University of Saskatchewan

RESEARCH PAPER FOR ECON 809

Gravity Models of Intra-EU Trade: Application to Heterogeneous Panels with Unobserved Common Time-Specific Factors for 91 Country Pairs.

NANA KWAKU ASAMOAH

ID: 11275489

1. INTRODUCTION

Panel data provides information on the behavior of individual units, both across the individual units and over time. Hence, they have both cross-sectional and time series dimensions. Panel data enable us to control for heterogeneity among said individual units and gives us more information about the data in terms of variability and allows for greater degrees of freedom as well as enhance greater efficiency. In panel data sets, we would ordinarily assume that whereas for a given individual unit there is correlation over time, however across units there is less collinearity among the different individual units. Panel data also allows us to account for aggregation biases as well as enables the researcher to examine dynamics in relation to adjustments.

Owing to the numerous benefits of using panel data, it is not surprising that it is popular in economic literature. However, its popularity is thwarted by the fact that its limitations act as stumbling blocks in carrying out analysis. The process of design and data collection in itself is a cumbersome process and usually the time series dimension is of short-panel since it is more difficult to obtain data on one individual over a very long period of time. And even if said data were obtained there might arise distortion of measurement errors, selectivity problems and cross-section dependence acting as hindrance to rightful inference. However despite these drawbacks, the importance of panel data cannot be overemphasized.

Empirically, the use of conventional cross-section estimation is criticized since it is not able to deal with bilateral heterogeneity. Much recent empirical studies have therefore emphasized the importance of explicitly allowing for the presence of time-specific effects in order to capture business cycle effects or deal with globalization issues. This brings us to the idea of panel data methods.

This project uses panel data to specify the “gravity model” of international trade and to estimate the effects of various pertinent variables to the volume of trade. It is well known that the flows of international trade can be well described by a ‘gravity equation’ in which bilateral trade flows are measured as a log-linear function of the incomes of and distance between trading partners. Indeed, the gravity equation modelling is one of the great success stories in empirical economics. Modelling and predicting foreign trade flows is considered for a long time an important task in international economics. One of the most fruitful ways to formalize this has been through the use of gravity type models.

One of the inevitable issues we face in conducting said analysis is that we do not know whether the unobserved time-invariant variables are correlated with the time-varying variables. In such conditions, the fixed effect model is suggestively most appropriate since it is almost always consistent no matter whether the true DGP is fixed or random. However with fixed effect models, all time-invariant variables get dropped.

Cheng and Wall (2002) argued that fixed-effects models are more appropriate than other panel data models. They argue that the major drawback of the fixed-effect model, that is not being able to estimate coefficients on time-invariant variables can be dealt with by estimating the regression of the individual-specific effects on the individual-specific variables by OLS. This however might not be a realistic approach since it ignores the potential correlation between the repressors and the unobserved heterogenic individual effects which may lead to biased estimates.

In contrast, Serlenga and Shin (2007) employed the Hausman-Taylor instrumental variable technique in which they capture certain degrees of cross-sectional dependence through heterogeneous time-specific factors to avoid the bias of uncorrected estimates. Their results indicated that their approach fits the data reasonably well and their estimation yields sensible results in comparison to the traditional approaches.

In this project, I thoroughly analyze the impact of the different variables on the volume of trade through the different models and draws conclusions on which is most appropriate to carry out the relevant inferences.

The Gravity model in detail

The gravity model is a commonly used model for estimating the impact of a wide range of policy issues in international trade in a number of related topics such as currency unions, regional trading groups and various trade distortions. In the context of international trade the gravity model has been applied since 1940s to explain the determinants of a diverse range of flows. It states that the size of trade flows between different countries is a factor of supply conditions in the point of origin, demand conditions in the point of destination and whether the forces related to the trade flows act as impediments or act as instigators between the two countries.

Despite its popularity in empirical use, until the seminal paper by Anderson (1979) the gravity model was criticized heavily on the basis that it lacked a theoretical foundation.

However, afterwards it became clearer that the theoretical grounds for the gravity model can be obtained by the classical trade models such as the Ricardian Model or the Hecksher Ohlin Model or even the new trade theory models such as Increasing Returns to Scale.

The gravity model does no derive its strength from its ability to test the validity of these three trade models against each other, but rather its capacity to reconcile theory and empirics in international trade by incorporating most of the empirical phenomena.

In recent decades this model has gained popularity as well as seen a lot of controversy in terms of its econometric specification. Many argue that using traditional cross-sectional specification may be inefficient since it might not be able to capture the issues of heterogeneity which comes with bilateral trade flows, thus a panel-based specification may be preferred. A panel-based approach may be able to capture heterogeneity issues by incorporating country-pair 'individual effects' and then test the assumption that unobserved individual effects are correlated with all the explanatory variables.

2. METHOD AND DATA

Many argued that the appropriate specification for a panel gravity model should consist of main (exporter, importer, and time) as well as time invariant exporter-by-importer (bilateral) interaction effects. The econometric representation of the gravity model I will use in this project takes the form of a triple-indexed model which is given below:

$$\boldsymbol{lnY_{ijt} = \alpha_i + \gamma_j + \lambda_t + \beta_1 lnX_{ijt} + \beta_2 lnX_{it} + \beta_3 lnXjt + \beta_4 lnZ_{ij} + U_{ijt}}$$

For i, j = 1,……….,N i≠j t= 1,……, T

Where:

Y_{ijt}= volume of trade from home country I to target country j at time t

X_{ijt}= explanatory variables with variations in all three dimentions

X_{it}, X_{jt}= explanatory variables with variations in i or j and t.

Z_{ij}= explanatory variables that do not vary over time but might vary in i and j

α_i, γ_j= local country and target country effect respectively that might be correlated with other explanatory variables

λ_t= time effect common to all cross sectional units

U_{ijt}= white noise disturbance term

The major source of data for this project is the Journal of Applied Econometrics Data Archive. This was the same dataset used by Serlenga and Shin (2007) whose paper was published in the Journal of Applied Econometrics.

There 91 observations of country pairs in Europe (European Union) who trade with each other. The country samples include all of the 15 EU countries: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxemburg, Netherlands, Portugal, Spain, Sweden and United Kingdom where Luxemburg and Belgium are treated the same thereby resulting in the 91 pairs.

The observations cover the period from 1960 to 2001, there are 6 time varying explanatory variables, 3 time invariant explanatory variables and 6 time specific common factors. We must note that time invariant explanatory variables have zero within variation and individual invariant explanatory variables have zero between variations.

Since it is a panel data, the data organization is a little more intricate than what we have seen so far. The first 42 columns represent all time observations for the first country pair, the next 42 represents all time observations for the second country pair and so on until the 91st country pair of the whole sample. In order for easier understanding of the nature of the data, I provided the detailed description of the data in the below:

Trade: Sum of logged exports and imports, bilateral trade flow

Gdp: the sum of logged real gross domestic products

Sim: a variable measuring the similarity between two trading countries

Rlf: a measure of relative factor endowments

Rer: the logged bilateral real exchange rate

Cee: a dummy variable equal to 1 when both belong to European community

Emu: a dummy variable equal to 1 when both adopt the common currency

Dist: a geographical distance between the capital cities of the two countries

Bor: a dummy variable equal to 1 when the trading partners share a boarder

Lan: a dummy variable equal to 1 when both countries speak the same language

RERT: a log of real exchange rates between European currencies and the US dollar

ftrade, **fgdp**, **fsim**, **frlf** and **frer** are the time specific common factors (individual means) of the variables **trade, gdp, sim, rlf** and **rer**.

Between and Within Variation in the Data.

From the dataset obtained, I observe that the first two variables are just used to represent the cross-section dimensions of the data hence their descriptions are not meaningful. So our variables of interest are the real variables which are trade, the dependent variable and all other variables which are the explanatory variables.

Looking at the variable of interest, trade, we notice that there is more between variation which is from one country pair to the next, than within variation which is the variation of particular country pairs over time which tells us there is more variation in trade flows between one particular pair to the other, as compared to variation in the same pair over time. This pattern runs through the dataset since we observe that the between effects are greater than the within effects.

We notice that for Border, Language and Distance the within variation is zero, which means that the variables are time-invariant which intuitively makes sense because we cannot expect common borders, similarity in language or the distance between two countries to change over time.

3. ECONOMETRIC ESTIMATIONS

Given the general econometric model given above, I will estimate the pooled OLS, the random and the fixed effects specifications.

a) Pooled OLS

Even though we are choosing panel-based estimation over cross sectional estimation we might not be able to get rid of heterogeneity by using a pooled regression model since it does not deal with issues of heterogeneity which might lead to model misspecification and possibly affect inference. The pooled model is the regular OLS estimation which specifies constant coefficients. I run two versions of this model, one with the regular standard errors and the other with the robust standard errors. Below is the general form of the OLS:

$$Y_{it}=\alpha+\beta X_{it}+u_{it}$$

Where Yit is the dependent variable, Xit is the independent variable and α is the constant term and β's are coefficients of the explanatory variables.

The major problem with this model is the inability to distinguish between the different pairs of countries that constitute our variables of interest. This implies that by combining all the pairs of countries by pooling we deny the heterogeneity or individuality that may exist among them. So if heterogeneity differ across the individual units, the unobserved heterogeneity induces autocorrelation, therefore any latent effects left out of the model will carry across all periods. This model is oversimplified by assuming that the individual units are homogenous and this may result in autocorrelation in panel data analysis which is possibly why pooled regressions are not widely used in the literature.

In this regression, I will pool all the observations together and run the regression model, not recognizing the cross section and time series nature of the data. I am estimating this model because I would compare its results with that the Random Effect model and the Fixed Effects. Further, the results from these models will be used to conduct the Hausman Test later in the paper to examine if the Random Effect model is more appropriate than the Fixed Effects.

The Group Means Estimator (Between Effects)

The pooled regression model can be estimated using sample means. The group mean estimator may be useful when the explanatory variables are measured with error, the group means estimator will average it out, if the error is random OLS will not be consistent.

b) Individual-Specific Effects Models

Taking into account the limitations of the pooled OLS, an individual-specific effects model might be a preferred alternative since it can deal with the issue of heterogeneity. In this kind of model we assume that the unobserved heterogeneity across individuals are captured by αi.

This model has the form:

$$Y_{it}=X'_{it}\beta+(\alpha_i+\varepsilon_{it})$$

If the individual-specific effects are correlated with the explanatory variables we use the fixed-effect model or else we use the random effects model.

Fixed Effect Model

The fixed effect model allows for heterogeneity or individuality among the different pairs of countries by having individual intercept values. Although the intercept may vary across the pairs, it does not vary across time thus making it time invariant. So, in the general model above, the fixed effect model has the αi possibly correlated with $X'it$ where $X'it$ could be endogenous. Hence this model is able to consistently estimate β for the time varying $X'it$.

Since the cultural, historical and political factors among the different countries are difficult to observe and measure, allowing each pair of countries to have their own dummy variables may lead to the inclusion of country-pair effects. While the dummy variables may be correlated with both the bilateral trade and the explanatory variables, the country pair effects allows for separation according to the direction of trade.

However, the time invariant issue of fixed effects does not allow for estimating coefficients on time invariant variables such as language dummies, distance or common border even though incorporating the effect of these can be quite important in certain contexts. In addition to dealing with unobserved heterogeneous individual effects we also need to deal with the correlation of those effects with both time invariant and time varying explanatory variables, so that no potential bias arises which will possibly affect inference.

I will therefore examine whether these time invariant variables have an effect on our model by dropping them from my estimation of the fixed effects model and running the regression again.

Random Effect Model

If the unobserved individual heterogeneity is assumed to be uncorrelated with the included variables then we may use the random effects model which essentially specifies that the explanatory variables have a common mean value for the intercept. What that means is that the individual specific-effects are distributed independently of the explanatory variables, hence it should be included in the error term. So, in the general model above, the random effect model assumes that αi is purely random and that $X'it$ must be exogenous. Hence this model is able to correct standard errors for auto correlated clustered errors and conduct predictions.

4. PRESENTATION AND DISCUSSION OF RESULTS

Following economic theory, GDP should have a positive relationship with trade flows. GDP is supposed to have a positive impact on the amount of trade that flows into a country. RLF measures the difference in terms of relative factor endowments between two countries. The larger the difference, the higher the volume of inter-industry trade will be, the opposite is true. In fact all the other explanatory variables are expected to have a positive impact on trade with the exception of distance. The effect of transportation costs proxied by geographical distance between capital cities is negatively related with trade. The further the distance between two countries the less likely they are to trade, all else constant. EMU is a situation where both countries adopt the same currency for transactions and it is expected to have a positive impact on trade flows. A single currency will reduce the transaction costs of trade within member countries. The common language dummy, LAN has a value equal to one when both countries speak the same official language and is meant to capture similarity in cultural and historical backgrounds between trading countries. Another consideration is the impact of bilateral real exchange rates, RER, which is defined as the price of the foreign currency per the home currency unit and is meant to capture the relative price effects. A depreciation of the home currency relative to the foreign currency (an increase in RER) is expected to lead to more export and less import for the home country. The effect of real exchange rates on total trade flow will be positive if the export component of the total trade is significantly larger than the import component, the opposite is true.

Table 4.1. Estimates from the five regression outputs

```
. estimates tab ols ols_robust random_effect between fixed_effect,b(%7.2f)se(%7.2f)stats(N r2 r2_a rms
```

Variable	ols	ols_rob~t	random_~t	between	fixed_e~t
Gdp	1.54	1.54	1.88	1.54	1.95
	0.01	0.01	0.03	0.08	0.04
Sim	0.85	0.85	1.19	0.81	1.23
	0.02	0.02	0.05	0.10	0.06
Rlf	0.02	0.02	0.03	0.01	0.03
	0.01	0.01	0.01	0.06	0.01
Rer	0.09	0.09	0.08	0.09	0.08
	0.00	0.00	0.01	0.02	0.01
Cee	0.23	0.23	0.32	-0.03	0.32
	0.02	0.02	0.02	0.24	0.02
Emu	0.21	0.21	0.06	-1.60	0.03
	0.05	0.04	0.03	1.84	0.03
Dist	-0.68	-0.68	-0.53	-0.71	(omitted)
	0.02	0.02	0.12	0.14	
Bor	0.53	0.53	0.44	0.61	(omitted)
	0.03	0.03	0.19	0.20	
Lan	0.25	0.25	0.41	0.24	(omitted)
	0.03	0.03	0.19	0.20	
RERT	0.18	0.18	-0.08	(omitted)	-0.12
	0.02	0.02	0.02		0.02
_cons	-10.93	-10.93	-15.11	-9.71	-19.48
	0.24	0.26	0.96	1.55	0.40
N	3822	3822	3822	3822	3822
r2	0.91	0.91		0.91	0.90
r2_a	0.91	0.91		0.90	0.90
rmse	0.58	0.58	0.29	0.52	0.29

legend: b/se

.

From the table above, observe that the signs of the coefficients in all the five models are the almost all the same, which indicates the relationship between the variables are the same across models. The estimates in the table above have been rounded to two decimal places which makes the estimates look exactly the same but I have presented the individual regression outputs in Appendix 1, so that I can explain certain variations in the results. From the results displayed in the appendix, the coefficients of the Fixed Effects Model and the Random Effects model are similar for some variable and very close for others, whereas the coefficients of the Pooled OLS is much smaller if we look at the GDP and SIM coefficients, where as it is larger if we look at the EMU coefficient.

The distinction between the pooled and the individual specific effects are expected, since by pooling the regression assumes that there is no relationship between the time-invariant variables and the time-varying variables and also robs the model of its heterogeneity by not treating it as the correctly specified panel-form.

Looking further into the table, we notice that the coefficient on RLF is almost identical in the all the models. This means that the similarity between factors of endowment are exogenous to other variables. From Appendix 1 we find the probability value is less than 0.05 (in all the models estimated) which means all the coefficients of these models are not equal to zero (jointly significant).We also notice that all the explanatory variables are statistically significant variables in explaining the dependent variable, Trade. The expected positive association between most of the independent variables and Trade were realized as well the negative relationship between Trade and Dist (distance). All coefficients are statistically significant.

Hausman Test

The Hausman Test is a test that helps us to choose between the fixed effect and the random effects model. The Hausman test-statistics can be only calculated for the time varying explanatory variables. The test is given by:

$$H=(\widetilde{\beta RE}-\widetilde{\beta FE})'(V(\widetilde{\beta RE})-V(\widetilde{\beta FE}))(\widetilde{\beta RE}-\widetilde{\beta FE})$$

The degrees of freedom are equal to the number of regressors for the time-varying explanatory variables. If I find a statistically significant p-value, then I will reject the null and use the fixed-effects model, otherwise I will use the random-effects model.

The test follows the following hypothesis:

Fixed effect model is consistent under H_o and H_a

Random effect inconsistent under H_a, efficient under H_o

Results from the Hausman Test is presented in the table below:

Table 4.3 Hausman Test

```
. hausman con eff
```

	Coefficients (b) con	(B) eff	(b-B) Difference	sqrt(diag(V_b-V_B)) S.E.
Gdp	1.951331	1.880274	.0710573	.0163096
Sim	1.227669	1.18895	.0387189	.0304738
Rlf	.0313344	.0332146	-.0018801	.0009717
Rer	.0839729	.0839463	.0000266	.00391
Cee	.3165067	.3246798	-.0081732	.0018318
Emu	.0274301	.055788	-.028358	.005024
RERT	-.1165067	-.0758605	-.0406462	.0094687

```
                        b = consistent under Ho and Ha; obtained from xtreg
         B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(7) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =       29.49
                Prob>chi2 =      0.0001
                (V_b-V_B is not positive definite)
```

From the results above we find the probability value is less than 5% therefore we can reject the null hypothesis. This means that according to the Hausman Test, fixed-effects model is the most appropriate model to use in this analysis.

5. CONCLUSION

In this empirical project, I have used panel data from 91 country pairs who trade with each other and analyzed how the results could differ in different panel models. The data covers the period from 1960 to 2001. The results from my estimations indicate that the Random Effects Model are not very useful in reality because of the assumption that the unobserved individual heterogeneity is uncorrelated with the explanatory variables. This assumption is often violated. The Hausman test also concluded that the Fixed-Effects model is the right way to go. An extension of this project could be done in the future by analyzing Chang and Wall's approach. However, we must take notice of the fact that unless certain assumptions are met, their results should be taken with skepticism. Given this limitation, their model may however be a better fit as compared to the random-effects model, since the Hausman test has indicated that the unobserved time invariant explanatory variables are indeed correlated with the time varying explanatory variables. In the interest of economic policy, it would be interesting to investigate the role of explanatory variables such as RLF and EMU in the gravity models of international trade over different time periods as well as the important role a factor like institutional quality could impact the inflow and returns from trade.

REFERENCES

1. Anderson, James E. (1979) "A Theoretical Foundation for the Gravity Equation", American Economic Review, 69(1), pp. 106-116

2. Cheng I, Wall HJ. 2002. Controlling heterogeneity in gravity models of trade. Working Paper 1999-010C: Federal Reserve Bank of St. Louis.

3. Laura Serlenga and Yongcheol Shin (2007), Gravity Models of Intra-EU Trade: Application of the CCEP-HT Estimation in Heterogeneous, Journal of Applied Econometrics, Vol. 22, No. 2, pp. 361-381

APPENDIX

APPENDIX 1: RESULTS FROM THE FIVE REGRESSION ESTIMATIONS

. reg Trade Gdp Sim Rlf Rer Cee Emu Dist Bor Lan RERT

Source	SS	df	MS
Model	12498.0818	10	1249.80818
Residual	1286.28189	3811	.337518208
Total	13784.3637	3821	3.60752778

Number of obs	= 3822
F(10, 3811)	= 3702.94
Prob > F	= 0.0000
R-squared	= 0.9067
Adj R-squared	= 0.9064
Root MSE	= .58096

Trade	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Gdp	1.543887	.0130767	118.06	0.000	1.518249	1.569525
Sim	.8494035	.0171976	49.39	0.000	.8156861	.883121
Rlf	.0249879	.0084132	2.97	0.003	.008493	.0414828
Rer	.0906828	.0039186	23.14	0.000	.0830001	.0983655
Cee	.233327	.0244463	9.54	0.000	.1853978	.2812562
Emu	.2137038	.0507735	4.21	0.000	.114158	.3132496
Dist	-.6828109	.0226465	-30.15	0.000	-.7272112	-.6384105
Bor	.5311422	.0337685	15.73	0.000	.4649363	.5973482
Lan	.2487985	.0340738	7.30	0.000	.1819939	.3156031
RERT	.1813633	.0209483	8.66	0.000	.1402924	.2224342
_cons	-10.93177	.244689	-44.68	0.000	-11.4115	-10.45203

. reg Trade Gdp Sim Rlf Rer Cee Emu Dist Bor Lan RERT, robust

Linear regression

Number of obs	= 3822
F(10, 3811)	= 3366.38
Prob > F	= 0.0000
R-squared	= 0.9067
Root MSE	= .58096

Trade	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
Gdp	1.543887	.0131458	117.44	0.000	1.518113	1.56966
Sim	.8494035	.0167181	50.81	0.000	.8166262	.8821808
Rlf	.0249879	.0079525	3.14	0.002	.0093963	.0405794
Rer	.0906828	.0035705	25.40	0.000	.0836824	.0976831
Cee	.233327	.022626	10.31	0.000	.1889668	.2776872
Emu	.2137038	.0412096	5.19	0.000	.1329088	.2944988
Dist	-.6828109	.0248209	-27.51	0.000	-.7314744	-.6341473
Bor	.5311422	.0331714	16.01	0.000	.4661068	.5961777
Lan	.2487985	.0333143	7.47	0.000	.183483	.3141141
RERT	.1813633	.0211936	8.56	0.000	.1398115	.2229152
_cons	-10.93177	.2584233	-42.30	0.000	-11.43843	-10.42511

```
. xtreg Trade Gdp Sim Rlf Rer Cee Emu Dist Bor Lan RERT, re

Random-effects GLS regression                   Number of obs      =      3822
Group variable: countrypair                     Number of groups   =        91

R-sq:  within  = 0.8982                         Obs per group: min =        42
       between = 0.8974                                        avg =      42.0
       overall = 0.8952                                        max =        42

                                                Wald chi2(10)      =  33466.38
corr(u_i, X)   = 0 (assumed)                    Prob > chi2        =    0.0000
```

Trade	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Gdp	1.880274	.0317116	59.29	0.000	1.81812	1.942427
Sim	1.18895	.0477741	24.89	0.000	1.095314	1.282585
Rlf	.0332146	.0078101	4.25	0.000	.0179071	.0485221
Rer	.0839463	.0091093	9.22	0.000	.0660925	.1018001
Cee	.3246798	.0159505	20.36	0.000	.2934174	.3559423
Emu	.055788	.0290003	1.92	0.054	-.0010516	.1126276
Dist	-.5303254	.1184368	-4.48	0.000	-.7624573	-.2981935
Bor	.4366102	.1907719	2.29	0.022	.0627041	.8105164
Lan	.4099015	.1852032	2.21	0.027	.0469099	.772893
RERT	-.0758605	.0230384	-3.29	0.001	-.1210149	-.0307061
_cons	-15.11136	.9592535	-15.75	0.000	-16.99146	-13.23125
sigma_u	.51584354					
sigma_e	.29210071					
rho	.75720375	(fraction of variance due to u_i)				

```
. xtreg Trade Gdp Sim Rlf Rer Cee Emu Dist Bor Lan RERT, be
note: RERT omitted because of collinearity

Between regression (regression on group means)  Number of obs      =      3822
Group variable: countrypair                     Number of groups   =        91

R-sq:  within  = 0.6656                         Obs per group: min =        42
       between = 0.9144                                        avg =      42.0
       overall = 0.8544                                        max =        42

                                                F(9,81)            =     96.18
sd(u_i + avg(e_i.))=  .5178089                  Prob > F           =    0.0000
```

Trade	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Gdp	1.540617	.0801364	19.22	0.000	1.381171	1.700063
Sim	.8056846	.1041487	7.74	0.000	.5984613	1.012908
Rlf	.0110735	.0604072	0.18	0.855	-.1091178	.1312648
Rer	.0866231	.0235	3.69	0.000	.0398654	.1333807
Cee	-.030312	.2441104	-0.12	0.901	-.5160151	.455391
Emu	-1.596042	1.841584	-0.87	0.389	-5.260216	2.068132
Dist	-.710384	.1355111	-5.24	0.000	-.9800086	-.4407595
Bor	.6085459	.2012652	3.02	0.003	.2080912	1.009001
Lan	.2359094	.2013823	1.17	0.245	-.1647782	.6365971
RERT	0	(omitted)				
_cons	-9.706384	1.55222	-6.25	0.000	-12.79481	-6.617952

```
. xtreg Trade Gdp Sim Rlf Rer Cee Emu Dist Bor Lan RERT, fe
note: Dist omitted because of collinearity
note: Bor omitted because of collinearity
note: Lan omitted because of collinearity

Fixed-effects (within) regression               Number of obs      =      3822
Group variable: countrypair                     Number of groups   =        91

R-sq:  within  = 0.8983                         Obs per group: min =        42
       between = 0.7987                                        avg =      42.0
       overall = 0.8204                                        max =        42

                                                F(7,3724)          =   4699.29
corr(u_i, Xb)  = -0.1343                        Prob > F           =    0.0000

------------------------------------------------------------------------------
       Trade |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         Gdp |   1.951331   .0356599    54.72   0.000     1.881416    2.021246
         Sim |   1.227669   .0566658    21.67   0.000      1.11657    1.338768
         Rlf |   .0313344   .0078703     3.98   0.000     .0159039     .046765
         Rer |   .0839729   .0099129     8.47   0.000     .0645375    .1034082
         Cee |   .3165067   .0160553    19.71   0.000     .2850285    .3479848
         Emu |   .0274301   .0294323     0.93   0.351    -.0302749     .085135
        Dist |          0  (omitted)
         Bor |          0  (omitted)
         Lan |          0  (omitted)
        RERT |  -.1165067   .0249083    -4.68   0.000    -.1653419   -.0676714
       _cons |  -19.47683   .3951507   -49.29   0.000    -20.25156   -18.70209
-------------+----------------------------------------------------------------
     sigma_u |  .76239634
     sigma_e |  .29210071
         rho |  .87199739   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(90, 3724) =   232.87            Prob > F = 0.0000

. estimates store fixed_effect
```