On the Problem of Endogenous Unobserved Effects in the Estimation of Gravity Models

Peter Egger

University of Innsbruck

Abstract

We propose to estimate gravity models by a Hausman and Taylor (1981) 2SLS error components approach. First, this allows to account for the possible correlation of some of the regressors with the unobserved effects, which has been ignored in previous research. Second and in contrast to the Within estimator, one obtains estimates for the time-invariant variables also in a framework with bilateral effects. We show that distance and country size are important sources of correlation. The properly estimated parameter of distance is about twice as high as its biased GLS counterpart and markedly higher than found by others.

- JEL Classifications : C33, F14
- Key words : Gravity equation, Panel econometrics

I. Introduction

Since a couple of years, it is convenient to analyze of bilateral trade flows (and more recently also of multinational sales, FDI, and migration) using panel econometric methods, since simple OLS estimates provide only limited information due to inconsistency and often deviate to a large extent from their panel econometric counterparts.

Egger (2000) mentions that the proper econometric representation of a gravity equation includes fixed exporter, importer, and time effects. However, Egger and Pfaffermayr (2002) show that a set-up with time and bilateral effects is even more

^{*}Corresponding address: Professor Peter Egger, Department of Economics, University of Innsbruck, Universitaetsstrasse 15, A-6020 Innsbruck, Tel: +43-512-507-7404 E-mail: Peter.Egger@uibk.ac.at ©2004-Center for International Economics, Sejong Institution, All Rights Reserved.

On the Problem of Endogenous Unobserved Effects in the Estimation of Gravity Models 183

general from an analysis of variance point of view.¹ Such an approach also gives more support to the random effects model than a framework with both exporter and importer effects. However, in this more general framework, the coefficients of time-invariant variables (most prominently distance, but also common borders, language, etc.) cannot be estimated in a standard fixed effects regression.

The present paper argues that simple GLS estimation of a random effects model is also of only limited value. The GLS model gives biased estimates, if there is some correlation of the explanatory variables with the unobserved bilateral effects, which should be interpreted as unobserved bilateral propensity to trade. In a sensitivity analysis, we identify the sources of correlation and propose to estimate the two-way gravity model with a Hausman and Taylor (1981) two-stage least squares (2SLS) error components model. For our data, the coefficient for distance (i.e., transport costs) is about twice as high in absolute value for the consistent Hausman and Taylor approach as for its biased simple GLS counterpart.

II. The Empirical Model: A Hausman and Taylor (1981) Approach

We follow a specification, which is motivated from the New Trade Theory (see Helpman, 1987; Hummels and Levinsohn, 1995; Egger, 2000; and others). In terms of general model structures, the basic specification reads:

$$Y_{ijt} = X_{ijt}\beta + Z_{ij}\gamma + \varepsilon_{ijt}$$

$$\varepsilon_{ijt} = \mu_{ij} + \lambda_t + v_{ijt},$$
(1)

where Y_{ijt} in our case represents real bilateral exports (*LREX*_{ijt}, the first letter "*L*" indicates that variables are expressed in logs). X_{ijt} and Z_{ij} denote $NT \times k$ and $NT \times g$ matrices, where the former vary not only in the cross-section (either over exporters, *i*; importers, *j*; or bilateral relationships, *ij*) but also in the time dimension (*t*), and the latter are time-invariant with distance as the most prominent example. *N* denotes the number of cross-sections and *T* is the number of years, where the typical bilateral export relation is observed. The specification only includes distance between two countries' capitals in Z_{ij} (*LDIST*_{ijt}⁽⁻⁾, hence, g = 1), reflecting a proxy

¹Of course, one would exhaust full variance information if modelling all types of interaction effects: exporter-time, importer-time, and exporter-importer (i.e., bilateral effects). However, we will focus only on the latter, as one typically is interested in parameter estimates of traditional time-variant gravity variables, which account for the major part of the variance in the other dimensions (see Egger and Pfaffermayr, 2002).

for transport costs.²

 X_{ijt} comprises the following variables. The overall bilateral country size in terms of real GDP, of a country pair $LGDT_{ijt}^{(+)} = \log(GDP_{it}+GDP_{jt})$, similarity in country size in terms of real GDP, $LSIM_{ijt}^{(+)} = \log[1 - (GDP_{it}/(GDP_{it}+GDP_{jt}))^2 - (GDP_{jt}/(GDP_{it}+GDP_{jt})^2]$, with $-\infty \le LSIM \le \log(0.5)$, which reaches its maximum, if two countries have the same size. Finally, a distance measure for relative factor endowments in terms of GDP per capita (GDP/POP), namely $LRLF_{ijt}^{(+)} = |\log(GDP_{it}/POP_{it}) - \log(GDP_{jt}/POP_{jt})|$, with $0 \le LRLF$, which reaches its minimum, if two countries exhibit identical relative factor endowments.

As in traditional gravity applications, other determinants are included, which besides time - vary either in the exporter or the importer dimension and reflect export enhancing or impeding factors. These variables are measures of the observable propensity to trade. In theoretical models, the latter is represented by trade costs in a broad sense. Since they do not vary in the bilateral but only the country-specific dimension of the cross-section, they are assumed to be uncorrelated with the unobserved bilateral effects. The corresponding set of variables contains exporter and importer viability of contracts ($LFV_{it}^{(+)}$, $LFV_{jt}^{(+)}$), and exporter and importer rule of law ($LFR_{it}^{(+)}$, $LFR_{jt}^{(+)}$), which are expected to have a positive impact on bilateral trade (exports). Finally, exporter and importer taxes on traded goods as a percentage of a country's overall trade ($LFT_{it}^{(-)}$, $LFT_{jt}^{(+)}$) are likely to impede exports the higher they are in the sending country or the lower they are in the receiving country. μ_{ij} are unobserved (random) bilateral effects and λ_t are time effects, which we propose to treat as fixed (see also Egger, 2000, for a detailed reasoning).

Of course and despite the correlation between some explanatory variables with the unobserved effects, the fixed effects approach obtains a consistent estimate of the β parameters, but it will not give an estimate of γ . In contrast, the random effects approach is more efficient and it obtains an estimate of γ , but it assumes that all elements of both Z and X are uncorrelated with μ_{ij} . Regarding the interpretation of μ_{ij} as the time-invariant bilateral propensity to trade, we propose to account for the possible lack of this required independence. First, distance (*LDIST*) may be one source of correlation. Second, part of the X (called X_2) may represent one. Possible candidates of the latter are size related variables like bilateral sum of GDP (*LGDT*) and relative country size (*LSIM*). Only part of the

²Superscripts (-) and (+) denote the theoretically expected sign of the impact of the respective variable on bilateral exports. See also the next section for more details on the explanatory variables.

time-variant regressors ($X_1 = n \times k_1$) might be uncorrelated with the unobserved bilateral effects. Ignoring the mentioned sources of correlation and running a traditional GLS regression - depending on the size of correlation - results inconsistent estimates.

In order to overcome this problem, Hausman and Taylor (1981), henceforth HT, suggest to premultiply the model by $\Omega^{-1/2}$:

$$\Omega^{-1/2} y_{ijt} = \Omega^{-1/2} X_{ijt} \beta + \Omega^{-1/2} Z_{ij} \gamma + \Omega^{-1/2} \varepsilon_{ijt}$$

$$\tag{2}$$

where $\Omega^{-1/2}$ has a typical element $y_{ijt} - \theta_{ij}y_{ij}$ for Y_{ijt} , and similarly for X_{ijt} and Z_{ij} with $\theta_{ij}=1-\sigma_v/\sigma_{1ij}$. An estimate of σ_v is derived from the Within estimation of (1) and an estimate of $\sigma_{1ij}=(T_{ij}\sigma_{\mu}^2+\sigma_v^2)^{0.5}$ stems from a 2SLS (between) regression of the bilateral averages of the Within residuals of (1) on *Z* (i.e., *LDIST*) with X_1 as the instruments.³

(2) is estimated by 2SLS using a proper set of instruments (called A_{HT}) in order to overcome the endogeneity bias. Following Breusch *et al.* (1989), we use a feasible set of instruments, which is equivalent to the idea of HT and consists of the Within transformed X (i.e., $\tilde{X} = [\tilde{X}_1, \tilde{X}_2]$) and the means of X_1 in the bilateral dimension (i.e., \overline{X}_1), hence, in our case $A_{HT} = [\tilde{X}, \overline{X}_1]$.⁴ 2SLS is only identified if $k_1 \ge g$, and the HT estimator is over-identified and more efficient than the Within estimator, if $k_1 > g$.

Summing up, the 2SLS error components model has two possible advantages with respect to the fixed effects approach. First, it is possibly more efficient without any loss in consistency, if $k_1>g$, and the instruments are legitimate. Second, if trade (or FDI, etc.) potentials are to be projected, which is a conventional method e.g., in the area of integration research (see Wang and Winters, 1992; Baldwin, 1994; etc.) it is more appropriate, since out-of-sample prediction conceptually contradicts the idea of the Within estimator.

III. Data and Estimation Results

We run regressions on bilateral exports from OECD countries to other economies (including OECD and non-OECD countries) covering the period 1986-

³Amemiya and MaCurdy (1986) and Breusch *et al.* (1989) suggest even more efficient sets of instruments than Hausman and Taylor (1981). However, as compared to Hausman and Taylor, their models require more exogeneity assumptions and they are less suited for the analysis of unbalanced panel data. Subscript ij indicates that the panel may be unbalanced.

⁴Note that we do not consider any time-invariant effects, which are uncorrelated with μ_{ij} .

1997. Nominal exports in current USD (from OECD, Monthly Statistics of International Trade; IMF, Direction of Foreign Trade; and the Vienna Institute for International Economic Studies, hereafter WIIW) are converted using export price (IMF, International Financial Statistics; OECD, Economic Outlook; and WIIW) and exchange rate indices (IMF, International Financial Statistics; and WIIW) to obtain real values with 1995 as the base year. Real GDP figures are based on nominal GDP in current USD (OECD, Economic Outlook and National Accounts Volume 1; IMF, International Financial Statistics; and WIIW), GDP deflators (same sources as GDP) and exchange rate indices. Population numbers are collected from OECD (Economic Outlook and National Accounts Volume 1), IMF (International Financial Statistics) and WIIW.

Three economic freedom variables are included for the exporters and the importers each, which are provided by Economic Freedom Network (Economic Freedom of the World) and account for legal structure and property rights (Area IV of the database) and international exchange (part of Area VI of the respective database). These variables are partly based on (zero-to-ten) ratings and partly on continuous data. They measure a country's economic freedom in several respects and can be interpreted as export impeding or enforcing determinants to control for the observed propensity to trade. The corresponding variables are viability of contracts (LFV_{ib} , LFV_{ji}), rule of law (LFR_{ib} , LFR_{jt}), and taxes on traded goods in percent of exports and imports (LFT_{ib} , LFT_{ij}).

All variables are in logs. The data base only contains bilateral trade relations, which cover a period of at least five years within 1986 and 1997. After removing a couple of outliers exhibiting excess studentized residuals with an absolute value larger than 3.5 in the Within regression of (1), we come up with 7337 observations. All regressions include time dummies.

In a first step, we estimate the Within model, which treats bilateral effects (μ_{ij}) as fixed and gives consistent parameter estimates independent of the possible correlation between μ_{ij} and some of the regressors. We find a significant positive effect of each of the three Heckscher-Ohlin variables (*LGDT*, *LSIM*, *LRLF*) on bilateral exports, being in accordance with the New Trade Theory (compare Helpman and Krugman, 1985). As one would have expected, the viability of contracts in both the exporting and the importing country (*LFV*) positively affects trade relations. The same holds true for higher levels of the rule of law in the importer country (*LFR_{ji}*). Contrary to our expectations, higher taxes on traded goods in the exporter country are positively related to trade activity. The remaining parameters

cannot be estimated precisely.

Second, we estimate the GLS (random effects) model to obtain an estimate of distance (*LDIST*) and find that the remaining coefficients deviate to a large extent from their fixed effects counterparts. This is an indication for correlation of some of the explanatory variables with the unobserved bilateral effects. It leads to biased estimates not only for these parameters (reflected by the highly significant Hausman test statistics) but also for distance.

To give a sensitivity analysis, we estimate four regressions in the spirit of Hausman and Taylor (1981). Specification HT-I in Table 1 assumes only distance (*LDIST*) to be the source of correlation. In HT-II, we treat both *LDIST* and the bilateral sum of GDP (*LGDT*) as only *singly exogenous* (this terminology has been introduced by Cornwell *et al.*, 1992, and it is due to X_2). In the next column of Table 1 (HT-III), similarity in country size (*LSIM*) is treated as *singly exogenous* besides *LDIST* and *LGDT*. Remember that we consider a possible relation between these size-related variables and μ_{ij} , since they also reflect some information on time-invariant trade-cost-like factors as distance (e.g., transaction between large economies typically involve larger bilateral distances). HT-IV enlarges the set of possible *singly exogenous* variables by bilateral distance in relative factor endowments (*LRLF*).

On the one hand, from the HT-I estimates we can see that distance indeed seems to be related to μ_{ij} , which shows up in parameter estimates, which lie in between their Within and their simple GLS counterparts. On the other hand, The HT test for over-identification⁵ together with the remaining relatively large deviations from the Within parameters reveal that there are additional sources of correlation of the explanatory variables with the μ_{ij} . HT-II indicates that *LGDT* indeed seems to exhibit an important correlation with μ_{ij} , and the respective estimation results are far closer to the Within estimator than in HT-I. Moreover, the corresponding test for over-identification in contrast to HT-I does not reject the null-hypothesis of the legitimacy of our set of instruments, and the geometric mean of canonical correlations shows a closer average relation between the singly exogenous variables and the corresponding instruments (i.e., A_{HT}) than in HT-I.⁶ As country

⁵This test is based on the squared standard error of the Within regression $(\hat{\sigma}_v^2)$ and the difference between the estimates and the variance-covariance matrices (*VC*) of the parameters between the Within and the HT regression $(\hat{m} = \hat{q}'[VC(\hat{q})]^{-1}\hat{q}$, where $\hat{q} = \hat{\beta}_{HT} - \hat{\beta}_{Within}$ and $VC(\hat{q}) = VC\hat{\beta}_{Within} - VC(\hat{\beta}_{HT})$. Under the null hypothesis $\hat{\sigma}_v^2 \hat{m}$ is distributed as χ_d^2 with $d = k_1 - g$.

⁶Canonical variate analysis is a standard tool of dimensionality reduction in multivariate statistics. The geometric mean of canonical correlations is a compact information of the association between two sets of variables (see e.g., Baltagi and Khanti-Akom, 1990).

Explanatory Variables:	Within	GLS	HT-I ^{a)}	HT-II ^{b)}	HT-III ^{c)}	HT-IV ^{d)}
Distance (LDIST _{ij})	-	-0.866**)	-0.991 ^{**)}	-1.623**)	-1.794 ^{**)}	-1.896**)
	-	0.032	0.125	0.142	0.155	0.171
Bilateral Sum of GDP (LGDT _{ijt})	0.384**)	0.919**)	$0.603^{**)}$	0.435**)	$0.426^{**)}$	$0.418^{**)}$
	0.029	0.023	0.027	0.029	0.029	0.029
Similarity in Size $(LSIM_{ijt})$	$0.064^{**)}$	0.231**)	$0.127^{**)}$	$0.066^{**)}$	$0.086^{**)}$	$0.079^{**)}$
	0.024	0.019	0.022	0.023	0.024	0.024
Distance in Relative Factor Endowments $(LRLF_{ijt})$	0.092**)	0.214**)	0.140***)	0.096**)	0.113**)	0.106**)
	0.021	0.017	0.019	0.020	0.021	0.021
Exporter Viability of Contracts (LFV_{it})	0.220**)	$0.102^{*)}$	$0.170^{**)}$	0.246**)	0.255**)	0.254**)
	0.052	0.053	0.051	0.052	0.052	0.052
Importer Viability of Contracts (LFV $_{jt}$)	0.573**)	$0.408^{**)}$	$0.507^{**)}$	$0.560^{**)}$	$0.564^{**)}$	$0.564^{**)}$
	0.033	0.035	0.033	0.033	0.033	0.033
Exporter Rule of Law (LFR _{it})	-0.032	$0.249^{**)}$	0.068	0.018	0.017	0.016
	0.048	0.048	0.047	0.047	0.048	0.048
Importer Rule of Law (LFR_{jt})	0.133**)	0.227**)	0.163**)	0.130**)	0.126**)	0.125**)
	0.027	0.027	0.026	0.026	0.027	0.027
Exporter Taxes in Percent of Trade (LFT_{it})	1.239**)	$0.568^{**)}$	$0.864^{**)}$	$1.158^{**)}$	$1.190^{**)}$	1.206**)
	0.158	0.162	0.155	0.158	0.159	0.159
Importer Taxes in Percent of Trade (LFT_{jt})	-0.065	-0.239**)	-0.152 ^{**)}	-0.088*)	-0.085*)	-0.086*)
	0.050	0.051	0.049	0.050	0.050	0.050
Number of Observations (NT)	7337	7337	7337	7337	7337	7337
Number of Bilateral Relationships (N)	779	779	779	779	779	779
R ^{2 e)}	0.99	0.94	0.95	0.94	0.94	0.94
Estimate of $\hat{\sigma}_{\mu}^2$	-	0.89	1.59	1.72	1.76	1.78
Estimate of $\hat{\sigma}_{v}^{2}$	0.28	0.28	0.28	0.28	0.28	0.28
Estimated average θ_{ij} = 1 - σ_v / σ_{1ij}	-	0.90	0.94	0.95	0.95	0.95
Estimated minimum θ_{ij}	-	0.87	0.93	0.93	0.93	0.94
Estimated maximum θ_{ij} Wald Tests: ^{f)}	-	0.91	0.95	0.95	0.95	0.95
Fixed Bilateral Effects (μ_{ij}): F(N = 778)	185.65**)	-	-	-	-	-
Fixed Time Effects (λ_t): $F(T = 11)$	84.91**)	682.44**)	$80.97^{**)}$	81.80**)	81.83**)	81.20**)
Hausman (Fixed versus Random Effects): $\chi^2(22)$) -	1969.46**)	-	-	-	-
Over-identification: $\chi^2(k_1-1)^{g}$	-	-	27.54 ^{**)}	7.91	8.26	10.32 ^{*)}
Canonical Correlationsh)	-	-	0.44	0.57	0.66	0.70

Table 1. Gravity model regression results (dependent variable is log of bilateral exports)

a) Only LDIST is singly exogenous, k_1 =9+T. - b) LGDT and LDIST are singly exogenous, k_1 =8+T. - c) LGDT, LSIM and LDIST are singly exogenous, k_1 =7+T. - d) LGDT, LSIM, LRLF and LDIST are singly exogenous, k_1 =6+T. - e) Calculated from ESS/TSS; for GLS and HT-I - HT-IV from the transformed model. - f) Based on the estimated variancecovariance matrix of the respective estimator; degrees of freedom = 6538 for the Within estimator and 7315 otherwise. - g) Based on Hausman and Taylor (1981), see a), b), c), and d) for information on k1. - h) Geometric mean of canonical correlations between the endogenous variables and the set of instruments, see Baltagi and Khanti-Akom (1995).

*) significant at 10 percent. - **) significant at 5 percent. Time dummies and constant not reported in order to save space. Standard errors are emphasized and reported below coefficients.

size in terms of bilateral GDP (*LGDT*) faces some correlation, we extend the set of singly exogenous variables by relative country size (*LSIM*) as well and find our presumption confirmed. Specification HT-III comes rather close to our Within estimates, it works well in terms of average canonical correlation, and the over-

On the Problem of Endogenous Unobserved Effects in the Estimation of Gravity Models 189

identification test again fails to reject the hypothesis of the appropriateness of the set of instruments. Hereafter, the treatment of distance in relative factor endowments (*LRLF*) in HT-IV as *singly exogenous* cannot improve the results.

The estimated parameter values in regressions HT-II to HT-IV are very close both to each other and to the Within parameters. Regarding the test statistics, HT-III is the most preferred specification. We may conclude from this analysis that distance and country size are the most important sources of correlation with the unobserved bilateral effects.⁷ Accounting for this relation and applying 2SLS techniques in the spirit of HT allows us to consistently estimate the impact of distance on bilateral exports, and this approach is more efficient than the Within framework. The coefficient of distance amounts to about -1.7, which in absolute value is about twice as high as its biased simple GLS counterpart and markedly (on average about two to three times) higher than the estimate found by other authors (see Oguledo and MacPhee, 1994, for an overview about the majority of cross-section estimates; and Baldwin, 1994; and Egger, 2000; for panel estimates). The difference between our and other panel data studies is supposed to have two main sources. First, previous research did not account for the fact of the superiority of a two-way framework (with time and bilateral effects) vis-à-vis a three-way setting (with exporter, importer and time effects). Second, the potential correlation between the explanatory variables and the unobserved effects has been ignored.

IV. Conclusions

A generalized gravity panel data model accounts for time and bilateral effects. Such a framework does not allow for a direct Within estimation of the parameters of time-invariant variables, where distance is the most prominent example of.

We propose to estimate gravity panel models by a 2SLS error components approach in the spirit of Hausman and Taylor (1981). (i) This set-up has several important advantages:

1. It allows to account for possible correlations of some of the explanatory variables with the unobserved effects, which in contrast to traditional GLS gives consistent parameter estimates.

⁷Regarding the estimated θ_{ij} values, this result is not so surprising. As compared to the GLS approach, the estimated θ_{ij} are larger and exhibit smaller variance in the HT models. This drives the random transformed variables closer to the Within transformed ones. Hence, the resulting parameter estimates are also closer to the Within model than to the traditional random effects (GLS) approach.

2. In contrast to the Within estimator, it is able to derive parameter estimates for the time-invariant variables, even in a setting with time and bilateral effects. Moreover, it is more efficient thant the fixed effects approach.

3. It is more appropriate, if trade (FDI, etc.) potentials are to be projected out of sample.

We show for a large panel of OECD and non-OECD countries that correlation between the explanatory variables and the unobserved bilateral effects matters a lot and gives heavily biased GLS estimates. Our Hausman and Taylor-type estimates lead us to the conclusion that country size and distance are the most important sources of this correlation, which has been ignored in previous research. With a value of -1.7, the properly estimated parameter of distance is about twice as high in absolute value as its biased simple GLS counterpart and between about twice and triple as high as most of the estimates found by other authors. This indicates that exports "die" twice to triple as fast with distance than previously assumed.

Acknowledgements

I am indebted in Badi Baltagi, Wilhelm Kohler and Michael Pfaffermayr for helpful suggestions and discussions. Additional comments by Hartmut Egger, Robert Kunst, Peter Steiner and an anonymous referee are gratefully acknowledged.

Received 6 February 2002, Accepted 17 August 2002

References

- Amemiya, T., MaCurdy, T.E. (1986) Instrumental-Variable Estimation of an Error-Components Model, *Econometrica*, 54, 869-880.
- Baldwin, R.E. (1994) Towards an Integrated Europe, CEPR, London.

Baltagi, B.H. (1995) Econometric Analysis of Panel Data, Wiley, Chichester.

- Baltagi B.H., Khanti-Akom, S. (1999) On Efficient Estimation With Panel Data, An Empirical Comparison of Instrumental Variables Estimators, *Journal of Applied Econometrics*, 5, 401-406.
- Breusch, T.S., Mizon, G.E., Schmidt, P. (1989) Efficient Estimation Using Panel Data, *Econometrica*, **57**, 695-700.

Cornwell, C., Schmidt, P., Wyhowski, D. (1992) Simultaneous Equations and Panel Data, *Journal of Econometrics*, **51**, 151-181.

Egger, P. (2000) A Note on the Proper Econometric Specification of the Gravity Equation,

On the Problem of Endogenous Unobserved Effects in the Estimation of Gravity Models 191

Economics Letters, 66, 25-31.

- Egger, P., Pfaffermayr, M. (2002) The Proper Econometric Specification of the Gravity Equation: A Three-Way Model With Bilateral Interaction Effects, *Empirical Economics*, forthcoming.
- Helpman, E., Krugman, P.R. (1985) *Market Structure and Foreign Trade*, MIT Press, Cambridge, Mass.
- Hummels, D., Levinsohn, J. (1995) Monopolistic Competition and International Trade: Reconsidering the Evidence, *Quarterly Journal of Economics*, **110**, 799-836.
- Hausman, J.A., Taylor, W.E. (1981) Panel Data and Unobservable Individual Effects, *Econometrica*, **49**, 1377-1398.
- Oguledo, V.I., MacPhee, C.R. (1994) Gravity Models: A Reformulation and an Application to Discriminatory Trade Arrangements, *Applied Economics*, **26**, 107-120.
- Wang, Z.K., Winters, A.L. (1992) The Trading Potential of Eastern Europe, *Journal of Economic Integration*, 7, 113-136.