***GRAVITY MODEL for Turkey-Germany***

The gravity theory suggests that a bilateral trade flow can be appropriately represented by taking account of the GDP of the trading countries and the distance between these two countries.

In fact, the gravity theory originates from Newtonian physics that imply the force of gravity between two matters can be determined by the size of these two matters and their distance.

Now, let’s test the gravity equation to see whether it accurately represents the trade between Germany and Turkey.

The gravity equation:

: All the trade (exports and imports of one country towards the other) between country i and j.

: Estimated coefficient of the “gravity term” or the “gravity ratio” for country i and j at time t.

: Gross Domestic Product of country i at time t.

: Distance between countries i and j. (by plane/air)

: the estimated exponent coefficient of the distance variable that will tell us how much explanatory power bilateral distances have in quantifying trade flows through the gravity model.

In other words, TradeFlowij will be regressed on the gravity term, along with an intercept term in the regression. Note that this is a nonlinear model due to the multiplication of two independent variables (GDPs of the two countries). Further, since is a denominator, a nonlinear estimator must be used to estimate the coefficients on the right-hand side (RHS).

Even though I began with using Maximum Likelihood Estimation (MLE), Python warned me about couple of numerical issues, and I was not able to calculate Mean Squared Error (MSE). Thus, I switched to **Nonlinear Least Squares** to estimate the parameters.

As the magnitude of the variables we are working with are large, I rescaled them by dividing each by their maximum value so that their range would be in (0, 1].

Estimated parameters:

Estimated beta: 0.6076928959513533

Estimated n: 1.0

Estimated constant: 0.39608622960928297

**Estimated beta (β)**: Represents the coefficient associated with the interactive variable that is the product of GDPs of Turkey and Germany at a given year. In this case, the estimated beta value of approximately 0.608 suggests that a one-unit increase in the product of GDPs is associated with an increase in bilateral trade flows between Turkey and Germany. However, since the estimated beta is close to zero and not statistically significant (see below), it implies that the product of GDPs may not have a significant impact on trade flows between the two countries, at least in the context of our model.

**Estimated n**: The exponent coefficient determines the degree to which distance influences trade flows; a value of 1 insinuates that trade flows decrease in proportion to the distance between countries. That is to say, the effect of distance on trade flows follows a linear relationship, where doubling the distance between two countries results in halving the expected trade flows between them.

**Estimated constant**: The intercept of the regression equation in the gravity model. It captures the baseline level of bilateral trade between Turkey and Germany when all other variables (product of annual GDPs and distance) are set to zero. This baseline level of trade may reflect factors that are not included in the model such as historical trade relationships, cultural ties, and geographical proximity between the two countries. That is, all the “omitted variables” will be here, and the downside is that these variables will be represented in the intercept term in a time invariant manner, as the intercept term is a constant thus is time invariant.

**Other Indicators to Interpret:**

P-values for Variables:

Product of GDPs: 0.8973610826324747

Distance: 0.8644004376758858

R-squared (R^2): 0.8829831023533649

Mean Squared Error (MSE): 0.0006897210443292836

A graph of a line

Description automatically generated with medium confidence

High R^2, low MSE, p-values far away from the conventional significance level (0.05) imply generally overfitting.

Overfitting may occur when a model captures noise/random fluctuations in the data rather than the underlying true relationship between the independent variables and the dependent variable. Particularly for our gravity model for modelling bilateral trade flows, overfitting is very likely to occur for simple models, since there are many variables that must be included to correctly explain trade flows between two countries. Obviously, the sizes of the two economies (GDP) and their distance from each other are not the mere influencers of their trade relationship.

In other words, including too few variables (two variables in our case) may lead to overreliance on these variables, directly overlooking the importance of other variables.

As mentioned partially above, the high R-squared value obtained may suggest our model explains a large portion of the variance in trade flows. However, in turn, the insignificance of the included variables (as indicated by high p-values) suggests that the model may be merely fitting noise rather than true underlying patterns.

Thus, the possibility of omitted variables, such as trade policies, exchange rates, cultural factors, and geopolitical considerations, may further complicate our model's ability to accurately capture the determinants of bilateral trade relations.

Therefore, while the model may provide a good fit to the observed data, it may not generalize well to other bilateral trade relations and/or accurately predict trade flows under different conditions (that may be captured by including other variables).

***BORDER EFFECTS***

Border effects apply for countries that share borders. Germany and Turkey does not share borders. However, we can add trade-restrictive independent variables such as tariff and non-tariff measures to see whether they are able to explain Germany-Turkey bilateral trade flows combined with the gravity term.

On the RHS, I will add the Measure of Aggregate Trade Restrictions (MATR) index that can be generated in a very self-tailored way from IMF’s AREAER database. I will compute MATR for Germany and Turkey and then interact these indices, as we did in GDP. To elaborate on MATR, it takes the unweighted sum of many binary variables each of which imply some type of trade-restrictiveness.

The main categories are:

*Exchange measures imposed for security reasons*

*Prescription of currency requirements*

*Payments arrangements*

*Administration of control*

*Payments arrears*

*Controls on trade in gold (coins and/or bullion)*

*Controls on exports and imports of banknotes*

*Foreign exchange accounts permitted*

*Accounts in domestic currency convertible into foreign currency*

*Documentation requirements for release of foreign exchange for imports*

*Import licenses and other nontariff measures*

*Import taxes and/or tariffs*

*State import monopoly*

*Repatriation requirements*

*Financing requirements*

*Documentation requirements*

*Export licenses*

*Export taxes*

*Controls on these transfers*

*Repatriation requirements*

*Restrictions on use of funds*

Further, there are subcategories of these categories for more comprehensiveness. You can see the whole list in the excel file “MATR.xls”.

The AREAER database stores all these subcategories for a given country and a year in categorical format, i.e., “yes”, “no”, “n.r.” (not regulated). The value “yes” implies the corresponding restriction/measure significantly exists, “no” implies the IMF found no measure, whereas “n.r.” implies the country itself informed the IMF that no such restriction/measure/regulation exists throughout a given year in a given country.

Dummy coding these values as “yes” equivalent to 1 and “no” and “n.r.” equivalent to 0 according to their trade-restriction implications is an appropriate way to convert them into usable form for quantitative data analysis. For previous implementations of MATR index, see Estefania-Flores et. al (2022) and Campos et. al (2023) (both are the works of the IMF staff).

A graph with orange and blue lines

Description automatically generated

***RESULTS FOR NEW MODEL***

Estimated beta1: 1.15406745732168e-05

Estimated n: 2.7933607126181954

Estimated constant: 27192357824.325413

Estimated beta2: -4120659.767225214 (coefficient of the product of MATR indices)

P-values for Parameters (with 3 decimals):

Estimated beta1: 0.000

Estimated n: 0.000

Estimated constant: 0.000

Estimated beta2: 0.000

R-squared (R^2): 0.7256402921164997

Mean Squared Error (MSE): 6.49841164470292e+17

A graph with blue and orange lines

Description automatically generated

The results suggest the updated model with the MATR index included has lower R-squared compared to the previous model which does not include the MATR index as an independent variable. However, the p-values for all parameters are significant, suggesting that the variables in the model are statistically significant predictors of bilateral trade between Turkey and Germany.

**Observations:**

1. Estimated beta1 (the coefficient for the GDP product) and n (the distance decay parameter) are both statistically significant with p-values close to zero, indicating that they have a significant impact on bilateral trade flows.

2. The estimated constant term is also statistically significant, suggesting that there is a baseline level of trade between Turkey and Germany when the above-mentioned variables are taken into account.

3. The coefficient of the product of MATR indices (beta2) is also statistically significant with a p-value close to zero. This indicates that trade restrictiveness -as measured by the MATR indices of both countries- may have a strong impact on bilateral trade flows.

4. The R-squared value is moderately lower compared to the previous model, albeit having statistically significant variables. In line with our argument of overfitting in the first model, this suggests that although the included variables explain a good amount of portion of the variance in trade flows, there may still be other factors not captured by the model. Note that we alleviated the overfitting issue since now we have significant variables and lower R^2.

5. The mean squared error (MSE) is quite high, indicating that there is still a considerable amount of unexplained variance in the model. This is natural for trying to explain such a macro variable like bilateral trade flows with only couple of variables.

As mentioned before, there are many other factors that influence trade relations. Further, MATR is an unweighted sum of binary variables, the format of which has been used in trade analysis previously but also questioned by many authors due to its noncontinuous, general nature (i.e., binary nature).

All in all, border effects (in this case, “international trade effects” rather than border) are quite significant when it comes to explaining bilateral trade flows between Turkey and Germany for the year interval 2013-2021. The new model lowered the overfitting of the pure gravity model but also suggested that there are other variables that we should account for. Further, using continuous and/or non-aggregate variables so as to not lose information (as may happen in MATR) may further improve the model. Nevertheless, since some trade measures are very hard to quantify (especially non-tariff measures), our model with the addition of MATR still does a pretty good job despite its simplicity.

**REFERENCES**

Campos, R. G., Estefania-Flores, J., Furceri, D., & Timini, J. (2023). Geopolitical fragmentation and trade. *Journal of Comparative Economics*, *51*(4), pp. 1289-1315. https://doi.org/10.1016/j.jce.2023.06.008

Estefania-Flores, J., Furceri, D., Hannan, S. A., Ostry, J. D., & Rose, A. K. (2022). *Measurement of aggregate trade restrictions and their economic effects*. International Monetary Fund WP/22/1.