CODE FILES

Topic: International Analysis Using EGM Neural Networks and Random Forests

Authors:

**Prakhar Pradhan**

**Aditya Gupta**

**Raj Aryan**

Data Consists of Top 18 Countries from 1948-2019 totaling 23328 and 25 columns.

**Saral Verma**

**Atul Umak**

**Chirag Sharma**

**Kaushal Chaudhary**

**Sumit Singh**

**Sandeep Gautam**

**Anurag Yadav**

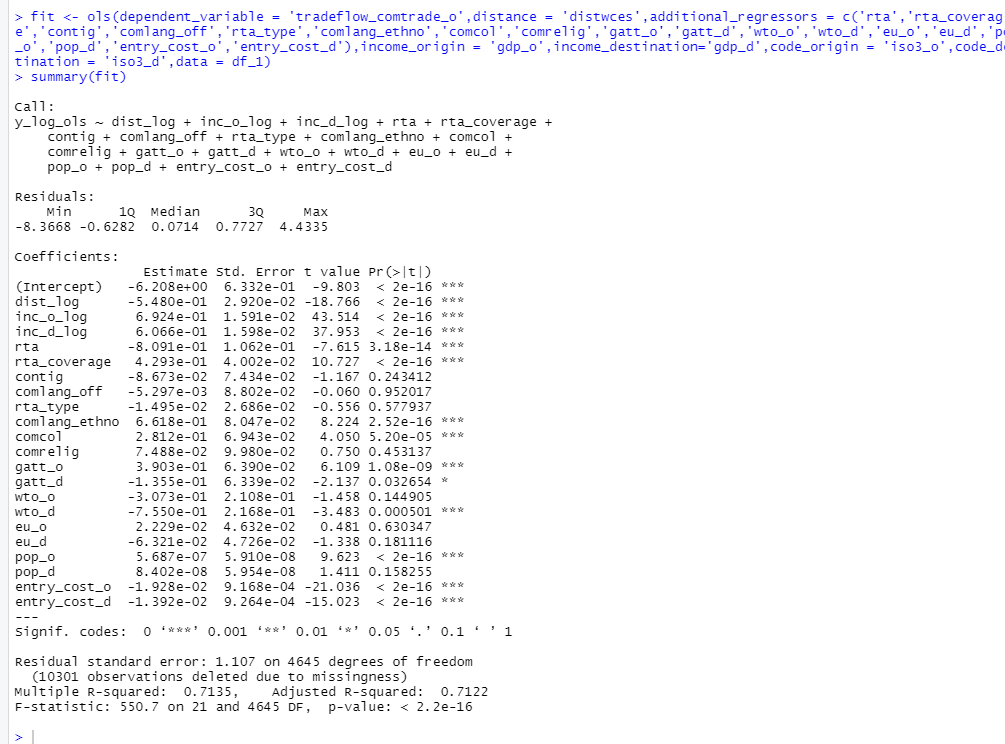
**Variable Description**

|  |  |
| --- | --- |
| Variables | Description |
| distwces | Using CES formulating Population-weighted distance between populated cities. |
| Gdp\_o / gdp\_d | GDP (nominal) of the origin or destination country (current thousands US$). |
| Iso3 | Country alphabetical and Numerical Code |
| Contig | Equal to 1 if countries shared common Border. |
| comcol | Dummy Column equal to 1 if countries share a joint colonizer post-1945. |
| comrelig | Religious proximity index |
| Rta | Equal to 1 if the country has an RTA agreement. |
| Rta\_type | Indicates the type of RTA (customs union, for instance) |
| Rta\_coverage | Equal to 1 if RTA covers both Goods and Services. |
| Gatt\_O / Gatt\_d | Dummy equal to1 if origin or destination country currently is a GATT member. |
| Wto\_o / Wto\_d | Dummy equal to 1 if origin or destination country currently is a WTO member. |
| Eu\_o / Eu\_d | Dummy equal to 1 if origin or destination country currently is an EU member. |
| Comlang\_off | Dummy equal to 1 if countries share a familiar official or primary language. |
| Comlang\_ethno | Equal to 1 if at least 9% of the population share the common language. |
| entry\_cost\_o / entry\_cost\_d | Cost of business (% of GNI per capita) of origin or destination country. |
| Tradeflow\_comtrade\_o | Trade flow as reported by the exporter (in thousands current US$) (source: Comtrade) |

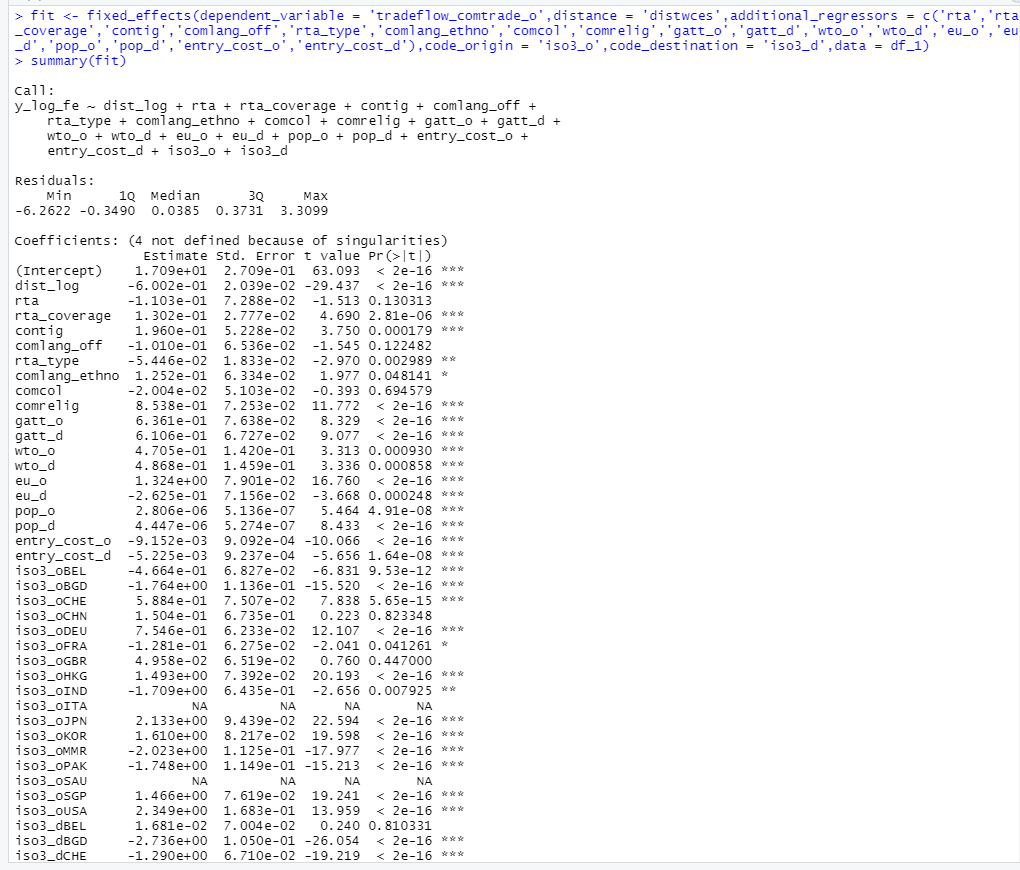
**For OLS, Fixed Effects, PPML, GPML We have used Special Package ‘Gravity’ available in R. Repository is CRAN. Using This Package, we have implemented the above models keeping the Conditions in mind for example, Distance is logged, and zero Trade Values are interpolated.**

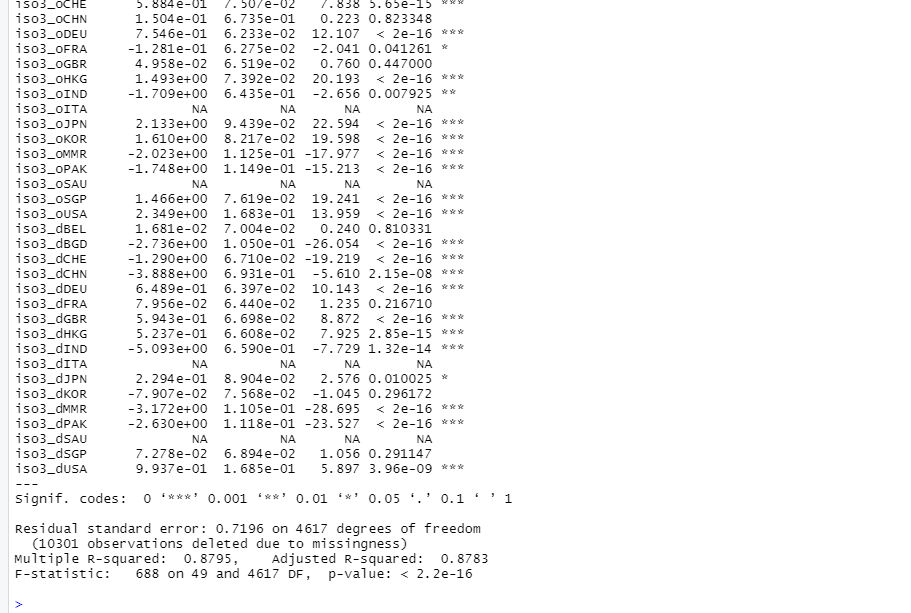
**Attached are the R codes:**

1. **OLS**

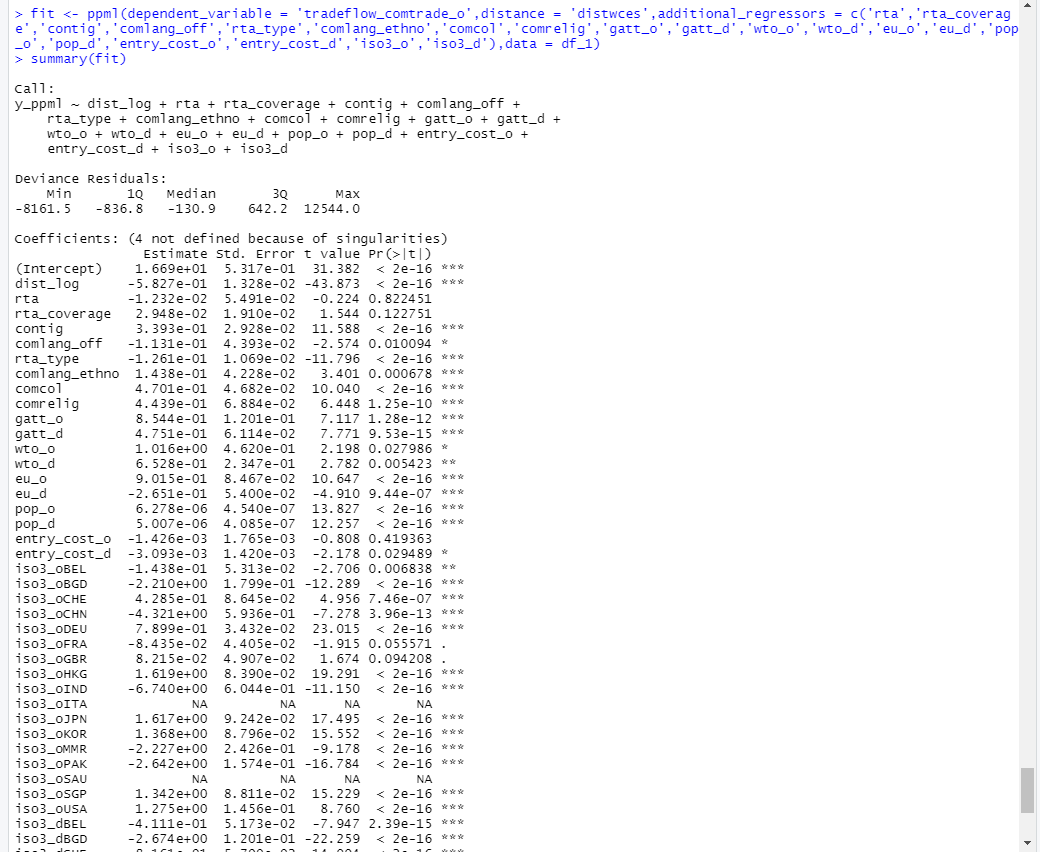


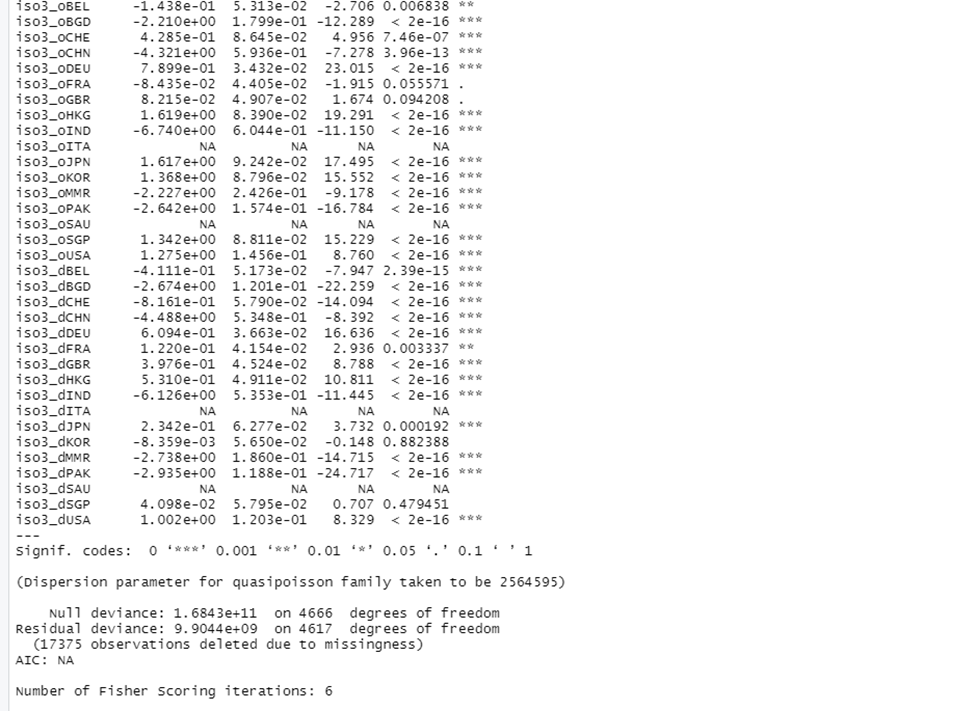
1. Fixed Effects



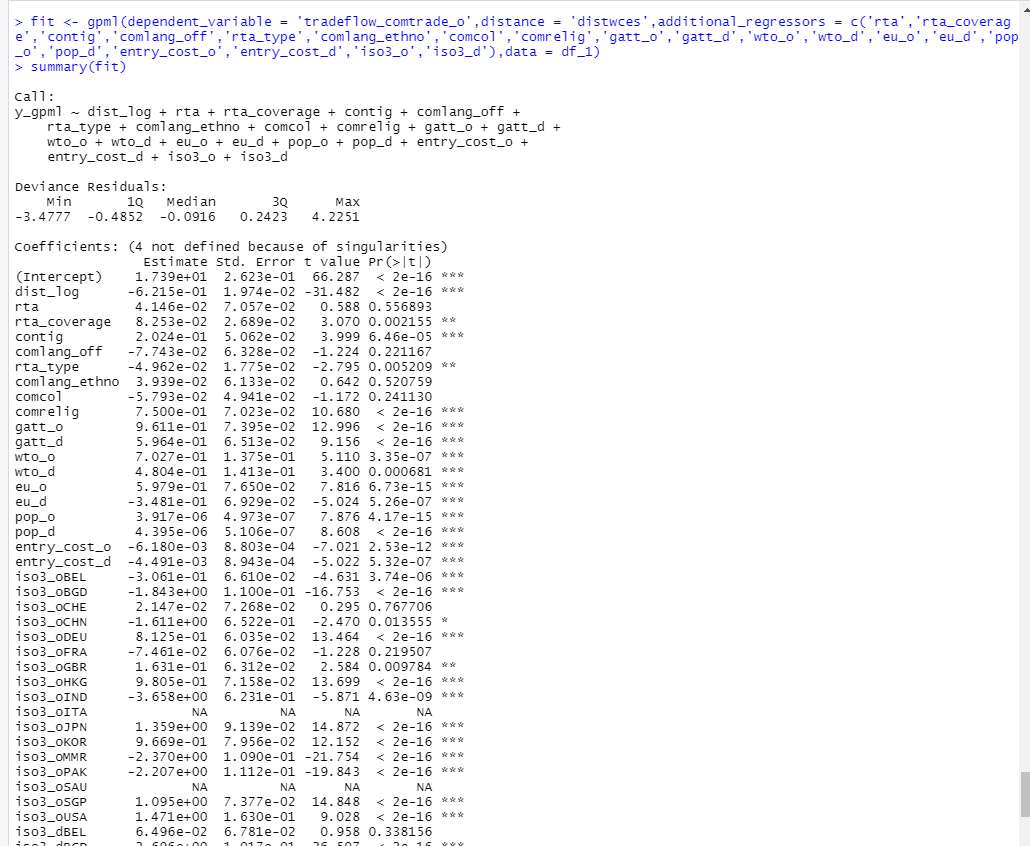


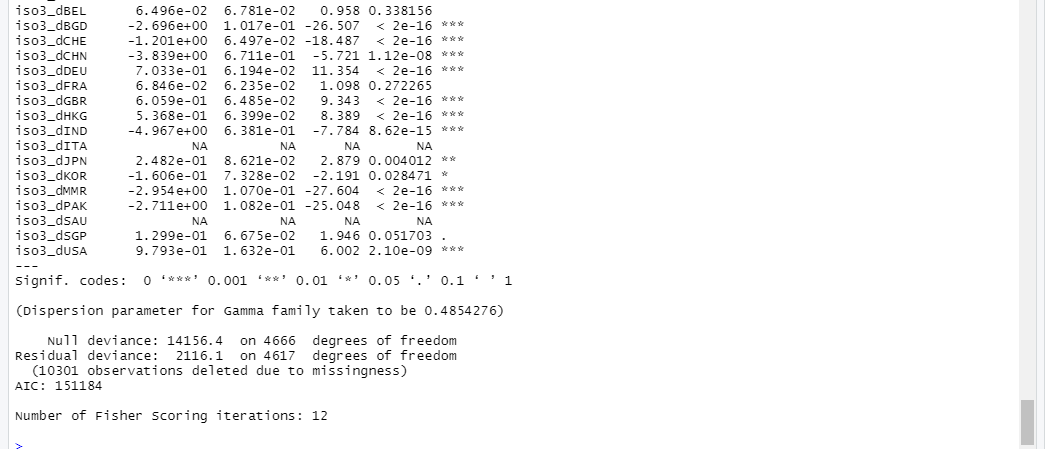
1. PPML





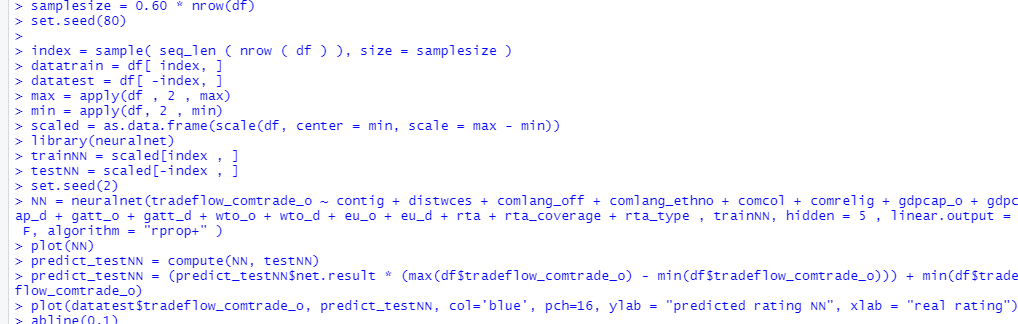
1. GPML

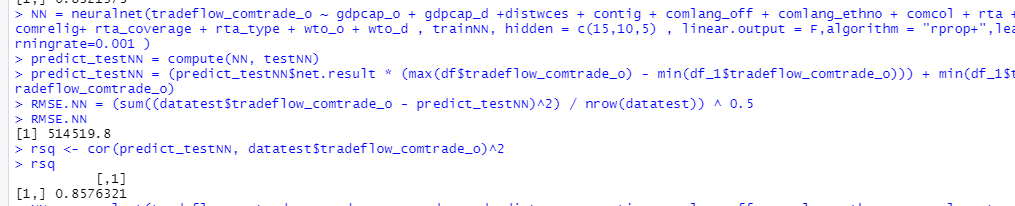




**For Neural Network, we have Used a special package called ‘Neural Net’ available in R. Following Steps Implemented:**

* **Divided Sample in Test and Train**
* **Scaled it.**
* **Divided in TrainNN and TestNN.**
* **Used Neural Net Function to train the model by adjusting Hyper Tuning Parameters like, number of hidden layers and learning rate. We used “Rprop” algorithm for loss function.**
* **Fitted the instance in TestNN and calculated the R2.**
* **Finally Plotted the graph of ‘Actual’ and ‘Predicted’.**





**For Random Forest We Have Implemented in Python using the ‘Sckit-learn’. Following Steps Were Taken:**

* **Mounted the drive and Loaded the csv File in Colab using Pandas Library in Python.**
* **Imported RandomForestRegressor from Sklearn.ensemble**
* **Formed X and Y by dropping the ‘Tradeflow\_comtrade\_o’ column.**
* **Using the Train Test Split Divided the Data into X\_train, Y\_train, X\_test, Y\_test.**
* **Fitted the Data and Predicted the X\_test.**
* **Used the Evaluation Metric R2 and RMSE.**

