paper studies linearized spatial structural gravity models of bilateral trade à la Behrens, Ertur and Koch (2012). We show that these models do in fact not require spatial econometric methods for estimation. This result follows from the nature of the specific spatial weights matrix, and from the exporter- or importer-specific nature of some regressors and the approximation error. All structural model parameters are identified from a linear regression that uses a spatial lag of the dependent variable as a control function.



**Keywords:** Gravity models; Multilateral resistance; Spatial models.

**JEL-codes:** F14; C23.



* *Corresponding author:* ETH Zurich, Department of Management, Technology, and Economics, Weinbergstr. 35, 8092 Zurich, Switzerland; E-mail: egger@kof.ethz.ch; + 41 44 632 41 08.

†Department of Economics, 111 Barry Street, The University of Melbourne, 3010 VIC, Australia; E-mail: kevin.staub@unimelb.edu.au; +61 3 903 53776.

# 1 Introduction

The gravity equation – describing aggregate demand for goods or services between any pair of countries – is among the most successful concepts in all of economics (see Leamer and Levinsohn, 1995). It is derived from utility maximization subject to income constraints and it can be represented for exporter $i$ and importer $j$ as

$$Z_{ij} \equiv \frac{X_{ij}}{Y_i Y_j} = \frac{W_i^{\alpha-1} T_{ij}^{\alpha}}{\sum_{k=1}^{N} L_k W_k^{\alpha} T_{kj}^{\alpha}}, \quad W_i = L_i^{-1} \sum_{j=1}^{N} X_{ij} = \sum_{j=1}^{N} \frac{W_i^{\alpha} T_{ij}^{\alpha} W_j L_j}{\sum_{k=1}^{N} L_k W_k^{\alpha} T_{kj}^{\alpha}}, \tag{1}$$

where $Z_{ij}$ are aggregate bilateral exports, $X_{ij}$, which are normalized by exporter and importer GDP, $Y_i$ and $Y_j$, respectively, $W_i$ are wages or producer prices, $T_{ij}$ are bilateral trade costs, $L_i$ is population or size of the labor force (see Arkolakis, Costinot and Rodríguez-Clare, 2012). Equation (1) shows that the log of $Z_{ij}$, $z_{ij}$, is a log-nonlinear function of $\{W_i; T_{ij}; L_i\}$. The parameter $\alpha = (-\infty, 0)$ (see Dixit and Stiglitz, 1977) reflects the partial response of trade with respect to changes in trade costs. Through (1), upon choice of a numéraire wage and for a given $\alpha$, $N-1$ values of $W_i$ are determined as implicit functions of all the $N$ values of $L_i$ and all the $N^2$ values of $T_{ij}^{\alpha}$. The model has a representation which nests a variety of isomorphic structural models of aggregate bilateral demand such as endowment-economy, Ricardian, and monopolistic-competition-increasing-returns-to-scale (see Eaton and Kortum, 2002; Anderson and van Wincoop, 2003; Arkolakis, Costinot and Rodríguez-Clare, 2012; Bergstrand, Egger and Larch, 2013; Baltagi, Egger and Pfaffermayr, 2015), hence its popularity. Its (log-)nonlinear structure is the main reason why gravity equations are estimated rarely in their structural form as in (1) but practitioners use either country fixed effects, which may be inefficient, or linear approximations in estimation. Behrens, Ertur and Koch (2012, henceforth BEK) present such an approximation around the point $\alpha = 0$, which leads to a spatial model. The purpose of this paper is to analyze this approximation.

# 2 The BEK spatial structural gravity model

## 2.1 Model outline

Define world population $L \equiv \sum_{k=1}^{N} L_k$ and the $N \times N$ matrices $W_{NN}$ which in column $j$ contains the same elements $L_j/L$ in all the rows, $D_{NN} = diag_N(L_j/L)$ which contains the diagonal elements of $W_{NN}$, $\tilde{W}_{NN} \equiv W_{NN} - D_{NN}$, and $I_{NN}$ which is an identity matrix. Let us generally use the convention that lower-case letters refer to variables in logs and refer to $N$-size vectors and matrices by subscripts $N$ and $NN$, respectively. Stacking all observations across exporters $i$ for a given importer $j$ in $z_{jN} = (z_{ij})$ $l_N = (l)$, $w_N = (w_i)$ and $t_{jN} = (t_{ij})$, BEK arrive at the log-transformed and linearized counterpart to (1):

$$z_{jN} \;=\; \alpha W_{NN} z_{jN} + (\alpha - 1)(l_N + w_N) + \alpha(I_{NN} - W_{NN}) t_{jN} + u_{jN}, \tag{2}$$

where $u_{jN}$ is an approximation error due to linearization, which only varies across $j$ but not $i$.[1] It is customary in empirical work to specify $t_{ij} = \sum_{h=1}^{H} \gamma_h d_{h,ij}$, where $d_{h,ij}$ are observable variables such as bilateral log-distance. What is then estimated on $d_{h,ij}$ are the compound parameters $\alpha\gamma_h$. The reduced form to (2) is

$$z_{jN} = (I_{NN} - \alpha W_{NN})^{-1}[(\alpha - 1)(l_N + w_N) + \alpha(I_{NN} - W_{NN})t_{jN} + u_{jN}]. \quad (3)$$

Existence of the latter requires $(I_{NN} - \alpha W_{NN})$ to have finite elements and to be invertible independent of the number of countries $N$. BEK reformulate the model resulting in the structural and reduced forms

$$\begin{aligned} z_{jN} &= (I_{NN} - \alpha D_{NN})^{-1} \\ &\quad \times[\alpha \tilde{W}_{NN} z_{jN} + (\alpha - 1)(l_N + w_N) + \alpha(I_{NN} - W_{NN})t_{jN} + u_{jN}], \end{aligned} \quad (4)$$

$$\begin{aligned} z_{jN} &= [I_{NN} - \alpha(I_{NN} - \alpha D_{NN})^{-1}\tilde{W}_{NN}]^{-1}(I_{NN} - \alpha D_{NN})^{-1} \\ &\quad \times[(\alpha - 1)(l_N + w_N) + \alpha(I_{NN} - W_{NN})t_{jN} + u_{jN}], \end{aligned} \quad (5)$$

respectively. The presence of $(I_{NN} - \alpha D_{NN})^{-1}$ in (4) makes the model nonlinear in $\alpha$. BEK tackle this by estimating a model with right-hand side variables $\{\tilde{W}_{NN} z_{jN}; (l_N + w_N); (I_{NN} - W_{NN})t_{jN}\}$ for each $i$ separately, since the coefficients on these variables are proportional to $(1 - \alpha L_i/L)^{-1}$.

## 2.2 Properties and novel insights

Clearly,

$$[I_{NN} - \alpha(I_{NN} - \alpha D_{NN})^{-1}\tilde{W}_{NN}]^{-1}(I_{NN} - \alpha D_{NN})^{-1} = (I_{NN} - \alpha W_{NN})^{-1},$$

and the reduced forms in (5) and (3) are the same. The matrix $I_{NN} - \alpha W_{NN}$ is invertible for any finite $\alpha \neq 1$ as then it has full rank. In any case, $\alpha = 1$ is outside the theoretically admissible parameter space. This suggests that the spatial model does not require the reformulation in (4) advocated by BEK. Since $W_{NN}$ is idempotent so that $W_{NN}^2 = W_{NN}$, the inverse $(I_{NN} - \alpha W_{NN})^{-1} = I_{NN} + \frac{\alpha}{1-\alpha} W_{NN}$ and $W_{NN}(I_{NN} - \alpha W_{NN})^{-1} = \frac{1}{1-\alpha} W_{NN}$. Moreover, since neither $u_{jN}$ nor $l_N$ vary across exporters, $W_{NN} u_{jN} = u_{jN}$ and $W_{NN} l_N = l_N$. Therefore,

$$W_{NN} z_{jN} = -W_{NN}(l_N + w_N) + \frac{1}{1 - \alpha} u_{jN}, \quad (6)$$

where $-W_{NN}(l_N + w_N)$ is a constant. Hence, there are no suitable instruments for $W_{NN} z_{jN}$ in this model as required, for instance, for a two-stage least-squares approach in

[1] After defining $Y \equiv \sum_{i=1}^{N} Y_i$, at the approximation point of the model $\alpha = 0$, we obtain $X_{ij} = \frac{L_i Y_j}{L}$ and $Y_i = \sum_{j=1}^{N} X_{ij} = \frac{L_i Y}{L}$, which implies wage equalization, $W_i = W$. Choosing the wage as the numéraire, we obtain $X_{ij} = \frac{L_i L_j}{L}$. Then, trade costs are irrelevant, and the variance of log bilateral exports, $x_{ij} = \ln X_{ij}$, is fully determined by the variation in exporter- and importer-specific log labor endowments (or population) across countries.

the spirit of Kelejian and Prucha (1998), Lee (2003), or Kelejian, Prucha and Yuzefovich (2004). In other words, the nature of $\{W_{NN}, l_N, w_N, u_{jN}\}$ and the parameter restrictions in the model imply that all of the variation in $W_{NN}z_{jN}$ is due to the approximation error, $u_{jN}$. Replacing $W_{NN}z_{jN}$ in (2) by the right-hand side in (6) and adding $l_N + w_N$ on both sides of the equation results in

$$\tilde{z}_{jN} \equiv z_{jN} + l_N + w_N \quad = \quad \alpha(I_{NN} - W_{NN})(w_N + t_{jN}) + \frac{1}{1-\alpha}u_{jN}. \tag{7}$$

We obtain the following five insights. First, BEK's linearization of (1) can be represented by (7), which relies exclusively on spatially weighted exogenous variables, but not on spatial lags of the dependent variable or of the disturbances. This corresponds to a spatial Durbin model. Second, while omitting a relevant spatial lag of the dependent variable from the right-hand side of a spatial autoregressive model usually leads to an omitted variables bias, this is not the case here, in a narrow sense. Omitting $\alpha W_{NN}z_{jN}$ from the right-hand side of (2) has only two consequences for equation (7): a rescaling of the constant and of the error term in (7) relative to (2). Third, since the approximation error $u_{jN}$ varies only across importers $j$ and exclusively depends on exogenous model variables, it appears natural to specify it as to be heteroskedastic or clustered by exporting country $i$. Hence, the device for parameter estimation is an OLS model with cluster-robust standard errors, which is much simpler than the first-order spatial-autoregressive-moving-average (SARMA) model in BEK.[2] Fourth, under BEK's assumptions, the model should be estimated for all countries jointly for the sake of efficiency gains, which is cumbersome with BEK's approach relying on (4). Fifth, to the extent that the approximation error is correlated with $w_N$ or $(I_{NN} - W_{NN})t_{jN}$, it may be preferable to estimate (2) instead of (7). The reason is that $W_{NN}z_{jN}$ depends linearly on $u_{jN}$ according to (6) and, hence, fully controls for BEK's model approximation error. The latter implies that $W_{NN}z_{jN}$ is a control function. Its parameter absorbs potential bias from the correlation of the other regressors with $u_{jN}$. As a result, the parameter on $W_{NN}z_{jN}$ should not be interpreted as an estimator for $\alpha$.

# 3 Simulation study

## 3.1 Design of experiments

We construct *worlds* of countries and country pairs according to (1) where everything is known to the simulator, while the researcher does not know the parameters on the regressors. We consider two configurations regarding country numbers with $N \in \{30; 60\}$ leading to numbers of country pairs of $N^2 \in \{900; 1,600\}$. This corresponds to typical data situations found in empirical structural work on gravity models (see Eaton and Kortum, 2002; Anderson and van Wincoop, 2003; Balistreri and Hillberry, 2007; Behrens, Ertur and Koch, 2012). For each of these worlds, we consider three configurations $\alpha \in \{-2; -4; -9\}$, which are supported quantitatively by a sizable body of

[2]Alternatively, one might want to estimate an exponential-family model for reasons outlined in Santos Silva and Tenreyro (2006).

work (see Arkolakis, Costinot and Rodríguez-Clare, 2012). Hence, there are six parameter configurations. For each of them, we randomly draw $1,000$ independent vectors of bilateral distances with typical element $DIST_{ij}$ and population sizes with typical element $L_i$ from the empirical distribution of these variables as published by the Centre d'Études Prospectives et d'Informations Internationales for $DIST_{ij}$ and by the World Bank's World Development Indicators for $L_i$ (using the year 2007). In line with the robust result of a coefficient on log distance of about $\alpha\gamma_{dist} = -1$ in empirical gravity models, we assume that log distance, $dist_{ij}$, is related to log trade costs $t_{ij}$ by a parameter of $\gamma_{dist} = -1/\alpha$. Based on the draws for $L_i$ and $t_{ij}$, the endogenous variables $W_i$ and $X_{ij}$ are solved from (1).

### 3.2 Features of model variables and the approximation error

Before turning to estimation, it is useful to study some moments and the correlations of key variables in the model across all experiments. For this purpose, we report on the averages of an analysis of variance of some variables in Table 1 and on average partial correlation coefficients in Table 2, each of them computed across all draws within one of the six parameter configurations in $\{N; \alpha\}$.

Table 1 reports on sums of squares of key variables and reveals the following features. First, the variation in the approximation error, $u_{ij}$, is large relative to normalized bilateral exports in logs, $z_{ij}$, and its size rises with the absolute level of $\alpha$; i.e., with the distance to the approximation point used by BEK to linearize the model.

The approximation error varies to a greater degree than log wages, $w_N$, whose variance is the same as that of $(I_{NN} - W_{NN})w_N$. The relative magnitude of the sum of squares of $u_{ij}$ relative to that of $z_{ij}$ declines as $N$, the number of countries, rises. The variance of $(I_{NN} - W_{NN})t_{jN}$ with typical element $\tilde{t}_{ij}$ is important relative to that of $u_{jN}$ and $w_N$. Clearly, while the exporter-and importer-specific components in $t_{ij}$ are symmetric by design (log-distance is symmetric), those of $\tilde{t}_{ij}$ are not. The pair-specific component of $\tilde{t}_{ij}$ is much bigger than the country-specific ones. Second, the variation in $u_{ij}$ is purely importer-specific. This is because BEK's approximation is about an importer-specific term, the log consumer-price index.

Table 2 suggests that there is a perfect correlation between the elements of $W_{NN}z_{jN}$ and the ones of the approximation error, $u_{jN}$, consistent with (6). There is some correlation between $(I_{NN} - W_{NN})t_{jN}$ and $u_{jN}$. This means that the parameter on $(I_{NN} - W_{NN})t_{jN}$ may exhibit some bias unless we condition on $W_{NN}z_{jN}$ (which means conditioning on $u_{jN}$, as mentioned before). However, we expect this bias to fade as $N$ rises. Not surprisingly, this problem becomes more pertinent if the approximation error is larger, which is the case with a bigger absolute value of $\alpha$.

Figure 1 visualizes the relationships in Table 2 based on one specific random draw for $N = 30$ and $\alpha = -4$. There are four general insights from an inspection of Figure 1 in conjunction with Table 2. First, the upper left panel of the figure documents that $W_{NN}z_{jN}$ is indeed perfectly correlated with $u_{jN}$ as suggested by (6). Second, all of the panels in Figure 1 illustrate the block structure of $u_{jN}$ which means it is not independently and identically distributed. Third, while the correlation between $u_{jN}$ and the other right-hand side model variables is weak on average, it may be stronger depending on the specific configuration of trade costs ($t_{jN}$) and population size ($W_{NN}$). From Table 2 we know that the risk of correlation between model variables and $u_{jN}$ is higher for $(I_{NN} - W_{NN})t_{jN}$ than for $w_N$. Figure 1, for instance, illustrates a case where $(I_{NN} - W_{NN})t_{jN}$ is negatively and $(\alpha - 1)(l_N + w_N) + \alpha(I_{NN} - W_{NN})t_{jN}$ is positively correlated with $u_{jN}$. In such a case we would expect the estimated parameter on $(I_{NN} - W_{NN})t_{jN}$ to be biased. Altogether we would expect a larger root mean squared error for the parameter on this variable than on $w_N$ or $(I_{NN} - W_{NN})w_N$, unless one controls for $u_{jN}$.

### 3.3 Parameter estimation

In this subsection, we generally employ $\tilde{z}_{jN}$ as the normalized dependent variable. We define the log-nonlinear term $M_{jN} \equiv \left(\ln(\sum_{k=1}^{N} L_k W_k^{\alpha} T_{kj}^{\alpha})\right)$ and formulate four types of models:

$$\tilde{z}_{jN} = \alpha_0 + \alpha_w(l_N + w_N) + \alpha_t t_{jN} - M_{jN}, \quad \text{(A)}$$

$$\tilde{z}_{jN} = \alpha_0 + \alpha_z W_{NN} z_{jN} + \alpha_w(l_N + w_N) + \alpha_t(I_{NN} - W_{NN})t_{jN} + u_{jN}, \quad \text{(B)}$$

$$\tilde{z}_{jN} = \alpha_0 + \alpha_w(I_{NN} - W_{NN})w_N + \alpha_t(I_{NN} - W_{NN})t_{jN} + \frac{1}{1-\alpha}u_{jN}, \quad \text{(C)}$$

$$\tilde{z}_{jN} = \alpha_0 + \alpha_w(l_N + w_N) + \alpha_t(I_{NN} - W_{NN})t_{jN} + \frac{1}{1-\alpha}u_{jN}. \quad \text{(D)}$$

Model (A) is the structural model directly corresponding to the log of (1), which we estimate by iterative least squares (cf. Anderson and van Wincoop, 2003). Models (B)-(D) can be estimated by simple OLS. In Section 2, we proved that Models (C) and (D) are equivalent so that there is no need to report on results for (D) apart from (C). Of all Models (A)-(C) we only present an unconstrained parameter-estimation version each, which does not enforce that $\alpha_w = \alpha$ and $\alpha_t = \alpha\gamma_{dist}$ are identical due to the chosen parametrization. We do so to mimic the situation of an empirical researcher who does not observe $t_{ij}$ but only $dist_{ij}$. The estimated parameters $\{\hat{\alpha}_w; \hat{\alpha}_t\}$ should be close to the true $\alpha$, especially, when being based on Models (A) or (B).

Apart from a process where the structural nonlinear Model (A) is true, we consider one where

$$\tilde{z}^*_{jN} = \tilde{z}_{jN} + \varepsilon_{jN} = \alpha_0 + \alpha_w(l_N + w_N) + \alpha_t t_{jN} - M_{jN} + \varepsilon_{jN},$$

with $\varepsilon_{ij} \sim i.i.d.N(0, \sigma^2_\varepsilon)$. We calibrate $\sigma^2_\varepsilon$ such that, in each experiment, the explanatory power as measured by the $R^2$ is 80% ($= (1 - \sigma^2_\varepsilon/\sigma^2_{\tilde{z}^*}) \times 100\%$), which is representative of a vast amount of empirical work on gravity models. The term $\varepsilon_{ij}$ adds stcchastics in a narrow sense which provides for a residual with Models (A) and (B) and one beyond the approximation (or linearization) error in Model (C).

We report on the average bias and root-mean-squared error (RMSE) in percent of the true $\alpha$ across all draws per configuration of $\{N; \alpha\}$ in Tables 3 and 4. Both tables are organized in three by two blocks. Each horizontal block contains estimates for the Models (A)-(C) for the cases $N = \{30; 60\}$. Vertically, we have three blocks corresponding to $\alpha = \{-2; -4; -9\}$. For each of the six blocks, we report on $\{\hat{\alpha}_{Mz}; \hat{\alpha}_w; \hat{\alpha}_t\}$ (where applicable).

We may summarize the simulation results in Tables 3 and 4 as follows. First, in the absence of $\varepsilon_{ij}$ (Table 3), both Models (A) and (B) correspond to the true one so that both the bias and the RMSE for $\{\hat{\alpha}_w; \hat{\alpha}_w\}$ in percent are zero. Recall that conditioning on $W_{NN}z_{jN}$ means conditioning on $u_{jN}$, according to (6). Clearly, Model (C) performs worst, but the bias and RMSE are still relatively small in Table 3. Second, the bias and RMSE of $\{\hat{\alpha}_w; \hat{\alpha}_w\}$ in percent rise in Model (C) with the absolute value of $\alpha$. This is a consequence of the linearization point of BEK's model being $\alpha = 0$. A greater distance to the approximation point means introducing endogeneity of $w_N$ and $(I_{NN} - W_{NN})t_{jN}$ in Model (C). Third, as expected, both the bias and the RMSE of $\{\hat{\alpha}_w; \hat{\alpha}_w\}$ in percent decline in Model (C) as the number of countries rises. However, then also the need for controlling for general equilibrium effects and nonlinear trade-cost effects as captured by $H_{jN}$ in (A) declines (see Egger and Staub, 2015). Fourth, adding a stochastic term $\varepsilon_{ij}$ to the log-transformed true model does not add a significant bias for Models (A)-(C), but it raises the corresponding RMSEs relative to Table 3, as expected.

## 4 Conclusions

This paper sheds light on the nature of structural linearized gravity models involving an endogenous spatial lag – other countries' population-share-weighted bilateral trade flows – as developed in Behrens, Ertur and Koch (2012). We demonstrate that the structure of the models, when considering their properties, is such that they do not require any use of spatial econometrics. Exporter-population-share-weighted log bilateral exports on the right-hand side of the models serve as a control function for the approximation error of the linearization, and this variable can be included without specific treatment (i.e., ignoring its endogeneity). One model version corresponds to a spatial Durbin model, which involves exporter-population-share-weighted exporter log wages (a constant) and log bilateral trade costs. These results should please the applied researcher, since estimation of such linearized models only involves OLS (on log-tranformed trade flows) with clustered standard errors at the level of exporters. For these particular models, there is no need for resorting to spatial econometric methods from a structural perspective, neither for point estimation nor for inference.

## References

Anderson JE, van Wincoop E. 2003. Gravity with gravitas: a solution to the border puzzle. *American Economic Review* **93**: 170–192.

Arkolakis C, Costinot A, Rodríguez-Clare A. 2012. New trade models, same old gains? *American Economic Review* **102**: 94–130.

Balistreri EJ, Hillberry RH. 2007. Structural estimation and the border puzzle. *Journal of International Economics* **72**: 451–463.

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.

Behrens K, Ertur C, Koch W. 2012. 'Dual' gravity: Using spatial econometrics to control for multilateral resistance. *Journal of Applied Econometrics* **27**: 773–794.

Bergstrand JH, Egger PH, 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**: 110–121.

Dixit AK, Stiglitz JE. 1977. Monopolistic competition and optimum product diversity. *American Economic Review* : 297–308.

Eaton J, Kortum SS. 2002. Technology, geography, and trade. *Econometrica* **70**: 1741–1779.

Egger PH, Staub KE. 2015. GLM estimation of trade gravity models with fixed effects. *Empirical Economics* : forthcoming.

Kelejian HH, Prucha IR. 1998. A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *Journal of Real Estate Finance and Economics* **17**: 99–121.

Kelejian HH, Prucha IR, Yuzefovich Y. 2004. Instrumental variable estimation of a spatial autoregressive model with autoregressive disturbances: large and small sample result. In LeSage J, Pace RK (eds.) *Advances in Econometrics*. New York: Elsevier, 163–198.

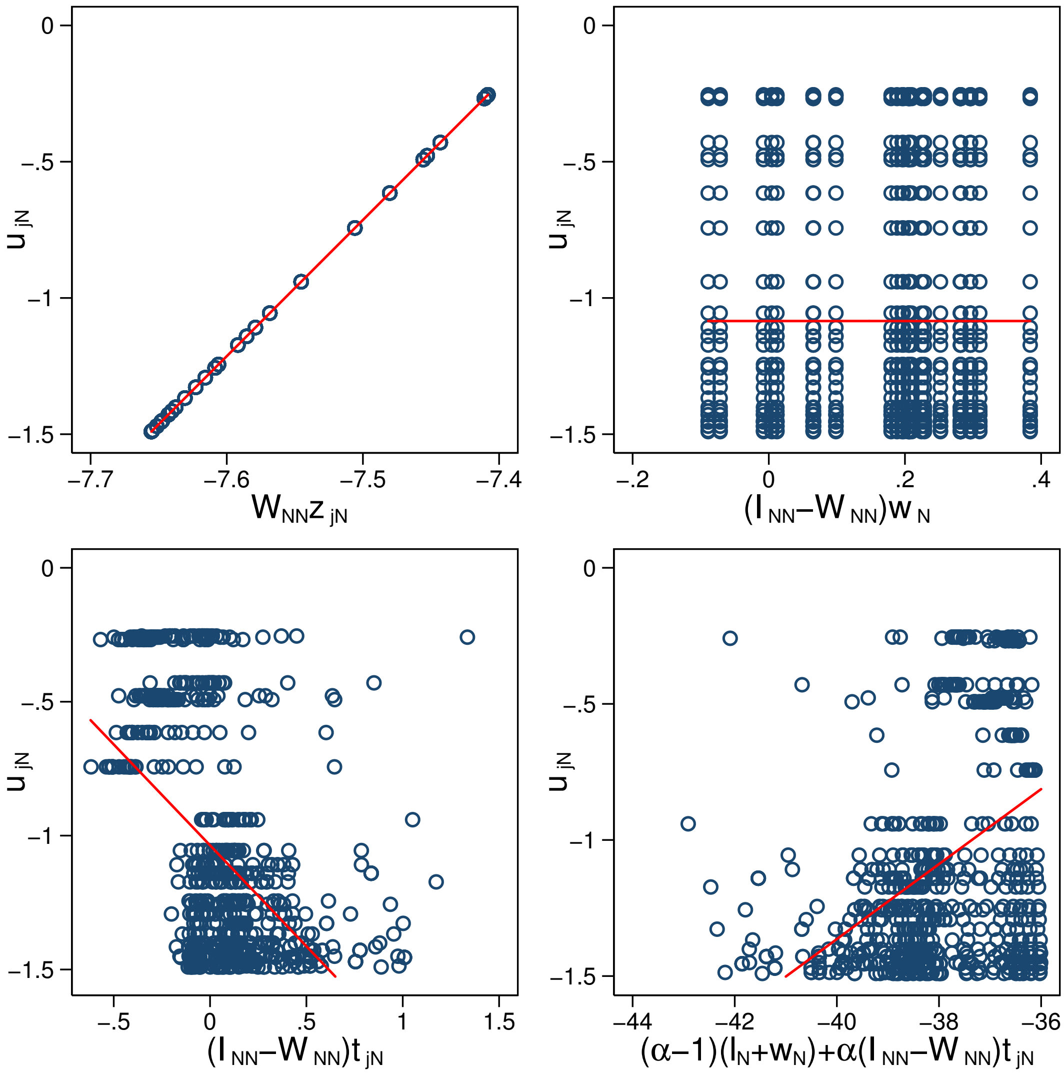
Leamer EE, Levinsohn J. 1995. International trade theory: the evidence. In Grossman GM, Rogoff K (eds.) *Handbook of International Economics Vol. 3*. North Holland, 1339–1394.

Lee LF. 2003. Best spatial two-stage least squares estimators for a spatial autoregressive model with autoregressive disturbances. *Econometric Reviews* **22**: 307–335.

Santos Silva JMC, Tenreyro S. 2006. The log of gravity. *The Review of Economics and Statistics* **88**: 641–658.

## Tables and figures

**Figure 1:** Scatterplot and linear fit of approximation error (u) and right-hand side variables from a random draw of the DGP with $\alpha = -4$ for $N = 30$




***Notes*:** The four panels of the figure display scatterplots of data obtained from one random draw of the DGP with $\alpha = -4$ for 30 countries (900 observations). The red line represents the fit from a linear regression.

**Table 1:** Analysis of variance for key model variables (mean sums of squares over 1,000 replications)

| | $N = 30$ | | | | $N = 60$ | | | |
|---|---|---|---|---|---|---|---|---|
| SS | $z_{ij}$ | $u_{ij}$ | $w_i$ | $\tilde{t}_{ij}$ | $z_{ij}$ | $u_{ij}$ | $w_i$ | $\tilde{t}_{ij}$ |
| | | | | $\alpha = -2$ | | | | |
| $i$ (exporter) | 112.26 | 0.00 | 14.99 | 20.47 | 370.46 | 0.00 | 52.21 | 77.74 |
| $j$ (importer) | 112.26 | 68.40 | 0.00 | 29.39 | 370.46 | 171.00 | 0.00 | 95.89 |
| residual | 738.85 | 0.00 | 0.00 | 184.71 | 2398.63 | 0.00 | 0.00 | 599.66 |
| total | 963.38 | 68.40 | 14.99 | 234.58 | 3139.55 | 171.00 | 52.21 | 773.29 |
| | | | | $\alpha = -4$ | | | | |
| $i$ (exporter) | 127.93 | 0.00 | 6.01 | 5.11 | 127.93 | 0.00 | 6.01 | 5.11 |
| $j$ (importer) | 127.93 | 205.19 | 0.00 | 7.47 | 127.93 | 205.19 | 0.00 | 7.47 |
| residual | 737.95 | 0.00 | 0.00 | 46.12 | 737.95 | 0.00 | 0.00 | 46.12 |
| total | 993.82 | 205.19 | 6.01 | 58.70 | 993.82 | 205.19 | 6.01 | 58.70 |
| | | | | $\alpha = -9$ | | | | |
| $i$ (exporter) | 139.89 | 0.00 | 1.63 | 1.01 | 139.89 | 0.00 | 1.63 | 1.01 |
| $j$ (importer) | 139.89 | 949.64 | 0.00 | 1.48 | 139.89 | 949.64 | 0.00 | 1.48 |
| residual | 737.95 | 0.00 | 0.00 | 9.11 | 737.95 | 0.00 | 0.00 | 9.11 |
| total | 1017.72 | 949.64 | 1.63 | 11.60 | 1017.72 | 949.64 | 1.63 | 11.60 |

***Notes***: SS refers to sum of squares. $\tilde{t}_{ij} \equiv t_{ij} - \sum_i \frac{L_i}{L} t_{ij}$ is a typical element of $(I_{NN} - W_{NN})t_{jN}$.

**Table 2:** Partial correlation coefficients of model variables with approximation error $u_{jN}$ (mean and standard deviations over 1,000 replications)

| | $N = 30$ | | $N = 60$ | |
|---|---|---|---|---|
| | Mean | SD | Mean | SD |
| | | $\alpha = -2$ | | |
| $W_{NN} z_{jN}$ | 1.00 | 0.00 | 1.00 | 0.00 |
| $l_N + w_N$ | -0.00 | 0.00 | 0.00 | 0.00 |
| $(I_{NN} - W_{NN})t_{jN}$ | 0.06 | 0.14 | 0.06 | 0.10 |
| | | $\alpha = -4$ | | |
| $W_{NN} z_{jN}$ | 1.00 | 0.00 | 1.00 | 0.00 |
| $l_N + w_N$ | -0.00 | 0.00 | -0.00 | 0.00 |
| $(I_{NN} - W_{NN})t_{jN}$ | -0.01 | 0.19 | -0.02 | 0.14 |
| | | $\alpha = -9$ | | |
| $W_{NN} z_{jN}$ | 1.00 | 0.00 | 1.00 | 0.00 |
| $l_N + w_N$ | 0.00 | 0.00 | -0.00 | 0.00 |
| $(I_{NN} - W_{NN})t_{jN}$ | -0.06 | 0.21 | -0.08 | 0.17 |

**Table 3:** Average bias and root mean squared error of estimated model parameters (1,000 replications): DGP with approximation error only

| | | $N = 30$ | | | $N = 60$ | | |
|---|---|---|---|---|---|---|---|
| | | (A) | (B) | (C) | (A) | (B) | (C) |
| | | | $\alpha = -2$ | | | | |
| $\hat{\alpha}_z$ | Bias | – | -150.00 | – | – | -150.00 | – |
| | RMSE | – | 150.00 | – | – | 150.00 | – |
| $\hat{\alpha}_w$ | Bias | -0.00 | 0.00 | -1.21 | -0.00 | -0.00 | -0.76 |
| | RMSE | 0.00 | 0.00 | 2.05 | 0.00 | 0.00 | 1.28 |
| $\hat{\alpha}_t$ | Bias | -0.00 | -0.00 | -0.79 | -0.00 | -0.00 | -0.60 |
| | RMSE | 0.00 | 0.00 | 1.56 | 0.00 | 0.00 | 1.05 |
| | | | $\alpha = -4$ | | | | |
| $\hat{\alpha}_z$ | Bias | – | -125.00 | – | – | -125.00 | – |
| | RMSE | – | 125.00 | – | – | 125.00 | – |
| $\hat{\alpha}_w$ | Bias | -0.00 | 0.00 | -0.81 | -0.00 | -0.00 | -0.36 |
| | RMSE | 0.00 | 0.00 | 1.67 | 0.00 | 0.00 | 0.99 |
| $\hat{\alpha}_t$ | Bias | -0.00 | -0.00 | -0.13 | -0.00 | -0.00 | 0.08 |
| | RMSE | 0.00 | 0.00 | 1.80 | 0.00 | 0.00 | 1.24 |
| | | | $\alpha = -9$ | | | | |
| $\hat{\alpha}_z$ | Bias | – | -111.11 | – | – | -111.11 | – |
| | RMSE | – | 111.11 | – | – | 111.11 | – |
| $\hat{\alpha}_w$ | Bias | -0.00 | -0.00 | -0.60 | -0.00 | 0.00 | -0.16 |
| | RMSE | 0.00 | 0.00 | 1.46 | 0.00 | 0.00 | 0.91 |
| $\hat{\alpha}_t$ | Bias | -0.00 | 0.00 | 0.40 | -0.00 | 0.00 | 0.64 |
| | RMSE | 0.00 | 0.00 | 2.25 | 0.00 | 0.00 | 1.74 |

***Notes***: Columns (A), (B) and (C) refer to Models (A), (B) and (C) in Section 3.3. Table entries are average biases and root mean squared errors in percent of the true $\alpha$.

**Table 4:** Average bias and root mean squared error of estimated model parameters (1,000 replications): DGP with approximation error and stochastic error

| | | $N = 30$ | | | $N = 60$ | | |
|---|---|---|---|---|---|---|---|
| | | (A) | (B) | (C) | (A) | (B) | (C) |
| | | | | $\alpha = -2$ | | | |
| $\hat{\alpha}_z$ | Bias | – | -128.50 | – | – | -128.95 | – |
| | RMSE | – | 129.33 | – | – | 129.51 | – |
| $\hat{\alpha}_w$ | Bias | 0.20 | -0.00 | -1.00 | -0.03 | -0.18 | -0.80 |
| | RMSE | 4.13 | 4.11 | 4.49 | 1.70 | 1.70 | 2.10 |
| $\hat{\alpha}_t$ | Bias | 0.01 | -0.07 | -0.79 | 0.01 | -0.12 | -0.59 |
| | RMSE | 1.00 | 1.11 | 1.76 | 0.43 | 0.59 | 1.13 |
| | | | | $\alpha = -4$ | | | |
| $\hat{\alpha}_z$ | Bias | – | -114.20 | – | – | -114.57 | – |
| | RMSE | – | 114.42 | – | – | 114.72 | – |
| $\hat{\alpha}_w$ | Bias | 0.20 | 0.12 | -0.61 | -0.05 | -0.04 | -0.41 |
| | RMSE | 3.32 | 3.30 | 3.62 | 1.38 | 1.38 | 1.68 |
| $\hat{\alpha}_t$ | Bias | 0.00 | 0.33 | -0.14 | 0.00 | 0.24 | 0.09 |
| | RMSE | 1.04 | 1.37 | 1.99 | 0.45 | 0.77 | 1.31 |
| | | | | $\alpha = -9$ | | | |
| $\hat{\alpha}_z$ | Bias | – | -106.46 | – | – | -106.70 | – |
| | RMSE | – | 106.51 | – | – | 106.73 | – |
| $\hat{\alpha}_w$ | Bias | 0.15 | 0.12 | -0.45 | -0.05 | 0.03 | -0.20 |
| | RMSE | 2.94 | 2.94 | 3.22 | 1.22 | 1.22 | 1.51 |
| $\hat{\alpha}_t$ | Bias | 0.00 | 0.63 | 0.39 | 0.00 | 0.50 | 0.65 |
| | RMSE | 1.06 | 1.67 | 2.41 | 0.46 | 1.01 | 1.80 |

***Notes***: Columns (A), (B) and (C) refer to Models (A), (B) and (C) in Section 3.3. Table entries are average biases and root mean squared errors, in percent of the true $\alpha$.