Stochastic Frontier Analysis of India’s Trade with Gulf Cooperation Council (GCC): Assessing Trade, Economic Integration and Export Potential

(Term Paper for the course Econometrics-2 (ECO342A) at IIT Kanpur, Spring 2024)

Research Group -

- Aagam Jain (210004)
- Harsh Verma (210415)
- Mihir Tomar (210607)
- Sampada Kalavakunta (210914)
- Utkarsh Gupta (211129)

Submitted to - Dr. Somesh K. Mathur
Indian Institute of Technology, Kanpur

Abstract and Introduction

This paper seeks to closely analyse India’s trade relations with the Gulf Cooperation Council (GCC), which comprises 6 nations - Kuwait, Bahrain, Oman, Qatar, Saudi Arabia and United Arab Emirates. It also explores India’s export potential to the GCC with which a Free Trade Agreement (FTA) was signed on 25th August 2004, and which is currently under the third round of negotiation (with the two previous rounds of negotiations being held in 2006 and 2008).

For India, the GCC as a whole has enormous geographic and economic significance. The Gulf nations make up India's "immediate" neighbourhood, with the Arabian Sea serving as the only physical barrier. The GCC has historically had strong trading ties with India, but those ties have recently been seeing new heights as the GCC has emerged as a major trading partner for India. It also has humongous potential to be India's investment partner in the near future. Another aspect of this relationship is that India's energy needs are greatly dependent on the GCC's large oil and gas reserves. Labour migration is also a crucial link between these two sides, with the Gulf countries being host to a very large Indian diaspora, approximately 8.5 million in number in 2021. As a result, the Indian expatriate community in the Gulf region is the source of huge amounts of remittances for India, leading to an inflow of large amounts of foreign exchange into India.

Considering the uniqueness and the importance of India's relation with the GCC, we'll add a few modifications to the stochastic frontier gravity model (SFGM) discussed by Ebaidalla M. Ebaidalla & Mohammed Elhaj Mustafa Ali (2023) to account for the finer and unique details in India-GCC trading relations, making our model more robust. The modifications would take the form of a few dummy variables accounting for important factors like diaspora, trade affinity, common colonisers, etc. These nuances are discussed in detail later. The "augmented" stochastic frontier gravity model would then be used to analyse India's trade relation with GCC, look at the export flows, calculate technical efficiency, etc. The coefficients thus obtained could be incorporated to predict India's export potential to the GCC, and also suggest steps to further boost India-GCC trading relation, giving us policy suggestions.

Literature Review and the Research Gap

The gravity model of trade has widely been used to analyse the determinants of bilateral trade flows. It was first proposed by Jan Tinbergen in his 1962 Nobel Prize-winning paper, "Shaping the World Economy: Suggestions for an International Economic Policy." However, the use of a gravity-like equation to explain bilateral trade flows was first suggested by economist Walter Isard in his 1954 book "Location and Space-Economy." Later, Pöyhönen (1963) applied a gravity-like equation to explain bilateral trade flows in the context of the European Free Trade Association (EFTA). Since then, the gravity model has been widely used in empirical studies of international trade.

The gravity model has evolved over the years, with researchers including additional variables to improve its accuracy. One such variable is the bilateral exchange rate, which reflects the relative prices of goods between two countries.

Anderson and van Wincoop (2003) introduced the idea of adding the exchange rate to the gravity model, showing that it can significantly improve the model's explanatory power. Anderson and van Wincoop (2003) also introduced the concept of the "border puzzle" which refers to the discrepancy between the predicted and observed trade flows across borders.

They further developed the gravity model by introducing a variable called the "multilateral resistance," which measures the overall trade barriers faced by a country. The authors argued that a country's trade flows are influenced not only by its direct bilateral relations but also by the trade barriers it faces with other countries. The multilateral resistance variable captures this effect and has been shown to significantly improve the model's explanatory power.

Baier and Bergstrand (2007, 2009) made significant contributions to the analysis of the impact of FTAs on trade flows. They used the gravity model with the multilateral resistance term to analyse the impact of FTAs on trade flows between countries.
Overall, the gravity model has proven to be a useful tool for analysing the determinants of bilateral trade flows. The model has evolved over the years, with researchers incorporating additional variables to improve its accuracy.

Now, turning our attention to our specific case of India-GCC trading ties. Since both GCC and India have just recently started seeing their immense economic potential materialise, it surely has generated some interest in the economists. Yet, as per our knowledge, there exist only a few studies analysing India-GCC trading ties. A few major and unique ones have been listed below.

Samir Pradhan (2009) explored India's export potential to the six-member countries of Gulf Cooperation Council (GCC) using a simple augmented gravity model, and used Ordinary Least Squares (OLS) technique with panel data (1994-2004) to estimate the model. Apart from all other basic explanatory and dummy variables used in the gravity model, he used "trading affinity" a new dummy variable to account for India's historically strong trade ties with gulf countries. Results depicted that India had the highest export potential with Oman, followed by Qatar, Bahrain, and Kuwait. Further results also showed that India overtrades with UAE and Saudi Arabia.

Then, Imran Alam and Shahid Ahmed (2018) made some improvements to the gravity model being used for analysing the trade between India and GCC, by introducing a few more relevant dummy variables like "diaspora", etc. They investigated determinants of India-GCC trade flow with the help of panel gravity model for the period 2001 to 2015. The result of their study showed that India has significant trade potential with all six GCC countries, with the highest trade potential with Kuwait, followed by Bahrain, Oman, UAE, Qatar and Saudi Arabia. They also concluded that the dummy variable "Diaspora" played a very positive role, hence urging the Indian government to take good care of their workers in the Gulf countries.

Given the literature review, our study seeks to provide a more comprehensive picture on the issue and fill in the literature gap regarding model, methodology and variable choices. We aim to use the stochastic frontier gravity model (SFGM) developed by Ebaidalla M. Ebaidalla & Mohammed Elhaj Mustafa Ali (2023), albeit with a few modifications (in the form of a few dummy variables, some of which have previously been used by researchers but only with traditional gravity models) to analyse India-GCC trade, and to evaluate trade determinants and trade potentials. The estimates from our model would then be used for assessing the export potential of India, also analysing the effects of FTA on India's trade with GCC. Overall it aims to be more robust (accounting for important variables like MTR terms, etc.) and use the recently developed modelling techniques (SFGM) for the gravity model, along with innovations like "diaspora" dummy variables of the past researchers, and also try to predict more things, which can be useful for policy making.

Objective

The objective of this research proposal is to evaluate the potential for trade integration between the GCC nations and India, and identify the factors affecting trade flows between them using the stochastic frontier gravity model (albeit with a few modifications) along with the technical efficiency. We will also analyse the impact of India-GCC FTA on India's trade with GCC countries. At last, to cap off our analysis, we'll calculate the export potential of India to the various GCC members, by calculating the technical efficiency, and give some policy recommendations.

Model and Methodology

Our paper would be based on an "augmented" (modified) stochastic frontier gravity model of trade. Gravity model of trade is usually used to find out the determinants which influence bilateral trade between two regions. The model will be estimated using a panel data framework that includes data on bilateral trade flows, economic size, distance, similar coloniser, diaspora, and trade agreements between the GCC countries and India. The stochastic frontier gravity model also considers inefficiencies in the trade process, such as trade barriers etc.

First let us understand Viner's concepts of TC (Trade creation) and TD (Trade Diversion) associated with a trade agreement. In simple words, when a trade agreement is formed, participating countries agree to reduce or eliminate trade barriers (tariffs, quotas etc). This can lead to creation of new trade flows among member countries and substitution of lower cost, more efficient producers for higher cost, less efficient producers and thus is beneficial for the countries. This is called the trade creation effect. Whereas, Trade Diversion refers to the redirection of trade from a lower cost, more efficient producer to a higher cost, less efficient producer as a result of trade agreement. Thus TD involves shifting of trade away from one country toward one's free trade partner and can be detrimental to countries in terms of national welfare. These are crucial concepts in developing a conceptual framework for evaluating trade implications of a trade agreement.

To model the trade creation effect, we can estimate the change in India's trade flows with GCC countries after the implementation of FTA, holding other factors constant. This can be done by comparing India's actual trade flows with GCC countries after the implementation of the FTA with the trade flows that would have occurred if the FTA had not been implemented. If India's actual trade flows with GCC countries are higher than the counterfactual trade flows, it suggests that the FTA has led to trade creation. This can be judged by plotting the aggregate trade of India with the GCC countries.

According to the standard gravity model of international trade and commerce, the volume of trade between two nations is proportional to their economic mass and an indicator of their relative trade frictions. The gravity model has been the workhorse model of international trade for more than 50 years, perhaps due to its intuitive appeal. In this generalised model, trade is positively correlated with the economic size and population of the countries and negatively correlated with the distance between the countries involved. It derives from Newtonian model of gravitation as in the simplest gravity model, the bilateral trade flows between two countries are assumed to be directly proportional to the product of their economic sizes (represented by their respective GDPs) and inversely proportional to the distance between them.

$$F_{ij} = G \cdot \frac{M_i M_j}{D_{ij}}$$

In this formula, G represents a constant, F represents trade flow, D represents distance, and M represents the economic aspects of the countries being measured. Using logarithms, the equation can be transformed into a linear form for econometric studies. Baier and Bergstrand (2007) and Soloaga and Winters (2001) enhanced the classic gravity model by introducing a set of dummy variables that capture the effect of trade policies of FTA membership nations. It is important to note that different types of dummy variables have been incorporated into the gravity models to account for unique aspects, such as geographical, cultural, and institutional factors, that can facilitate or impede bilateral flows of goods. To address the inherent bias of the standard gravity model of trade and to estimate potential trade flows, this study uses the stochastic production frontier approach presented by Kalirajan (1999) to overcome the limitations of the conventional gravity model. The SFGM measures the trade frontier as the maximum possible level of trade for a given bilateral trading pair, which is influenced by a random error term that can be positive or negative, allowing the stochastic frontier trade level to fluctuate relative to the deterministic portion of the gravity equation. The stochastic frontier gravity model incorporates the gravity model, the stochastic frontier technique, and a non-negative error factor. In particular, the non-negative error term refers to inefficiencies "behind the border" in the exporting nation that hinder it from reaching its trading frontier. On the other hand, the random word encompasses all other disturbances, including "beyond the border." We can use the gamma coefficient to understand the variation in total trade due to 'behind the border' effects.

$$\gamma = \left[\frac{\sum_t \sigma_{ut}^2}{\sum_t \sigma_{ut}^2 + \sigma_{vt}^2} \right]$$

square(σ{ut}) is the variance of the one-sided error term at time t, whereas square(σ{vt}) is the variance of the two-sided error term at time t.

After estimating the parameters, the point estimates of technical efficiency can then be measured using Battese and Coelli (1988) formula:

$$TE\left[- \mu_{ijt} | \epsilon_{ijt}\right] = \left[\frac{1-(\sigma^* - \mu_{ijt}^* / \sigma^*)}{1-(-\mu_{ijt}^* / \sigma^*)} \right] exp(- \mu_{ijt}^* + \sigma^{*2}/2)$$

where square(μ*{ijt}) = ε{ijt} − square(σ{v})/σ{μ} and (.) is the standard normal density function. Estimates of technical efficiency for each country-pair range from zero to one. A TE value of one would suggest that actual and potential trade levels are identical, however values trending towards zero would indicate that there is room to increase real trade levels, i.e. a TE value of 1 indicating that it is fully efficient and values less than 1 indicating inefficiency.

The conventional gravity model can be converted to an SFGM variant as follows:

$$Trade_{ijt} = f(GDP_{it}, GDP_{jt}, POP_{it}, POP_{jt}, DIS_{ij}, Z_{ij}, X_{ijt})exp(V_{ijt} - u_{ij})$$

Where V{ijt}−u{ij} is the error term. Trade{ijt} is the bilateral trade between two countries i and j over a period of time t; GDP{it} and GDP{jt} are the economic sizes of both countries; POP{it} and POP{jt} are the populations of countries i and j, respectively; and DIS{ij} is the distance between two countries. There are two components to the error term: the two-sided (V{ijt}) and one-sided error terms (u{ijt}). The two-sided error element (V{ijt}) represents the influence on trade flows of other variables, such as measurement mistakes and the implicit beyond-the-border constraints that are beyond the control of the exporting country and are randomly distributed throughout the sample data. Whereas the one-sided error term (u{ijt}) reveals the cumulative effects of "behind the border" restraints on trade and identifies the degree to which actual trade volumes depart from the maximum possible level. The full gravity model to investigate the determinants of India's trade performance can be stated as follows, based on the available literature and the preceding discussion, and also adding a few relevant dummy variables as was done by Imran Alam and Shahid Ahmed (2018) for simple gravity model (here I'm adding those dummy variables to the SFGM). I will utilise both time and country fixed effects to add the MTR term. (i refers country i, j refers country j and t refers time period (year))

lnTrade{ijt} = α{ij} + β1lnGDP{it} + β2lnGDP{jt} + β3lnPOP{it} + β4lnPOP{jt} + β5lnDIS{ij} + β6lnREX{it} + β7lnREX{jt} + β8lnINFR{it} + β9lnINFR{jt} + β10lnINS{it} + β11lnINS{jt} + β12TA_FTA{ij} + β13lnMR{ij} + β14COMCOL{ij} + β15TradeAffinity{ij} + β16Diaspora{ij} + v{ij} − μ

Here, MR{ij} is the multilateral trade resistance term defined as follows -

$$MR_{ij} = \sum_{k} \left(\frac{Y_k}{Y}\right)^{\theta} (1 + D_{ik}^{\gamma})^{-\alpha} (1 + D_{jk}^{\gamma})^{-\beta}$$

where Y{k} is the economic size of country k, Y is the total world GDP. θ, α, β are parameters to be estimated. γ is a distance decay parameter.

The dummy variable TA_FTA{ij}, which reflects the vector of time-invariant explanatory factors, is added to the gravity model. TA_FTA{ij} is a dummy variable that indicates membership in a trade integration or a free-trade-agreement (FTA); it has a value of 1 if both the reporting country and the partner country are members of the same sub-regional trade integration, or have a FTA between themselves. (It is the practical effect of India-GCC FTA in our case).

In addition, a vector of time-varying explanatory factors, such as the actual exchange rate, infrastructure level (INFR), and institutional quality (INS), for both the reporting and trading partners are added to the model. Finally, v{ijt}−μ is the error term as previously defined.

The other dummy variables which have been added are (they have been taken from Imran Alam and Shahid Ahmed (2018)) -

TradeAffinity (Trading Affinity) – If a country has on average more than one per cent share in India's global trade from 2004 to 2022, then dummy value will be one otherwise it will be zero. Expected sign of trading affinity is positive.

Diaspora - If average numbers of country i Diaspora population in country j is more than 10 percent of total population of country j for the period of 2004 to 2022, dummy value will be one otherwise it will be zero. Expected sign of Diaspora is positive.

COMCOL (common coloniser) – if country i and country j were colonies with the same coloniser then it will be one otherwise zero.

Apart from this, the other variables are -

GDPit/jt : Gross Domestic Product of country i/j
POPitj/t : Population of country i/j
DISij : Distance between countries i & j
REXit/jt : Real Exchange Rate of country i/j
INFRit/jt : Level of Infrastructure of country i/j
INSit/jt : Institutional quality for both reporting and trading partner of country i/j
Vijt - u : Error term

Battese and Coelli (1988) proposed a formula for calculating technical efficiency (TE) in a stochastic frontier gravity model. The formula is:

TE(i) = exp(-u(i))

where i refers to the i-th firm in the sample and u(i) is the inefficiency term for the firm. The inefficiency term is assumed to be a random variable that captures the deviation of firm i's observed output from its optimal output level given its inputs.

In the Battese and Coelli (1988) formula, the inefficiency term u(i) is decomposed into two parts: a stochastic component v(i) and a deterministic component x(i)β, where x(i) is a vector of firm i's input values and β is a vector of unknown parameters. The formula is written as:

u(i) = x(i)β + v(i)

The stochastic component v(i) is assumed to follow a half-normal distribution with zero mean and variance σ(v^2), while the deterministic component x(i)β is non-negative.

Using this formula, TE(i) is estimated by taking the ratio of the observed output y(i) to the estimated optimal output level y(i)*. The estimated optimal output level is calculated as y(i)* = exp(x(i)β), and is obtained by maximising the likelihood function of the model given the data.

Therefore, the Battese and Coelli (1988) formula for TE(i) can be written as:

TE(i) = y(i) / exp(x(i)β + v(i)

This formula measures the extent to which firm i is operating at its optimal output level given its inputs, with a TE value of 1 indicating that the firm is fully efficient and values less than 1 indicating inefficiency.

We will be estimating the above model using the following 2 estimators:

1. True fixed effects model
2. Random effects GLS model

We will be using Stata for the estimation.

The gravity model's capacity to forecast future trade flows between two sides is another helpful feature. Using the coefficient value from the augmented gravity model, the trade potential between India and the GCC can be determined. The entire trade potential for the most recent time, 2022, can be estimated. India's potential for trade with the GCC can be determined by the ratio of the computed trade value from the augmented model (P) to the actual trade value between India and the GCC countries (A). Stated differently, a P/A score more than one indicates that India has the potential to trade with that nation. The difference, or P-A, between the computer trade value (P) and real trade value (A), or trade potential, can also be calculated to see the trade potential in absolute numbers. {this method for gauging export potential was used by Imran Alam and Shahid Ahmed (2018)}. This exercise can be used by policy-makers to gather insights and hence align their policy with recommendations from theory.

We use panel data (2004-2022) for our analysis. The use of panel data helps us in eliminating any bias caused by differences among the countries. The data obtained is for aggregate volume of trade, populations of the countries, distance between the countries, GDP data of the countries, the real exchange rate between the countries at various points in time, diaspora data, etc.

Random and Fixed Effects -

Fixed effects and random effects refer to different ways of modelling unobserved heterogeneity. Fixed effects are estimated using least squares (or maximum likelihood) and capture individual differences by differences in the intercept parameter. They are used when there is a dependence between unobserved factors and the observed independent variables. Random effects are estimated with shrinkage and capture individual differences by the intercept parameters, but treat the individual differences as random rather than fixed. They are used when the unobserved effects are independent of the observed independent variables.

Now we move onto one of the methods of estimation; the fixed effects model. We know that the fixed effects models are useful because they control for unobserved heterogeneity between individuals or entities that is constant over time, which OLS cannot do. We will run the code for the fixed effects model and see if the results are mostly of sobriety and matching the expectations.

To check and account for unobserved heterogeneity, we also try the random effects model using the respective code for it and check its results too. We will also store the results data of both the codes for comparing the two effects.

Data Used and its Sources

We use panel data (2004-2023) for our analysis. The use of panel data helps us in eliminating any bias caused by differences among the countries. The data obtained is for aggregate volume of trade, populations of the countries, distance between the countries, GDP data of the countries, the real exchange rate between the countries at various points in time, diaspora data, etc. The data is procured from trustable sources like -

- World Bank's World Development Indicators (for data on Real GDP, population, diaspora and infrastructure): World Development Indicators | DataBank (worldbank.org)
- World Bank's World Integrated Trade Solution (WITS): World Integrated Trade Solution (WITS) | Data on Export, Import, Tariff, NTM (worldbank.org)
- International Monetary Fund's Direction of Trade Statistics (for data on data on trade volume): Direction of Trade Statistics (DOTS) - dataset by imf | data.world
- UNCTAD database (for data on both the tariffs and trade): UNCTADstat
- Ministry of Commerce and Industry, Govt. of India (for export import data): Export Import Data Bank (Monthly) - Mcommerce
- Centre D' Etudes Prospectives et D' Informations Internationales (CEPII) (COMLANG and COMCOL data): CEPII - Gravity

- GCC-Stat (for high quality GCC specific data): [GCC Statistical Center - Home]
- Geographical distances were procured using Google Maps satellite data: [Google Maps]
- For Institutional quality we have used the control of corruption variable as a proxy Worldwide Governance Indicators from worldbank.org

Link to our dataset -
https://docs.google.com/spreadsheets/d/1Z6loGOQy5pdX8Y8fRH8HGFbi4QqwTfhZRoLYNGw6I0Y/

ID	Year	Countries	trade_ijt	gdp_it	gdp_it (in billion $)	gdp_jt	gdp_jt (in billion $)	pop_it	pop_jt
1	2004	India-Kuwait	421.44	709,000,000,000.00	709.00	59,400,000,000.00	59.40	1,140,000,000.00	2,153,481.00
1	2005	India-Kuwait	513.73	820,000,000,000.00	820.00	80,800,000,000.00	80.80	1,150,000,000.00	2,235,403.00
1	2006	India-Kuwait	614.81	940,000,000,000.00	940.00	102,000,000,000.00	102.00	1,170,000,000.00	2,363,409.00
1	2007	India-Kuwait	681.54	1,220,000,000,000.00	1,220.00	115,000,000,000.00	115.00	1,190,000,000.00	2,506,769.00
1	2008	India-Kuwait	797.50	1,200,000,000,000.00	1,200.00	147,000,000,000.00	147.00	1,210,000,000.00	2,650,930.00
1	2009	India-Kuwait	782.45	1,340,000,000,000.00	1,340.00	106,000,000,000.00	106.00	1,220,000,000.00	2,795,550.00
1	2010	India-Kuwait	1,856.01	1,680,000,000,000.00	1,680.00	115,000,000,000.00	115.00	1,240,000,000.00	2,943,356.00
1	2011	India-Kuwait	1,181.41	1,820,000,000,000.00	1,820.00	154,000,000,000.00	154.00	1,260,000,000.00	3,143,825.00
1	2012	India-Kuwait	1,061.08	1,830,000,000,000.00	1,830.00	174,000,000,000.00	174.00	1,270,000,000.00	3,394,663.00
1	2013	India-Kuwait	1,061.14	1,860,000,000,000.00	1,860.00	174,000,000,000.00	174.00	1,290,000,000.00	3,646,518.00
1	2014	India-Kuwait	1,198.89	2,040,000,000,000.00	2,040.00	163,000,000,000.00	163.00	1,310,000,000.00	3,761,584.00
1	2015	India-Kuwait	1,247.51	2,100,000,000,000.00	2,100.00	115,000,000,000.00	115.00	1,320,000,000.00	3,908,743.00
1	2016	India-Kuwait	1,497.99	2,290,000,000,000.00	2,290.00	109,000,000,000.00	109.00	1,340,000,000.00	4,048,085.00
1	2017	India-Kuwait	1,365.66	2,650,000,000,000.00	2,650.00	121,000,000,000.00	121.00	1,350,000,000.00	4,124,904.00
1	2018	India-Kuwait	1,333.92	2,700,000,000,000.00	2,700.00	138,000,000,000.00	138.00	1,370,000,000.00	4,317,185.00
1	2019	India-Kuwait	1,286.56	2,840,000,000,000.00	2,840.00	136,000,000,000.00	136.00	1,380,000,000.00	4,441,100.00
1	2020	India-Kuwait	1,054.20	2,670,000,000,000.00	2,670.00	106,000,000,000.00	106.00	1,400,000,000.00	4,360,444.00
1	2021	India-Kuwait	1,241.93	3,150,000,000,000.00	3,150.00	137,000,000,000.00	137.00	1,410,000,000.00	4,250,114.00
1	2022	India-Kuwait	1,560.45	3,420,000,000,000.00	3,420.00	175,000,000,000.00	175.00	1,420,000,000.00	4,268,873.00
2	2004	India-Bahrain	156.46	709,000,000,000.00	709.00	13,200,000,000.00	13.20	1,140,000,000.00	833,451.00
2	2005	India-Bahrain	192.25	820,000,000,000.00	820.00	16,000,000,000.00	16.00	1,150,000,000.00	901,921.00
2	2006	India-Bahrain	184.52	940,000,000,000.00	940.00	18,500,000,000.00	18.50	1,170,000,000.00	970,981.00
2	2007	India-Bahrain	252.47	1,220,000,000,000.00	1,220.00	21,700,000,000.00	21.70	1,190,000,000.00	1,040,532.00
2	2008	India-Bahrain	286.52	1,200,000,000,000.00	1,200.00	25,700,000,000.00	25.70	1,210,000,000.00	1,110,343.00
2	2009	India-Bahrain	250.21	1,340,000,000,000.00	1,340.00	22,900,000,000.00	22.90	1,220,000,000.00	1,179,453.00
2	2010	India-Bahrain	651.83	1,680,000,000,000.00	1,680.00	25,700,000,000.00	25.70	1,240,000,000.00	1,213,645.00
2	2011	India-Bahrain	439.99	1,820,000,000,000.00	1,820.00	28,800,000,000.00	28.80	1,260,000,000.00	1,212,077.00
2	2012	India-Bahrain	603.47	1,830,000,000,000.00	1,830.00	30,700,000,000.00	30.70	1,270,000,000.00	1,224,939.00
2	2013	India-Bahrain	639.36	1,860,000,000,000.00	1,860.00	32,500,000,000.00	32.50	1,290,000,000.00	1,261,673.00
2	2014	India-Bahrain	472.98	2,040,000,000,000.00	2,040.00	33,400,000,000.00	33.40	1,310,000,000.00	1,311,134.00
2	2015	India-Bahrain	654.14	2,100,000,000,000.00	2,100.00	31,100,000,000.00	31.10	1,320,000,000.00	1,362,142.00
2	2016	India-Bahrain	471.71	2,290,000,000,000.00	2,290.00	32,200,000,000.00	32.20	1,340,000,000.00	1,409,661.00
2	2017	India-Bahrain	556.82	2,650,000,000,000.00	2,650.00	35,500,000,000.00	35.50	1,350,000,000.00	1,456,834.00
2	2018	India-Bahrain	742.14	2,700,000,000,000.00	2,700.00	37,800,000,000.00	37.80	1,370,000,000.00	1,487,340.00
2	2019	India-Bahrain	559.09	2,840,000,000,000.00	2,840.00	38,700,000,000.00	38.70	1,380,000,000.00	1,494,188.00

ID	Year	dis_ij	rex_it	rex_jt	infr_it (1 to 5)	infr_jt (1 to 5)	ins_it (scaled)	ins_it	ins_jt (scaled)	ins_jt	cb_ij	ta_ij (FTA)	mtr_ij	COMLANG	COMCOL	TradeAffinity	Diaspora
1	2004	3,305.00	45.32	0.29	2.82	2.49	0.41	-0.45	0.65	0.77	0.00	0.00	11.94	0.00	1.00	0.00	0.00
1	2005	3,305.00	44.10	0.29	2.83	2.56	0.43	-0.36	0.60	0.50	0.00	0.00	10.15	0.00	1.00	0.00	0.00
1	2006	3,305.00	45.31	0.29	2.86	2.59	0.45	-0.27	0.59	0.43	0.00	0.00	9.22	0.00	1.00	0.00	0.00
1	2007	3,305.00	41.35	0.28	2.90	2.64	0.42	-0.40	0.57	0.36	0.00	0.00	10.61	0.00	1.00	0.00	0.00
1	2008	3,305.00	43.51	0.27	2.90	2.67	0.43	-0.34	0.58	0.41	0.00	0.00	8.16	0.00	1.00	0.00	0.00
1	2009	3,305.00	48.41	0.29	2.95	2.69	0.41	-0.45	0.56	0.30	0.00	0.00	12.64	0.00	1.00	0.00	0.00
1	2010	3,305.00	45.73	0.29	2.91	3.45	0.41	-0.46	0.56	0.30	0.00	1.00	7.30	0.00	1.00	0.00	1.00
1	2011	3,305.00	46.67	0.28	2.89	3.22	0.39	-0.54	0.52	0.08	0.00	1.00	5.91	0.00	1.00	0.00	1.00
1	2012	3,305.00	53.44	0.28	2.88	3.23	0.40	-0.51	0.46	-0.20	0.00	1.00	5.26	0.00	1.00	0.00	1.00
1	2013	3,305.00	58.60	0.28	2.99	3.34	0.40	-0.52	0.46	-0.20	0.00	1.00	5.34	0.00	1.00	0.00	1.00
1	2014	3,305.00	61.03	0.28	2.88	3.34	0.41	-0.46	0.45	-0.25	0.00	1.00	6.26	0.00	1.00	0.00	1.00
1	2015	3,305.00	64.15	0.30	3.01	3.41	0.42	-0.41	0.45	-0.25	0.00	1.00	9.13	0.00	1.00	0.00	1.00
1	2016	3,305.00	67.20	0.30	3.34	3.47	0.43	-0.34	0.44	-0.30	0.00	1.00	10.50	0.00	1.00	0.00	1.00
1	2017	3,305.00	65.12	0.30	3.21	3.48	0.44	-0.29	0.43	-0.36	0.00	1.00	10.95	0.00	1.00	0.00	1.00
1	2018	3,305.00	68.39	0.30	2.91	3.39	0.45	-0.23	0.44	-0.32	0.00	1.00	9.78	0.00	1.00	0.00	1.00
1	2019	3,305.00	70.42	0.30	3.01	3.48	0.44	-0.30	0.47	-0.16	0.00	1.00	10.44	0.00	1.00	0.00	1.00
1	2020	3,305.00	74.10	0.31	3.02	3.48	0.44	-0.29	0.48	-0.09	0.00	1.00	12.59	0.00	1.00	0.00	1.00
1	2021	3,305.00	73.92	0.30	3.11	3.56	0.44	-0.32	0.49	-0.06	0.00	1.00	11.50	0.00	1.00	0.00	1.00
1	2022	3,305.00	78.60	0.31	3.25	3.59	0.44	-0.32	0.53	0.13	0.00	1.00	9.77	0.00	1.00	0.00	1.00
2	2004	2,955.00	45.32	0.38	2.82	3.20	0.41	-0.45	0.59	0.44	0.00	0.00	53.71	0.00	1.00	0.00	0.00
2	2005	2,955.00	44.10	0.38	2.83	3.21	0.43	-0.36	0.58	0.38	0.00	0.00	51.25	0.00	1.00	0.00	0.00
2	2006	2,955.00	45.31	0.38	2.86	3.24	0.45	-0.27	0.54	0.19	0.00	0.00	50.81	0.00	1.00	0.00	0.00
2	2007	2,955.00	41.35	0.38	2.90	3.41	0.42	-0.40	0.54	0.18	0.00	0.00	56.22	0.00	1.00	0.00	0.00
2	2008	2,955.00	43.51	0.38	2.90	3.44	0.43	-0.34	0.54	0.18	0.00	0.00	46.69	0.00	1.00	0.00	0.00
2	2009	2,955.00	48.41	0.38	2.95	3.36	0.41	-0.45	0.54	0.18	0.00	0.00	58.52	0.00	1.00	0.00	0.00
2	2010	2,955.00	45.73	0.38	2.91	3.36	0.41	-0.46	0.54	0.18	0.00	1.00	32.68	0.00	1.00	0.00	1.00
2	2011	2,955.00	46.67	0.38	2.89	3.00	0.39	-0.54	0.54	0.21	0.00	1.00	31.60	0.00	1.00	0.00	1.00
2	2012	2,955.00	53.44	0.38	2.88	3.08	0.40	-0.51	0.57	0.37	0.00	1.00	29.80	0.00	1.00	0.00	1.00
2	2013	2,955.00	58.60	0.38	2.99	3.24	0.40	-0.52	0.59	0.43	0.00	1.00	28.62	0.00	1.00	0.00	1.00
2	2014	2,955.00	61.03	0.38	2.88	3.04	0.41	-0.46	0.55	0.27	0.00	1.00	30.54	0.00	1.00	0.00	1.00
2	2015	2,955.00	64.15	0.38	3.01	3.28	0.42	-0.41	0.52	0.11	0.00	1.00	33.76	0.00	1.00	0.00	1.00
2	2016	2,955.00	67.20	0.38	3.34	3.10	0.43	-0.34	0.49	-0.05	0.00	1.00	35.56	0.00	1.00	0.00	1.00
2	2017	2,955.00	65.12	0.38	3.21	2.72	0.44	-0.29	0.47	-0.17	0.00	1.00	37.32	0.00	1.00	0.00	1.00
2	2018	2,955.00	68.39	0.38	2.91	3.73	0.45	-0.23	0.46	-0.18	0.00	1.00	35.71	0.00	1.00	0.00	1.00
2	2019	2,955.00	70.42	0.38	3.01	3.63	0.44	-0.30	0.49	-0.04	0.00	1.00	36.69	0.00	1.00	0.00	1.00

STATA Code

```
* first normalise the variables *

gen logY = log(trade_ijt/1000)
gen logX1 = log(gdp_it/100000000000)
gen logX2 = log(gdp_jt/100000000000)
gen logX3 = log(dis_ij/1000)
gen logX4 = log(pop_it/1000000000)
gen logX5 = log(pop_jt/1000000000)
gen logX6 = log(rex_it/100)
gen logX7 = log(rex_jt/100)
gen logX8 = log(mtr_ij)
gen X9 = Diaspora
gen X10 = TradeAffinity
gen X11 = log(infr_it1to5)
gen X12 = log(infr_jt1to5)
gen X14 = log(ins_itscaled)
gen X16 = log(ins_jtscaled)
gen X18 = ta_ijFTA
gen X20 = COMCOL



* define panel data *

xtset id Year

* run panel sfa *

sfpanel logY logX1 logX2 logX3 logX4 logX5 logX6 logX7 logX8 X9 X10 X11 X12 X14 X16  X18  X20 Year,
model(tfe) distribution(tn) efficiency(eta_ij)
```

```
. predict eff, u
(876 missing values generated)

. sort eff

. predict effi
(option xb assumed; fitted values)
(876 missing values generated)

. predict te
(option xb assumed; fitted values)
(876 missing values generated)

. egen fegrp = cut(effi), at(0, .3, .6, .9, 1)
(950 missing values generated)

. tab fegrp

      fegrp |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |         13       32.50       32.50
         .3 |         17       42.50       75.00
         .6 |         10       25.00      100.00
------------+-----------------------------------
      Total |         40      100.00

.
```

Explanation of the code -

1. import excel "C:\Users\HP\Downloads\ECO342A_project_data.xlsx", sheet("Sheet1") firstrow
    - This line imports data from an Excel file located at the specified path (C:\Users\HP\Downloads\ECO342A_project_data.xlsx) and selects the sheet named "Sheet1". firstrow indicates that the first row of the Excel sheet contains variable names.

2. gen logY = log(trade_ijt/1000)
    - This line creates a new variable named logY by taking the natural logarithm (log()) of the variable trade_ijt divided by 1000.

3. gen logX1 = log(gdp_it/100000000000)
    - Similarly, this line creates a new variable named logX1 by taking the natural logarithm of the variable gdp_it divided by 100,000,000,000.

4. Lines 3 to 14
    - These lines follow the same pattern as the previous ones, creating new variables (logX2, logX3, ..., X20) by taking the natural logarithm of various variables divided by different constants.

5. xtset id Year
    - This line sets up the data as panel data, where id represents the individual identifier and Year represents the time variable.

6. sfpanel logY logX1 logX2 logX3 logX4 logX5 logX6 logX7 logX8 X9 X10 X11 X12 X14 X16 X18 X20 Year, model(tfe) distribution(tn) efficiency(eta_ij)
    - This line runs a stochastic frontier analysis (SFA) using panel data. It specifies the dependent variable (logY) and independent variables (logX1, logX2, ..., X20). model(tfe) specifies the model to be estimated (in this case, time-fixed effects), distribution(tn) specifies the distribution (in this case, truncated normal), and efficiency(eta_ij) specifies that the efficiency will be estimated.

7. predict eff, u
    - This line predicts the efficiency residuals (u) and stores them in a new variable named eff.

8. sort eff
   - This line sorts the dataset based on the variable eff (efficiency).

9. predict effi
   - This line predicts the fitted values of efficiency and stores them in a new variable named effi.

10. predict te
   - This line predicts the technical efficiency scores and stores them in a new variable named te.

11. egen fegrp = cut(effi), at(0, .3, .6, .9, 1)
   - This line creates a new categorical variable named fegrp based on the variable effi. It categorizes the efficiency scores into four groups: 0 to 0.3, 0.3 to 0.6, 0.6 to 0.9, and 0.9 to 1.

12. tab fegrp
   - This line displays a frequency table of the variable fegrp, showing the distribution of efficiency groups.

Hypothesis/Expectations/Tentative Conclusions

We expect India to have a good trade potential with all the GCC countries. The order is difficult to predict beforehand. Moreover, when the FTA (Free Trade Agreement) dummy takes the value of one, we expect the results to show a sharp increase in the magnitude of export potential. This can be checked by plotting a bar chart.

A positive and significant relationship is expected to exist between bilateral trade flows and economic, geographic, and cultural factors among the GCC countries and India. Countries that are more technically efficient in producing and exporting goods are expected to have higher levels of bilateral trade with their trading partners. The level of infrastructure, population, GDP and Institutional quality for both reporting and trading partners of the country are expected to influence bilateral trade flows positively. The coefficient of the real exchange rate is expected to be negative. Physical distance is expected to have a negative effect. The effect of the diaspora, common border, common coloniser and sub-regional integration is expected to be positive. It is expected that a common coloniser will help to improve trade negotiation, historical ties, etc and further it will reduce transaction cost. So, the expected sign of COMCOL is positive. The coefficient for the population parameters is hard to predict and I will reserve my judgement till the empirical analysis.

We also do not expect the variables population and distance to be statistically significant, as with the advancements in technology, these two factors have been rendered a bit less relevant to determining trade, as technology and modern transport can easily bridge global gaps, and enable a smaller population to produce more.

Results

Results from true fixed effects model :

```
True fixed-effects model (truncated-normal)         Number of obs =        114
Group variable: id                               Number of groups =          6
Time variable: Year                            Obs per group: min =         19
                                                              avg =       19.0
                                                              max =         19

                                                    Prob > chi2   =     0.0000
Log likelihood =      9.1807                        Wald chi2(15)  =     619.40
```

logY	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Frontier						
logX1	.0947607	.5483247	0.17	0.863	-.9799359	1.169457
logX2	.6086709	.2660964	2.29	0.022	.0871316	1.13021
logX3	-.2419998	123.0267	-0.00	0.998	-241.37	240.886
logX4	.6445064	6.641343	0.10	0.923	-12.37229	13.6613
logX5	.0641234	.2585405	0.25	0.804	-.4426066	.5708535
logX6	.1022291	.6263802	0.16	0.870	-1.125453	1.329912
logX7	.0460733	1.800039	0.03	0.980	-3.481938	3.574084
logX8	-.1937004	.2172358	-0.89	0.373	-.6194749	.232074
X9	.2537039	.1416092	1.79	0.073	-.0238449	.5312528
X10	.3007104	.1572222	1.91	0.056	-.0074395	.6088604
X11	.0308992	.7115091	0.04	0.965	1.425431	1.363633
X12	.3670527	.3897349	0.94	0.346	-.3968137	1.130919
X14	.0044278	.7837612	0.01	0.995	-1.531716	1.540572
X16	.4600722	.3242184	1.42	0.156	1.095529	.1753843
X18	0	(omitted)				
X20	.0837597	.	.	.	.	.
Year	.0001553	.0736659	0.00	0.998	-.1442272	.1445377
Mu						
_cons	-2.980144	7.946468	-0.38	0.708	-18.55493	12.59465
Usigma						
_cons	-1.648198	2.923832	-0.56	0.573	-7.378804	4.082409
Vsigma						
_cons	-3.074647	.2923984	-10.52	0.000	-3.647737	-2.501556
sigma_u	.4386301	.6412404	0.68	0.494	.0249869	7.699876
sigma_v	.2149557	.0314263	6.84	0.000	.1614002	.2862819
lambda	2.040561	.6556019	3.11	0.002	.7556045	3.325517

Results from random effects GLS regression:

```
Random-effects GLS regression                   Number of obs      =       114
Group variable: id                              Number of groups   =         6

R-sq:                                           Obs per group:
     within  = 0.8247                                        min =        19
     between = 0.9995                                        avg =      19.0
     overall = 0.9727                                        max =        19

                                                Wald chi2(15)      =   3486.07
corr(u_i, X)   = 0 (assumed)                    Prob > chi2        =    0.0000
```

logY	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
logX1	.8707557	.5874571	1.48	0.138	-.280639	2.02215
logX2	.0547162	.2632483	0.21	0.835	-.461241	.5706734
logX3	-6.064565	.38128	-15.91	0.000	-6.81186	-5.31727
logX4	4.158651	5.119696	0.81	0.417	14.19307	5.875769
logX5	1.054302	.1378174	7.65	0.000	.784185	1.324419
logX6	.1719895	.6675622	0.26	0.797	-1.136408	1.480387
logX7	-.1286734	.0449439	-2.86	0.004	-.2167619	-.040585
logX8	-.2934638	.2400911	-1.22	0.222	-.7640336	.177106
X9	.0411595	.1547966	0.27	0.790	.3445553	.2622362
X10	.8321559	.1287204	6.46	0.000	.5798686	1.084443
X11	.0976656	.8044049	0.12	0.903	1.67427	1.478939
X12	.2624903	.3989972	0.66	0.511	-.5195299	1.044511
X14	.8332194	.8411046	0.99	0.322	2.481754	.8153153
X16	.4929033	.3626136	1.36	0.174	1.203613	.2178063
X18	0	(omitted)				
X20	1.394566	.1456252	9.58	0.000	1.109146	1.679986
_cons	9.207773	1.603777	5.74	0.000	6.064428	12.35112
sigma_u	0					
sigma_e	.23885373					
rho	0	(fraction of variance due to u_i)				

The results of the regression using true fixed effects and random effects are both quite similar and are as per the expectations. Let us look at the coefficients of the various variables to see how.

- First of all, the coefficients of the GDP of both countries are positive, indicating that trade is positively proportional to the size of the countries, as per our expectations.
- Then we have a negative coefficient of the distance between the countries, which is also as per the logical expectations.
- INFR and INS too have a positive effect, as was expected.
- Next, we look at the coefficients of the populations of both countries; they too are positive, as was expected.
- The real exchange rate is negatively proportional to the trade; this too satisfies our prediction.
- Now comes the important part: the coefficients of the Multilateral Trade Resistance (MTR) term, Diaspora, and Trade Affinity. We expect the coefficients of these to be negative, positive, and positive, in that order, respectively. The regression results satisfy our expectations, with MTR having a negative coefficient, Diaspora having a positive coefficient, and Trade Affinity also having a positive coefficient.

```
. list id Year eff effi te in 1/10
```

	id	Year	eff	effi	te
1.	1	2010	.0348214	.2595987	.2595987
2.	5	2013	.0358773	2.49545	2.49545
3.	5	2014	.0384515	2.496979	2.496979
4.	2	2010	.0405552	-.949262	-.949262
5.	5	2012	.0413189	2.45526	2.45526
6.	6	2010	.0417411	1.536267	1.536267
7.	4	2021	.0420773	.8106366	.8106366
8.	3	2016	.0430962	.1730329	.1730329
9.	6	2011	.0439841	1.658005	1.658005
10.	3	2021	.0446317	.3564975	.3564975

Technical Efficiency Table

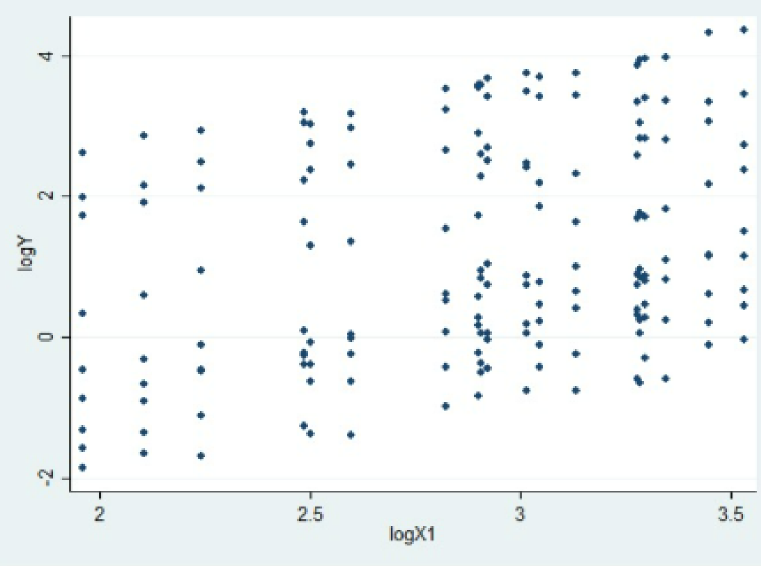
Explanation of the code used for generating the technical efficiency table : This code will generate an efficiency table displaying the first 10 observations for the variables id, Year, eff (efficiency residuals), effi (fitted values of efficiency), and te (technical efficiency scores). The list command parameters (in 1/10) can be adjusted to display more or fewer observations as needed

- The coefficient for Technical Inefficiency comes out to be negative. Thus, we can say that entering into Trade Agreements has a positive impact on Technical Efficiency .
- Checking the proportion of variation in trade accounted for by technical inefficiency:
- $\gamma = \sigma^2u / (\sigma^2u + \sigma^2v) = 0.8003$
- Thus, almost 80% of the variation in exports from India to members of GCC arises due to Technical Inefficiency.

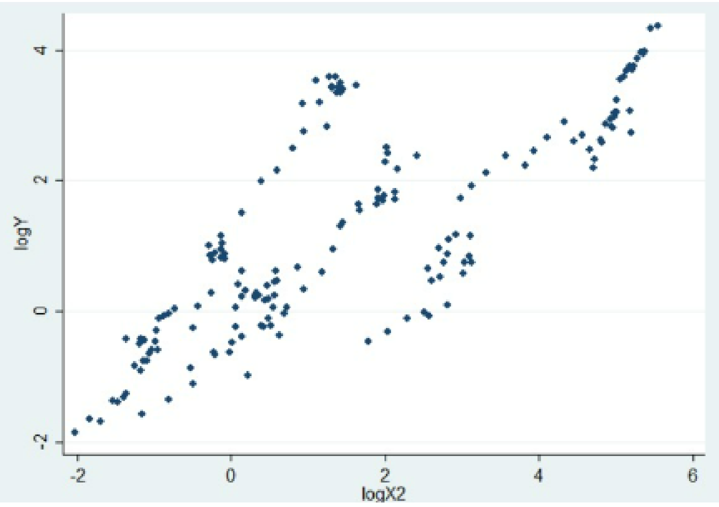
It can be seen from the technical efficiency table that Oman has the highest trade potential with India.

Also on average, India's trade with GCC has a Technical efficiency of 30%, hence there is a lot of scope for improving trade with the GCC, the FTA would be useful in that.

For robustness checks, we would check for the statistical significance of the various coefficients obtained by comparing their p-values with their significance level.



Scatter Plot of India's GDP with Trade



Scatter Plot of GCC Countries' GDP with Trade

Both plots show a positive trend, which is as per the usual predictions made by the gravity model.

Finally, we move on to analysing the effect of the FTA on the trade volume between India and GCC Countries, for this, we have plotted a bar chart for the aggregate trade of India with GCC versus year. As can be seen, using the plot attached above, trade saw a significant rise 2004 onwards, which is when the FTA was signed, trade also saw a significant jump in 2009, which was the year when some of the terms of the FTA between India and GCC countries were brought into effect. After that, further talks with the GCC have been withheld, as GCC has deferred its negotiations with all countries and economic groups and is currently reviewing its negotiations with all countries and economic groups. Efforts are being made at various bilateral/multilateral forums for early resumption of the negotiations. This can be seen in the form of stagnation in trade after 2009, leading us to the conclusion that not much change in ease of trade has happened between the GCC and India after 2009, and there certainly is scope for further negotiations. The sudden drop in 2020 can be attributed to the COVID Pandemic.

The policy recommendation hence is to re-start FTA negotiations with GCC, as an FTA would be beneficial for India's export to the GCC countries. Our model also shows that India's export potential is highest with Oman, followed by Qatar and then Bahrain. Hence, if bilateral agreements are to be undertaken, then India should mainly focus on these members of the Gulf Cooperation Council.

References

- Tinbergen, Jan. Shaping the World Economy: Suggestions for an International Economic Policy. Books (Jan Tinbergen). Twentieth Century Fund, New York, 1962. http://hdl.handle.net/1765/16826
- Pöyhönen, Pentti. A Tentative Model for the Volume of Trade between Countries. Weltwirtschaftliches Archiv 90 (1963): 93–100. http://www.jstor.org/stable/40436776
- Anderson, James E., and Eric van Wincoop. Gravity with Gravitas: A Solution to the Border Puzzle. The American Economic Review 93, no. 1 (2003): 170–92. http://www.jstor.org/stable/3132167
- Baier, Scott & Bergstrand, Jeffrey. (2009). Estimating the Effects of Free Trade Agreements on International Trade Flows Using Matching Econometrics. Journal of International Economics. 77. 63-76. https://doi.org/10.1016/j.jinteco.2008.09.006
- Yotov, Y. V., Piermartini, R., Monteiro, J.-A., & Larch, M. (2016). An Advanced Guide to Trade Policy Analysis: The Structural Gravity Model. Geneva: World Trade Organization. https://www.wto.org/english/res_e/booksp_e/advancedwtounctad2016_e.pdf
- Samir Pradhan (2009). India's Export Potential to the Gulf Cooperation Council: A Gravity Model Exploration. https://www.researchgate.net/publication/46563584
- **[Base Paper]** Imran Alam, Shahid Ahmed (2018). India's Trade with Gulf Cooperation Council (GCC) Countries: a Panel Gravity Model Analysis. https://www.researchgate.net/publication/328429188
- Ebaidalla M. Ebaidalla & Mohammed Elhaj Mustafa Ali (2023). Assessing Intra-Arab Trade Integration and Potential: Evidence from the Stochastic Frontier Gravity Model, The International Trade Journal, 37:2, 221-239, DOI: 10.1080/08853908.2022.2029725: erf.org.eg/app/uploads/2018/02/Ebaidalla-M.-Ebaidalla-Mohamed.pdf

Plagiarism Report

https://drive.google.com/drive/folders/1M8pYXL4sPEKEVFsqH_zzj2CtlmMarfWs?usp=sharing