Estimating the Gravity Models of Trade for EU Panel Data

Lawrence Agyepong (Student ID: 11350279)

Econ809: Project 4

Due: March 31, 2023

1.0 Introduction

This project aims to estimate gravity models for the European Union using various panel data analysis techniques. The dataset, comprising 1960-2001 bilateral trade flows between 15 EU members, includes 91 country pair observations, 42-time observations, six time-varying and three time-invariant regressors, and six time-specific common factors. I follow Serlenga & Shin (2007)to estimate pooled regression, fixed effects, random effects, and Hausman-Taylor models. The results show all explanatory variables significantly impact trade as expected. Furthermore, I estimate the first difference between estimator models for robustness. The project is organized into six sections: section 2 describes the gravity model; section 3 discusses the empirical specification of the models; section 4 describes the data used; section 5 presents and discusses the results; and section 6 concludes.

2.0 Overview of the Gravity Model

The gravity model of international trade flows serves as a fundamental model for assessing the effects of various policy matters, such as regional trading groups, currency unions, and trade distortions (Serlenga & Shin, 2007). This model, inspired by the concept of gravity in physics, has been employed extensively in international trade to predict and elucidate trade flows between nations. In physics, the attractive force between two objects is contingent upon their mass and the distance separating them. Similarly, within the realm of international trade, the gravity model posits that the trade volume between two countries is positively correlated with the size of their economies and negatively correlated with the distance between them.

Serlenga & Shin (2007) attribute the empirical success of the gravity model to its capacity to encompass most of the empirical phenomena typically witnessed in international trade. The gravity relationship arises from any trading model featuring trade costs that increase with distance. Chaney (2018) points out that the majority of empirical studies utilizing the gravity model prioritize economic size over distance. However, it is important to note that the distance factor significantly impacts bilateral trade flows between two countries. In its simplest form, the Gravity Model can be expressed as:

Trade_ij = (GDP_i × GDP_j) / Distance_ij

Where: Trade_ij represents the trade flow between country i and country j; GDP_i and GDP_j are the respective Gross Domestic Products of country i and country j; and Distance_ij represents the distance between the two countries. The gravity model has gained extensive use in international trade literature and empirical analyses, primarily due to its proficiency in predicting trade patterns accurately. It elucidates why nations with larger economies and closer geographical proximity are more inclined to engage in trade with each other. Furthermore, the model has been expanded to encompass other factors that influence trade, including shared languages, common borders, colonial connections, and trade agreements. These supplementary factors contribute to explaining trade patterns that the basic Gravity Model may not entirely capture.

3.0 Empirical Model Specification

Early empirical studies that attempt the estimate the gravity model have relied on cross-section estimation techniques. However, it is widely acknowledged that cross-section OLS estimation overlooks heterogeneous characteristics associated with bilateral trade relationships. Consequently, these cross-section estimates may not fully account for the heterogeneous factors, potentially resulting in significant heterogeneity bias. In recent studies, the panel data approach has been favored, as it enables the modeling of such heterogeneity by incorporating country pair 'individual' effects. Brief descriptions of the empirical specifications for the OLS Pooled regression, Fixed Effects, and Random Effects models used to estimate the gravity model are provided below. Additionally, other models, such as the Hausman-Taylor, First Difference, and Between Estimator, are briefly discussed.

3.1 Pooled Regression (POLS)

The method of pooling all observations together in panel data essentially treats it as a straightforward cross-sectional dataset. This approach disregards the panel structure of the data, which entails multiple observations for each country pair over time, as well as any potential unobserved heterogeneity across countries. If unobserved country-specific factors are correlated with the explanatory variables, this method may yield biased estimates. The model specification is as follows:

$$Trade\_ijt = \beta 0 + \beta 1 * GDP\_it + \beta 2 * GDP\_jt - \beta 3 * DIST\_ij + X\_ijt' * \beta 4 + \varepsilon\_ijt$$

Where t represents time, X_ijt' is a vector of other control variables, and ε_ijt is the error term.

3.2 Fixed Effects (FE)

The fixed effects method accommodates unobserved country-specific effects by incorporating fixed effects for each country pair within the model. These fixed effects capture time-invariant factors that impact trade flows between countries. This method generates consistent estimates, even when unobserved country-specific factors correlate with the explanatory variables. However, a limitation of the fixed effects model is its inability to estimate the effects of time-invariant explanatory variables (e.g., distance), as they are absorbed by the fixed effects. The model specification is:

$$Trade\_ijt = \alpha\_ij + \beta 1 * GDP\_it + \beta 2 * GDP\_jt - \beta 3 * DIST\_ij + X\_ijt' * \beta 4 + \varepsilon\_ijt$$

Where α_ij is the fixed effect for each country pair (i, j), capturing unobserved time-invariant heterogeneity.

3.3 Random Effects (RE)

The random effects method assumes that unobserved country-specific effects are random and uncorrelated with the explanatory variables. Compared to the fixed effects model, the random effects model enables more efficient estimation. However, the assumption that random effects are uncorrelated may not always hold in practice. The model specification is:

$$Trade\_ijt = \beta 0 + \beta 1 * GDP\_it + \beta 2 * GDP\_jt - \beta 3 * DIST\_ij + X\_ijt' * \beta 4 + \alpha\_ij + \varepsilon\_ijt$$

Where α_ij is the random effect for each country pair (i, j), assumed to be uncorrelated with the explanatory variables.

In estimating the Gravity Model with panel data, the choice of estimation technique hinges on the research question, data structure, and assumptions concerning unobserved heterogeneity. Pooled regression is the most straightforward method but may yield biased estimates if unobserved country-specific factors correlate with the explanatory variables. FE models address unobserved time-invariant heterogeneity, but they cannot estimate the effects of time-invariant explanatory variables. In contrast, RE models are more efficient and allow for the estimation of time-invariant variables, but they rely on the assumption of uncorrelated random effects. Prior studies often employ the Hausman test to decide between FE and RE models, as it examines the validity of the random effects assumption. If the null hypothesis of the Hausman test is rejected, the FE model is favored; otherwise, the RE model is deemed more suitable.

3.4 Hausman & Taylor Model (HT)

The model overcomes certain limitations of the fixed effects and random effects models by including both time-variant and time-invariant variables while addressing potential correlations between the independent variables and the unobserved, time-invariant factors. The model employs instrumental variables (IV) estimation to tackle the correlation between endogenous variables and unobserved effects. Time-variant and exogenous variables are used as instruments for time-variant and endogenous variables, while time-invariant and exogenous variables serve as instruments for time-invariant and endogenous variables. The Hausman-Taylor model offers a more efficient estimation compared to the fixed effects model when time-invariant variables are of interest, and it allows for the inclusion of time-invariant variables even when they correlate with the unobserved individual effects. However, the model depends on the availability of a sufficient number of valid instruments and assumes that the selected instruments are uncorrelated with the error term.

Applying a Hausman-Taylor model to the gravity model in a panel data setting, where trade data is available for multiple country pairs over time, results in the following basic structure for the Hausman-Taylor gravity model:

$$Trade\_it = β\_1 * GDP\_i * GDP\_j + β\_2 * DIST\_ij + β\_3 * X1\_ijt + β\_4 * X2\_ij + u\_i + ε\_it$$

where:

- T_it represents the trade flow between country i and country j at time t
- GDP_i and GDP_j are the GDPs of countries i and j, respectively
- DIST_ij is the distance between countries i and j
- X1_ijt represents time-variant variables (e.g., GDP, exchange rates)
- X2_ij represents time-invariant variables (e.g., common language, shared border)
- β_1, β_2, β_3, and β_4 are the coefficients associated with these variables
- u_i is the unobserved, time-invariant effect for the country pair i
- ε_it is the error term

3.5 First Difference

An alternative approach to estimating panel data involves taking the first difference of both dependent and explanatory variables. This method is assumed to eliminate unobserved heterogeneity when the data is first differenced. By taking the first difference, changes in the data from one period to the next are considered, effectively removing all time-invariant variables. This technique has the advantage of eliminating latent heterogeneity that may arise from fixed effects or random effects. On the downside, the model only accounts for time-varying cross-section units. Assuming strict endogeneity, the first difference estimator is unbiased and consistent. Although our data spans multiple periods, the first difference method is employed in this study to compare results with other panel data models discussed previously.

3.6 Group Means (Between Estimator)

The Group Means panel estimator, also known as the Between Effects estimator, is an econometric method employed in panel data analysis. Contrasting with Fixed Effects (within) and Random Effects estimators, which emphasize within-entity variation, the Group Means estimator focuses on between-entity variation. This entails that the Group Means estimator examines the disparities in the averages of the dependent variable and the independent variables across distinct entities (e.g., individuals, firms, or countries) in the panel data. The Group Means estimator proves especially valuable when the research question centers on the association between the average values of the dependent and independent variables across entities, rather than the relationship within entities over time.

4.0 Data Description

The dataset under analysis encompasses bilateral trade flows among 15 EU member nations, including Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxemburg, Netherlands, Portugal, Spain, Sweden, and the United Kingdom. Spanning a timeframe of 42 years from 1960 to 2001, it comprises ninety-one country pair observations, forty-two-time observations, six time-varying regressors, three time-invariant regressors, and six time-specific common factors. The dependent variable in this study is the sum of logged exports and imports of bilateral trade flow. Among the six time-varying regressors are the sum of the logged real gross domestic product, the degree of similarity between two trading nations, relative factor endowments, and the logged bilateral real exchange rate. Additionally, a dummy variable equal to 1 is used to indicate when both countries are part of the European Community and when both countries adopt the common currency. Table 1 presents a description of the main variables used in this project.

Table 1: Description of Variables

Variable	**Description**
TRADE	the sum of logged exports and imports, bilateral trade flow
GDP	the sum of the logged real GDPs
SIM	a measure of similarity between two trading countries
RLF	a measure of relative factor endowments
RER	the logged bilateral real exchange rate
CEE	a dummy equal to 1 when both belong to European Community
EMU	a dummy equal to 1 when both adopt the common currency
DIST	the geographical distance between capital cities
BOR	a dummy equal to 1 when the trading partners share a border
LAN	a dummy equal to 1 when both speak the same language

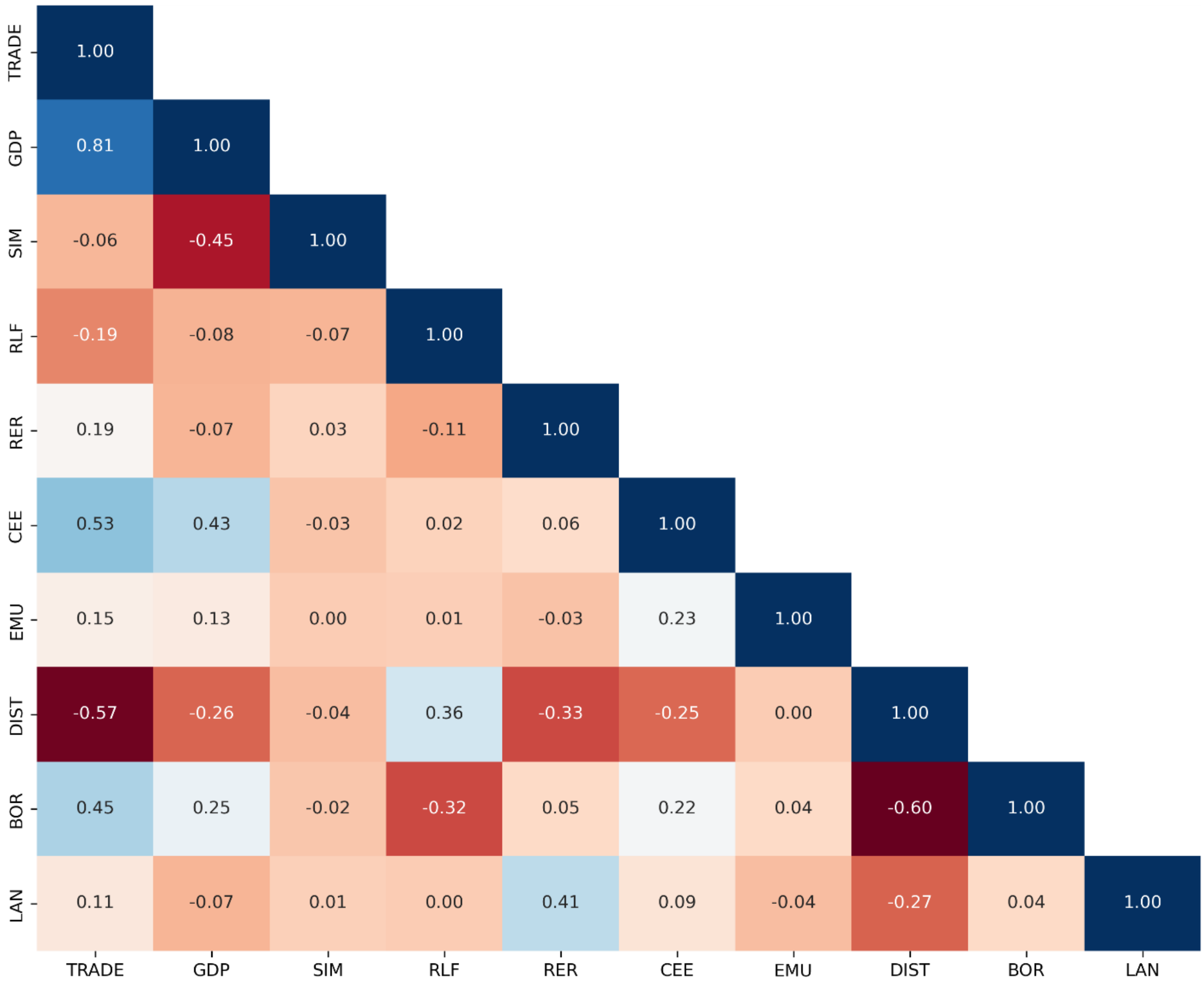
5.0 Results

In this section, the results of the analysis are presented and discussed. First, I present the correlation between the variables used in the study. Next, I present the results from the pool regression model, fixed effects, random effects, and Hausman-Taylor models. Furthermore, I present the first differenced and group means models to check the robustness of the other models.

5.1 Correlation between the variables of the study

Figure 1 shows the correlation between the variables of the study. A cursory look at the figure shows a considerable association between the variables. Our interest is how the independent variables correlate with trade volume. We observe that the most important predictor of trade between two countries is the sum of their GDPs i.e., GDP has the highest correlation coefficient (0.81) compared to other variables. This suggests that higher GDP is associated with higher trade volume. The geographical distance between the two trading countries is another important predictor of trade volumes as shown in the figure. The closer the countries, the higher the trade volume. Membership of both trading countries to the European Community is positively associated with high trade volume with a correlation coefficient of 0.53. Also, countries sharing a common border have high trade volumes. Other factors such as bilateral real exchange rate, and having a common currency and language are positively correlated with trade, while the similarity between the two trading countries and relative measure of factor endowments associates negatively with trade.

Figure 1: Correlation between the variables of the study



5.2 Panel Regression Model

Table 2 presents the results of the estimations of all the models. The models of most interest are the pooled ols regression (POLS), fixed effects (FE), and random effects (RE) models. The Hausman-Taylor models (HT-IV1 and HT-IV2) were replicated as in Serlenga & Shin's (2007) paper. The first differenced (FD) and Group Means or between estimator (BE) was included to serve as a robustness check. The pooled regression results show that all the variables significantly affect trade. We notice that the sum of logged GDP has the most effect on trade across all the models since its coefficient has the highest magnitude. Also, the signs of bilateral real exchange rate, relative factor endowments, the similarity between the countries, European community membership, common currency, distance, and common language are consistent across models 1 to 5.

The pooled regression estimation results are likely to be biased since it does not account for the individual heterogeneity amount the countries. The fixed effects model takes into account the heterogeneity in bilateral trade. It assumes that the independent variables are correlated with the individual fixed effects and hence produces consistent estimates. However, the FE model wipes out all time-invariants variables such as distance, border, and common language from the model. Unlike the fixed effects model, the random effects model assumes that the unobserved individual heterogeneity effects are uncorrelated with the independent variables. Hence, the random effects model allows us to estimate the parameters of both time-varying and time-invariant variables.

Table 2: Estimation Results of the various Panel Models

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Trade	POLS	FE	RE	HT-IV1	HT-IV2	FD	BE
RER	.1***	.06***	.07***	.06***	.06***		.09***
	(0)	(.01)	(.01)	(.01)	(.01)		(.02)
GDP	1.58***	1.81***	1.79***	1.81***	1.81***		1.54***
	(.01)	(.02)	(.02)	(.02)	(.02)		(.08)
RLF	.03***	.03***	.03***	.03***	.03***		.01
	(.01)	(.01)	(.01)	(.01)	(.01)		(.06)
SIM	.88***	1.17***	1.14***	1.17***	1.2***		.81***
	(.02)	(.06)	(.05)	(.06)	(.05)		(.1)
CEE	.32***	.31***	.32***	.31***	.31***		-.03
	(.02)	(.02)	(.02)	(.02)	(.02)		(.24)
EMU	.2***	.09***	.09***	.09***	.09***		-1.6
	(.05)	(.03)	(.03)	(.03)	(.03)		(1.84)
DIST	-.65***		-.59***	-.38**	-.37**		-.71***
	(.02)		(.12)	(.19)	(.19)		(.14)

BOR	.52***		.44**	.60**	.61**		.61***
	(.03)		(.19)	(.26)	(.26)		(.2)
LAN	.23***		.42**	1.56**	1.56**		.24
	(.03)		(.18)	(.71)	(.68)		(.2)
D.RER						.31***	
						(.01)	
D.GDP						2***	
						(.09)	
D.RLF						0	
						(0)	
D.SIM						.84***	
						(.1)	
D.CEE						.06***	
						(.01)	
D.EMU						-.02	
						(.02)	
CONSTANT	-10.95***	-18.21***	-13.93***	-18.21***	-18.21***	0	-9.71***
	(.25)	(.29)	(.89)	(.29)	(.29)	(0)	(1.55)
Observations	3822	3822	3822	3822	3822	3731	3822

Standard errors are in parentheses
**** p<.01, ** p<.05, * p<.1*

To decide between fixed effects and random effects models, I employ the Hausman Test which assesses the validity of the random effects model's assumption of no correlation between independent variables and the individual heterogeneity effects. Rejecting the null of the Hausman test implies that the fixed effects model is the appropriate specification. Figure 2 presents the results of the Hausman test. We reject the null hypothesis at the 5% significance level; thus, the fixed effects model is the rights specification. Below, I discuss the results from the fixed effects model.

Figure 2: Hausman Test

```
. hausman fe re

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |       fe           re         Difference       Std. err.
-------------+----------------------------------------------------------------
         rer |   .0609803     .0689902       -.0080099        .0034872
         gdp |   1.812493     1.794872        .0176209         .007598
         rlf |   .0325083     .0334115       -.0009033        .0010354
         sim |   1.172255     1.142659        .0295964        .0316216
         cee |   .3093359     .3181826       -.0088466        .0023945
         emu |   .0852087     .0926522       -.0074435               .
------------------------------------------------------------------------------
                           b = Consistent under H0 and Ha; obtained from xtreg.
            B = Inconsistent under Ha, efficient under H0; obtained from xtreg.

Test of H0: Difference in coefficients not systematic

    chi2(6) = (b-B)'[(V_b-V_B)^(-1)](b-B)
            =  13.34
Prob > chi2 = 0.0380
(V_b-V_B is not positive definite)
```

As shown in Table 2, all the time-varying variables in the fixed effects model have a highly significant effect on trade. A 10% increase in the sum of the GDP of the trading countries increases the trade flow by 18% on average. On average, a 10% increase in the bilateral real exchange rate will result in a 0.6% increase in trade volume between the two trading countries. This suggests that the export component of total trade in the home country is significantly larger than the import component. Relative factor endowments have a positive impact on trade flow implying that the larger the difference between the two countries' factor endowments, the higher the volume of inter-industry trade. Also, when the similarity index increases by 10%, trade flow increases by 11.7%. European Community membership and common currency also have a positive and significant effect on trade flows. This is expected since a single currency will reduce the transaction costs of trade within member countries.

Although the Hausman test indicates that the fixed effects model is preferred, it wipes out the time-invariant variables. However, these variables are important for our model. To address this issue, I estimate the Hausman-Taylor (HT) model as described in Serlenga & Shin (2007). Columns 4 and 5 of table 1 present the two HT models. The set of instrument variables used in the HT estimation is RER for IV1 and RER, TGDP, and RLF for IV2. The estimation results show that distance negatively affects trade flows whiles common language and border have a positive effect on trade flow. These results are consistent with the results from the random effects model.

I also estimate the first difference model and between estimators to assess the robustness of the FE model. All but the sign of EMU in the first difference model is not consistent with the results found in the fixed effects model. Also, RLF and EMU are not significant in the first difference model. In the BE model, the signs of the estimates are consistent with the other models except for the CEE and EMU which have a negative relationship with trade flows.

6.0 Conclusion

In conclusion, this project employed various estimation techniques, including pooled regression, fixed effects, random effects, and Hausman-Taylor models, to estimate gravity models for European Union trade. The Hausman test favored the fixed effects model, which demonstrated that all explanatory variables significantly influenced trade flows. It was found that the combined GDPs of home and foreign countries positively affected real total trade, while home currency depreciation led to increased trade flows. Additionally, the similarity in size promoted real trade flows, highlighting the importance of intra-industry trade in the EU. The impact of relative differences in factor endowments between trading partners on trade flows was small but significantly positive. Furthermore, both trade and currency union memberships significantly boosted real trade flows. These findings are consistent with those reported by Serlenga & Shin (2007) and reinforce the importance of considering various factors when analyzing EU trade patterns.

References

Chaney, T. (2018). The gravity equation in international trade: An explanation. Journal of Political Economy, 126(1), 150-177.

Serlenga, L., & Shin, Y. (2007). Gravity models of intra-EU trade: application of the CCEP-HT estimation in heterogeneous panels with unobserved common time-specific factors. Journal of applied econometrics, 22(2), 361-381.

Appendix

Appendix 1: Pooled Regression Model

trade	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
rer	.099	.004	25.65	0	.091	.106	***
gdp	1.579	.013	125.89	0	1.555	1.604	***
rlf	.032	.008	3.75	0	.015	.048	***
sim	.885	.017	52.47	0	.852	.918	***
cee	.318	.023	14.05	0	.273	.362	***
emu	.204	.051	3.99	0	.104	.305	***
dist	-.646	.022	-28.76	0	-.69	-.602	***
bor	.525	.034	15.39	0	.458	.592	***
lan	.234	.034	6.80	0	.166	.301	***
Constant	-10.947	.247	-44.31	0	-11.432	-10.463	***

Mean dependent var	4.895	SD dependent var	1.899
R-squared	0.905	Number of obs	3822
F-test	4027.901	Prob > F	0.000
Akaike crit. (AIC)	6778.581	Bayesian crit. (BIC)	6841.066

**** $p<.01$, ** $p<.05$, * $p<.1$*

Appendix 2: Fixed Effect Model

trade	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
rer	.061	.009	7.06	0	.044	.078	***
gdp	1.812	.02	91.46	0	1.774	1.851	***
rlf	.033	.008	4.12	0	.017	.048	***
sim	1.172	.056	21.10	0	1.063	1.281	***
cee	.309	.016	19.30	0	.278	.341	***
emu	.085	.027	3.18	.001	.033	.138	***
Constant	-18.206	.288	-63.26	0	-18.771	-17.642	***

Mean dependent var	4.895	SD dependent var	1.899
R-squared	0.898	Number of obs	3822
F-test	5448.322	Prob > F	0.000
Akaike crit. (AIC)	1376.337	Bayesian crit. (BIC)	1420.077

**** p<.01, ** p<.05, * p<.1*

Appendix 3: Random Effects Model

trade	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
rer	.069	.008	8.74	0	.054	.084	***
gdp	1.795	.018	98.06	0	1.759	1.831	***
rlf	.033	.008	4.27	0	.018	.049	***
sim	1.143	.046	25.01	0	1.053	1.232	***
cee	.318	.016	20.08	0	.287	.349	***
emu	.093	.027	3.46	.001	.04	.145	***
dist	-.591	.117	-5.06	0	-.82	-.362	***
bor	.441	.19	2.32	.02	.068	.815	**
lan	.417	.185	2.26	.024	.055	.78	**
Constant	-13.93	.889	-15.68	0	-15.672	-12.189	***

Mean dependent var	4.895	SD dependent var	1.899
Overall r-squared	0.898	Number of obs	3822
Chi-square	33369.017	Prob > chi2	0.000
R-squared within	0.898	R-squared between	0.900

**** p<.01, ** p<.05, * p<.1*

Appendix 4: Group Means

trade	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
rer	.087	.023	3.69	0	.04	.133	***
gdp	1.541	.08	19.22	0	1.381	1.7	***
rlf	.011	.06	0.18	.855	-.109	.131	
sim	.806	.104	7.74	0	.598	1.013	***
cee	-.03	.244	-0.12	.901	-.516	.455	
emu	-1.596	1.842	-0.87	.389	-5.26	2.068	
dist	-.71	.136	-5.24	0	-.98	-.441	***
bor	.609	.201	3.02	.003	.208	1.009	***
lan	.236	.201	1.17	.245	-.165	.637	
Constant	-9.706	1.552	-6.25	0	-12.795	-6.618	***

Mean dependent var	4.895	SD dependent var	1.899
R-squared	0.914	Number of obs	3822
F-test	96.178	Prob > F	0.000
Akaike crit. (AIC)	147.870	Bayesian crit. (BIC)	210.356

**** p<.01, ** p<.05, * p<.1*

Appendix 5: First Difference

D.trade	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
D	.307	.012	25.42	0	.283	.33	***
D	2.003	.093	21.60	0	1.821	2.185	***
D	.001	.005	0.17	.865	-.009	.011	
D	.835	.103	8.08	0	.633	1.038	***
D	.063	.015	4.30	0	.034	.092	***
D	-.021	.018	-1.19	.233	-.055	.013	
Constant	-.005	.004	-1.25	.21	-.012	.003	

Mean dependent var	0.069	SD dependent var	0.152
R-squared	0.280	Number of obs	3731
F-test	241.933	Prob > F	0.000
Akaike crit. (AIC)	-4686.560	Bayesian crit. (BIC)	-4642.989

**** p<.01, ** p<.05, * p<.1*

Appendix 6: Hausman Taylor Estimation (HT- IV1)

```
Hausman–Taylor estimation                       Number of obs     =      3,822
Group variable: id                              Number of groups  =         91

                                                Obs per group:
                                                              min =         42
                                                              avg =         42
                                                              max =         42

Random effects u_i ~ i.i.d.                     Wald chi2(9)      =   32924.65
                                                Prob > chi2       =     0.0000
```

trade	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
TVexogenous						
rer	.0609803	.0086367	7.06	0.000	.0440526	.077908
TVendogenous						
gdp	1.812493	.0198284	91.41	0.000	1.77363	1.851356
rlf	.0325083	.0078923	4.12	0.000	.0170397	.0479769
sim	1.172255	.0555967	21.08	0.000	1.063288	1.281223
cee	.3093359	.0160348	19.29	0.000	.2779083	.3407636
emu	.0852087	.0268029	3.18	0.001	.032676	.1377415
TIexogenous						
dist	-.3815958	.1917947	-1.99	0.047	-.7575065	-.0056851
bor	.6013204	.2586476	2.32	0.020	.0943805	1.10826
TIendogenous						
lan	1.558674	.7070635	2.20	0.027	.1728551	2.944493
_cons	-15.75991	1.502941	-10.49	0.000	-18.70562	-12.81419
sigma_u	.65096562					
sigma_e	.29268254					
rho	.83184173	(fraction of variance due to u_i)				

Note: TV refers to time varying; TI refers to time invariant.

Appendix 7: Hausman Taylor Estimation (HT-IV2)

```
. xthtaylor trade rer gdp rlf sim cee emu dist bor lan, endog(sim cee emu lan)

Hausman-Taylor estimation                        Number of obs     =      3,822
Group variable: id                               Number of groups  =         91

                                                 Obs per group:
                                                               min =         42
                                                               avg =         42
                                                               max =         42

Random effects u_i ~ i.i.d.                      Wald chi2(9)      =   33167.80
                                                 Prob > chi2       =     0.0000

------------------------------------------------------------------------------
       trade | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
TVexogenous  |
         rer |   .0625332   .0085438     7.32   0.000     .0457877    .0792786
         gdp |   1.805507   .0192419    93.83   0.000     1.767794    1.843221
         rlf |   .0325371   .0078427     4.15   0.000     .0171656    .0479086
TVendogenous |
         sim |   1.197439   .0528833    22.64   0.000     1.093789    1.301088
         cee |   .3109107   .0159875    19.45   0.000     .2795758    .3422455
         emu |   .0855045   .0267842     3.19   0.001     .0330084    .1380005
TIexogenous  |
        dist |  -.3775983   .1913561    -1.97   0.048    -.7526495   -.0025472
         bor |   .6101144   .2636025     2.31   0.021     .0934629    1.126766
TIendogenous |
         lan |   1.559897   .6754906     2.31   0.021     .2359596    2.883834
             |
       _cons |  -15.66188   1.496354   -10.47   0.000    -18.59468   -12.72908
-------------+----------------------------------------------------------------
     sigma_u |  .67259975
     sigma_e |  .29268254
         rho |  .84079079   (fraction of variance due to u_i)
------------------------------------------------------------------------------
Note: TV refers to time varying; TI refers to time invariant.
```

Stata Code:

// Project 4 - Gravity Model

clear

```
import delimited "C:\Users\Lawrence Agyepong\OneDrive - University of Saskatchewan\Usask Econ\Winter 2023\Econometrics 2\Projects\Project 4\gravity_data.csv", numericcols(1)

xtset id year, yearly
//Table II: Alternative Panel Data
// Pooled Regression - OLS
reg trade rer gdp rlf sim cee emu dist bor lan
// FE Model
xtreg trade rer gdp rlf sim cee emu, fe
est store fe
// RE Model
xtreg trade rer gdp rlf sim cee emu dist bor lan, re
est store re
// Hausman Test
hausman fe re
// Other models
// Group Means (Between Effects)
xtreg trade rer gdp rlf sim cee emu dist bor lan, be
// First difference
reg d.trade d.rer d.gdp d.rlf d.sim d.cee d.emu
// Hausman and Taylor Model
xthtaylor trade rer gdp rlf sim cee emu dist bor lan, endog(gdp rlf sim cee emu lan) // Instrument is rer
xthtaylor trade rer gdp rlf sim cee emu dist bor lan, endog(sim cee emu lan)
xtoverid
```