DEPOCEN

Working Paper Series No. 2009/18

Gravity Equation for Different Product Groups: A study at product level

Thuy Nguyen*
Jean-Louis Arcand**

* Graduate Institute of International and Development Studies, Geneva, Switzerland
Thuythu.nguyen@graduateinstitute.ch
** Graduate Institute of International and Development Studies, Geneva, Switzerland

Gravity Equation for Different Product Groups: A study at product level*

Thuy Nguyen
Graduate Institute of International and Development Studies
Geneva, Switzerland

Supervisor: Professor Jean-Louis Arcand
Graduate Institute of International and Development Studies
Geneva, Switzerland

June, 2009

Abstract

Do homogeneous and heterogeneous goods response the same way to changes in income and different measures of distance? Running country-fixed-effect gravity equation for different product groups, I find that homogeneous goods are less responsive to changes in income than heterogenous goods. I also find that export volume of all product types is significantly hindered by geographical distance between countries. However, exports of homogeneous goods are not affected by social distance measures such as common language and colonial relationship, while exports of heterogeneous goods significantly improve if trading parties speak the same official language and have colonial relationship. Fixed effect quantile estimation (Koenker 2004) with bootstrapped standard errors confirms the above finding for income and geographical distance. Regarding two social distance measures, common language and colonial relationship, median quantile (Tobit) estimation suggests that common language does not have impact on exports of any product type, while colonial relationship significantly influences export of heterogeneous goods. At higher levels of quantiles the impact of common language increases for all product types, and even strongest on exports of homogeneous goods. Colonial relationship loses its impact as being evaluated at 90th percentile.

*I owe great gratitude to professor Marcelo Olarreaga, University of Geneva, for suggesting me to do this study at product level.

I: Introduction

The purpose of this paper is to evaluate whether trade flow of different product classifications (homogeneous, reference priced, and differentiated) responses the same way to changes in income and distance.[1] Measure of income often being used is countries' gross domestic product (GDP), while measures of distance include geographical distance and an indicator of whether countries share a common border. There are also social distance measures such as variables indicating whether two trading partners speak a common language, and whether they share a colonial relationship. Running gravity equation for different product groups, we can assess how well trade flow of each product group is explained by GDPs, and different measures of distance.

In words, gravity equation states that export flow from one country to another should be proportional to the product of the two countries' GDPs and be inversely proportional to the distance between them. Notice that in this definition, there is no element of neo-classical trade theory in predicting trade flow of products such as technology or factor endowment difference. Instead, the definition of gravity equation reminds us of new trade theory, which explains trade among identical countries (i.e. same level of income, taste, and even latitude.) In fact, the most crystallized mathematical model of gravity equation is derived from new trade theory's assumptions of imperfect competition (in differentiated products) and increasing returns to scale. Much less in theory is talked about the export flow of homogeneous products except in neo-classical trade theory. Neo-classical trade theory asserts that flow of homogeneous goods should be between countries of different endowment and technology.

Should homogeneous goods also be traded more among countries of similar income and in proximity to each other? We are interested in knowing how income and distance would impact the flow of homogeneous compared to differentiated products? Thus, this paper attempts to empirically assess whether gravity equation predicts the trade flow of homogeneous goods the same way it does for differentiated goods.

The paper would run three gravity equations, one for each product group with level of observation at 4-digit SITC goods.[2] Looking at coefficients of income and different measures of distance, we can evaluate how well explanatory variables in gravity equation predicts the flow of different product groups. We are interested in seeing whether variables explaining the export flow of homogeneous goods have coefficients' signs that follow the prediction of gravity equation. If yes, then we would assess whether the income and distance elasticity of homogeneous goods' exports are as strong as those of heterogeneous goods.

The paper is organized as follows: Section II explains how all internationally traded products are classified, and some features of each product type. Section III gives literature review on gravity model and trade flow of goods. Section IV

[1] All international traded commodities are classified into three groups by Rauch (1999).

[2] SITC or Standard International Trade Classification is a system that encodes all internationally trade products. The system makes it easier for compiling and also promoting the comparability of international trade statistics. SITC-revision 2 is examined in this paper.

explains the methodology this paper uses to assess whether all product groups follow assumption of gravity equation the same way. Section V describes data and regression estimation methods used in this paper. Section VI reports empirical results and Section VII concludes the paper.

II: Categorizing products and characteristics of each product type

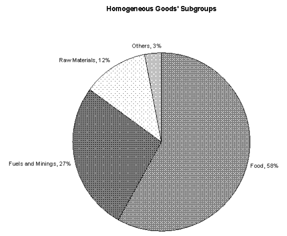
In trade, homogeneous products are considered to be identical, especially in quality, by consumers across nations, whereas heterogeneous goods has distinctive features that separate one from another. Rauch (1999) systematically categorizes all internationally traded products into three groups: goods traded on an organized exchange (homogeneous goods), goods that are not traded on an organized exchange but possessing "reference prices", and goods that are not traded on an organized exchange (heterogeneous goods).[3] Explaining this classification, Rauch says that there are costs associated with setting up markets, or organized exchanges, that is independent of volume of transactions of a good. Thus, if one product can be disaggregated into types and the market for each type is too "thin", then that product would not set up and be traded on organized exchange. Footwear, for example, is a product that can be disaggregated into different types such as slippers, flip-flops, tennis shoes, high-heels, etc. On the one hand, there is no organized exchange for footwear, because the product is not clearly defined. On the other hand, setting up organized exchange market for each type of footwear is too costly in comparison with small volume of transaction for each type. As a result, we can classify footwear as differentiated products since their names and prices are not listed on any board of organized exchange. Similarly, we can also identify homogeneous products as they are traded on organized exchanges. Those are products such as soya beans (SITC-2222), coffee (SITC-0711), gas oil (SITC-3343) or crude oil (SITC-3330). Between homogeneous and heterogeneous product types is a group of products that are not traded on an organized exchange but nevertheless possessing what Rauch calls "reference prices" such as chemical wood pulp (SITC-2518) and copper ore and concentrates (SITC-2871). Based on above reasoning, he classifies all international traded products at 4-digit SITC level into three groups: homogeneous goods, reference prices, and differentiated products.

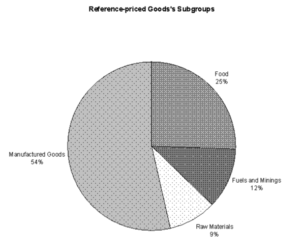
In Rauch list, there are totally 1189 products that are traded in the world. Of which, 146 are classified as homogeneous, 353 as reference-priced, and 690 products as differentiated. Differentiated products account for 58% of total goods traded internationally.

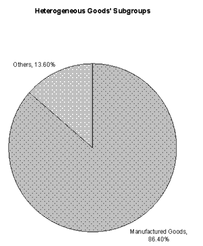
Product Type	**Homogeneous**	**Reference priced**	**Heterogeneous**
Percentage	12%	30%	58%

[3] Commodities traded on an organized exchange means that these commodities are bought and sold in standardized contracts. Examples of organized exchange marketplace for commodities are The Chicago Board of Trade, London Metal Exchange, or the Dalian Board of Trade.

Out of 146 homogeneous goods, 58% are food products, 27% are fuels and minings, and 12% are raw materials. 54% of reference-priced goods are manufactured goods, 25% are food products. Regarding 690 differentiated commodities, 86.4% of which are manufactured goods.[4]







[4] According to the WTO, food includes all goods classified in 1-digit SITC sections 0, 1, 4 and division 22. Raw materials are goods in divisions 21, 23, 24, 25, 26, and 29. Fuels and minings are goods in SITC divisions 27, 28, SITC 3, and division 68. Manufactured products include all goods in SITC 5 through 8 (except division 68 and group 891).

Rauch's classification reveals that homogeneous goods are closely related to food and primary commodities. These products are traded for consumption or production as they were found in nature. Whereas heterogenous goods are linked to manufactured products, in whose process primary goods are transformed into different types of other goods.

Trade flow of food and primary goods, thus homogeneous goods, is different from that of manufactured or heterogeneous goods in several ways. Firstly, the supply of primary and agricultural goods is more volatile since it is likely to be under control of non-human factors such as resource constraint, seasonality and natural disasters. If one big supplier of rice loses one season, it would mean that many other countries would not have rice to buy even though they are willing to spend more on it. Secondly, export flow of primary and food products depends strictly on border control because of food safety regulations or health and sanitation rules.[5] Exports of poultry may subject to lots of inspection and even be quarantined in countries with high toll of Avian Flu deaths during peak time of the disease. Hence, even if one country is productive in raising poultry, its exports might not increase substantially because of food safety regulations.

For many of such reason, Engel's law proposes that primary commodities, especially food, are income-inelastic (income elasticity <1) while manufactured products are income-elastic (income elasticity >1). Manufacturing products are not so often subjected to either natural factors or human-creating rules, thus more easily to get through international borders and traded on larger scale, especially when income levels of exporting or importing countries increase. Also, differentiated goods tend to be traded under imperfect competition. Their prices, hence trade volume (price times quantity), can increase very easily. Whereas homogeneous goods are usually traded under perfect competition. Their trade volume does not respond as strongly to increases in income because of price rigidity and unchangeable quantity produced due to resource dependency. As a result, we should expect that homogeneous goods respond less strongly to changes in income than heterogeneous goods.

II : **Gravity Model and Trade Flow of Goods**

Section II suggests that income plays some role in determining trade volume of commodities, whether they are homogeneous or heterogeneous. However, economists had not used this role to explain trade flow of goods until mid-1960s. Instead, since the time of Ricardo and Adam Smith, they had paid significant attention to comparative advantage to explain trade flow. In Ricardian model, technology difference is the source of comparative advantage that

[5] Article 20 of the General Agreement on Tariffs and Trade (GATT) allows governments to act on trade in order to protect human, animal or plant life or health, provided they do not discriminate or use this as disguised protectionism. The Sanitary and Phytosanitary Measures Agreement by the WTO after the Uruguay Round is a separate agreement on food safety and animal and plant health standards, which allows countries to set their own standards of health and safety regulations.

determines trade. That is, exporting nations are those who are relatively more efficient in making the good. Moving to Ricardo-Viner and Heckscher-Ohlin (HO) models, endowment difference is the source of comparative advantage that determines trade among nations. Countries that are abundant in labor supply should export labor-intensive goods, while those that are abundant in capital should export capital-intensive goods. HO model made huge success in explaining trade between rich and poor countries, or North-South trade. However, it does not explain whether trade occurs between countries that have the same level of endowment, or income.

The reason why income had not been given adequate attention is because neo-classical trade models mentioned above did not take into account the variety of products. For instance, HO model would consider a T-shirt made in Italy the same as a T-shirt made in China. If both China and Italy are equally efficient in making those T-shirts, those T-shirts may be called homogeneous labor-intensive goods in HO model and will not be traded. Nevertheless, in our world today consumers have many reasons to distinguish a T-shirt made in Italy from the one made in China. For instance, the Italian-made T-shirt may be more environmental friendly in the way that it is made of organic cotton and dyed without chemicals. Similarly, consumers can also distinguish a Japanese car from an American one as the Japanese car consumes less gas per mile. Hence, both of these capital-rich countries, Japan and the U.S., still trade with each other. They trade because they have income to produce many varieties and because of their love to consume a vast variety of products.

In 1974, Grubel's & Lloyd's research showed that most of the world trade was not actually explained by HO model. Statistics from the OECD, which show how much overlapped it is between what a country imports and exports, reflect that 75.9% of what France imported and exported during 1988-91 was manufactured goods. Similar numbers also applied to other countries such as Canada, Austria, and the U.K. Hence, in contradiction to assumption of HO model, which predicts that France, for instance, should export manufactured goods and import agricultural goods, most of France's trade turned out to be two-way in similar goods or Intra-Industry Trade. Moreover, while HO model predicts little trade between nations with similar factor endowment, fact shows that biggest trade volumes are between nations that have similar factor endowments, or France trades more with Germany than with Vietnam. Since Grubel & Lloyd discovered such contradictions against neoclassical trade theory, economists then developed New Trade theory to explain intra-industry trade. Grubel & Lloyd thought the source of comparative advantage was increasing returns to scale (IRS). IRS explains why production of a particular good is concentrated in a single nation (i.e. France) rather than dispersed among all nations (i.e. Austria, Germany, and Italy). This, plus the broad similarity of tastes among rich nations explains Intra-Industry Trade (France produces and exports Peugeot to Italy while importing Italy's Ferrari). More formally, Dixit-Stiglitz (1977) developed models that explains why nations would each make some unique varieties to avoid direct competition and buys some of every variety available in the market so as to increase their utility of consumption.[6]

Key source of comparative advantage in this model is imperfect competition (firm has some market power or it can produce at a price that is greater than marginal cost). Goods have to have some unique qualities that cannot be perfectly substituted, which are called differentiated products, in order to be traded in the world of imperfect competition. Such unique quality may just be brand's name, different colors or symbols from competitor's products, etc. Consumers themselves have their own definition of unique quality or variety.

In this Dixit-Stiglitz model, if we take exporting country's income is a proxy for the range of varieties to be sold and the importing country's income as a proxy for demand for those varieties, then we can see that countries with higher GDPs should export and import more varieties. This model is more advanced than neo-classical models like Ricardian and HO models in the way that it suggests that income level should play a significant role in explaining trade flow of goods. Higher income countries would trade more. Also taking the importance of income level into account, Helpman-Krugman(1985) prove that volume of intra-industry trade is larger when countries have equal size.[7]

[6] Dixit-Stiglit model uses Constant Elasticity of Substitution (CES) function to derive the love of variety effect. The more varieties a country has (either through the capacity to produce or import), the higher utility it gains. So countries produce to export and import more and more varieties to increase their satisfaction of consumption.

Dixit-Stiglitz specification of CES function is to maximize Utility function

$$U = [\Sigma_{i=1}^{N} (C_i)^{1-\frac{1}{\sigma}}]^{\frac{1}{1-\frac{1}{\sigma}}} \text{ subject to a budget constraint } \Sigma_{i=1}^{N} p_i C_i = E \qquad (1)$$

C_i is consumption of good i, with i= 1, 2, ..., N. E is total expenditure.

σ is the elasticity of substitution and $\sigma > 1$ to make sense of monopolistic competition in differentiated products in which the marginal revenue p$(1 - \frac{1}{\sigma}) > 0$.

The higher σ is, the higher the degree of substitution. In other words, if σ is close to infinity, we have a world of perfect subsitutablity.

If all varieties are priced at p, common understanding is that comsumers would buy an equal amount of each variety.

Thus, comsumption of a typical variety is $\frac{E}{Np} = C$. Replacing this into Utility equation:

$$U = [\Sigma_{i=1}^{N} (\frac{E}{Np})^{1-\frac{1}{\sigma}}]^{\frac{1}{1-\frac{1}{\sigma}}} \qquad (2)$$

$$= [N(\frac{E}{Np})^{1-\frac{1}{\sigma}}]^{\frac{1}{1-\frac{1}{\sigma}}} \qquad (3)$$

$$= N^{\frac{1}{1-\frac{1}{\sigma}}} \frac{E}{Np} \qquad (4)$$

$$= N^{\frac{\sigma}{\sigma-1}} N^{-1} \frac{E}{p} \qquad (5)$$

$$= N^{\frac{1}{\sigma-1}} \frac{E}{p} \qquad (6)$$

So if N or the number of varieties increases, Utility will increase also.

[7] Define s is country Home's share of world resources. (1-s) is country Foreign's share of world resouces.

The role of income in determining trade was also taken into account in the empirical work of Timbergen (1962), Leamer and Stern (1970). These works assumed that nations produce their goods and throw them all together into a pile. Each nation would then draws its consumption out of the pile in proportion to their income. The higher the income, the higher the level of comsumption (or trade). These empirical works proposed some foundation for the popular gravity equation, which states that trade volume should be proportional to income and inversely proportional to distance between two nations.

However, they do not have clear microfoundations since we cannot define which trade theory stands behind it. It is only in 1979 when Anderson suggested that gravity equation should be built under assumption of differentiated goods. The most crystalized mathematical derivation of gravity equation in trade was then developed by Anderson and Wincoop in 2001 using Dixit-Stiglitz's CES functions.[8]

$$V_{od} = \frac{1}{\Omega_o P_d^{1-\sigma}} \frac{Y_o E_d}{\tau_{od}^{\sigma-1}} \tag{7}$$

or:

$$V_{od} = G \frac{Y_o Y_d}{(\text{distance})^{\sigma-1}} \left(G = \frac{1}{\Omega_o P_d^{1-\sigma}} = \frac{1}{\Sigma_{k=1}^{R} \tau_{ok}^{1-\sigma} \frac{E_k}{P_k^{1-\sigma}}} \frac{1}{P_d^{1-\sigma}} \right) \tag{8}$$

which says that the volume of export of differentiated products from country o (origin) to country d (destination) is proportional to the product of two countries' incomes Y_1, Y_2 and inversely proportional to distance or trade costs between them.

Y_o, E_d are country o's and d's total income/expenditure
τ_{od} : Trade cost between country o and country d
(k except o and d is the rest of the world (RoW))
τ_{ok} : Trade cost between country o and countries other than country d
E_k : Expenditure or income of countries other than o and d
$P_{d:}$ Perfect Price Index in country d
P_k : Perfect Price Index in countries other than o and d

The equation takes into account the importance of the multilateral resistance term, G, or the openness of the exporting country to all other countries different from country d. The term G is important because it signifies how country d

Hence, Foreign share of of all varieties is $(1-s)\,n^w$, where n^w is number of all varieties (differentiated products) existing in the world. Home's share of consumption of each variety is sY^w, where Y^w is total world income.

Thus, volume of trade of all differentiated products between Home and Foreign is:

$$VT = 2s(1-s)n^w Y^w$$

Maximizing above equation with respect to s, we find at $s = \frac{1}{2}$, or volume of Intra-Industry trade is greatest when Home and Foreign have same level of resources/income.

[8] Detailed derivation in Appendix B

is interacted with the rest of the world. G changes for each pair of trading partners. k except o and d is the rest of the world (RoW). If country i is far from RoW, or $\frac{\tau_{ok}}{P_k}$ is large, then ($\frac{\tau_{ok}}{P_k}$)$^{1-\sigma}$ is small, G is large, which mean that the exports from nation o to nation d will increase. In other words, if country o is "remote " from RoW, then country o would trade more with country d. Moreover, if RoW have high GDPs (large E_k,), then G is small, country o would trade less with country d (because RoW has higher demands for goods from country o).

Taking natural logarithm of equation (8) :

$$\ln V_{od} = \ln (Y_o E_d) - (\sigma - 1) \ln(\tau_{od}) - \ln(\Omega_o P_d^{1-\sigma}) \tag{9}$$

Based on data that we have on countries' GDPs and distance, we can run an econometric models as the following to assess whether the coefficients of income and distance follows the signs in equation (9).

$$\ln V_{od} = \alpha_0 + \alpha_1 \ln (Y_o E_d) + \alpha_2 (1 - \sigma) \ln(\tau_{od}) + \alpha_3 \ln(\Omega_o P^{1-\sigma}) + u_{od} \tag{10}$$

As it is complicated to obtain data on $\Omega_o P^{1-\sigma}$, Rose & Van Wincoop (2001) introduce an approach of using country-specific dummies so as to take into account the interaction among countries. This method is considered to be consistent by Anderson and Van Wincoop (2003). Baldwin & Taglioni (2006) suggests that country-specific dummies can only replace $\ln(\Omega_o P^{1-\sigma})$ in cross-sectional data. For panel data, country dummies cannot take into account time-series effect. Their specification of gravity equation for cross-sectional data using country-specific dummies suggested by Baldwin & Taglioni (2006) is:

$$\ln V_{od} = \alpha_0 + \alpha_1 \ln (Y_o E_d) + \alpha_2 (1 - \sigma) \ln(\tau_{od}) + \alpha_3 D_o + \alpha_4 D_d + u_{od} \tag{11}$$

where both the Ys and Es are incomes (GDPs) of exporting and importing countries, and the Ds are nation dummies.

One thing we need to keep in mind is that the derivation of gravity model is solely based on assumption of differentiated products. It does not tell us whether homogeneous goods would follow assumption of equation (8). Hence, we barely know whether homogeneous goods should be traded more if incomes of two trading partners get larger and closer to each other, or if distance between two countries is shorter. Also, the degree of product differentiation σ does not tell us much. If σ is large or there is high substitutability between goods produced either in o and d, the term (distance)$^{\sigma-1}$ will be large, thus reducing trade volume between two nations of interest. However, σ also appears in the term $\frac{1}{P_d^{1-\sigma}}$, which means that an increase in σ would lead to an increase in trade volume.

Nevertheless, Hummels and Levinsohn (1995) empirically showed that gravity equation works for both OECD and non-OCED countries (which export a high proportion of homogeneous primary goods). Also, Feenstra, Markusen and Rose (1998) explain that gravity equation can also be derived from a "reciprocal dumping" model of trade in homogeneous goods. In the reciprocal dumping model of Brander and Krugman (1983), one country may choose to sell at a lower price in its foreign market when the elasticity of demand in the foreign market is higher than at home. In other words, quantity sold in foreign market will increase by a lot if one firm takes into account the effect of transportation costs but does not raise the price of its product to a higher level than the price at home market. Intra-trade in homogeneous goods is possible in this way, which encourages Feenstra, Markusen and Rose (1998) to build a gravity equation based on this assumption.

Moreover, Feenstra, Markusen and Rose (2001) empirically find that the elasticity of exports with respect to own GDP of differentiated products is higher than of homogeneous products. Explanation is because homogeneous products have greater barriers to entry (resource constraint, seasonality, etc.), thus having lower home market effect. Home market effect reasons that when one country has higher level of income, the number of firms located there grows more rapidly than output, and the country becomes a net exporter of the good, despite the increase in domestic demand. Feenstra, Markusen and Rose (2001) thus imply that countries exporting homogeneous goods are less likely have number of firms increased, and less likely to increase their volume of homogeneous exports than of differentiated-goods when their incomes rise by the same level. As a result, we should expect to observe lower elasticity of exports with respect to own GDP for homogeneous products than for differentiated products.

The methodology that Rauch (1999) and Feenstra, Markusen and Rose (2001) used to assess the export flow of homogenous, referenced priced product groups is putting the assumption of differentiated products aside. Rauch (1999) derives a gravity equation that can be applied to all product groups. His derivation of gravity equation is very much like what was understood about gravity equation by Timbergen in 1960s, but takes into account the effect of "remoteness" or the term G in equation (8).

That is, we would consider a world of N countries where exporting country o and importing country d can take value o,d = 1, 2, ..., N.

Let Y_o denote total GDP of country o.

World GDP will then be:

$$Y_W = \sum_{1}^{N} Y_o \tag{12}$$

Call s_d country d's share of world expenditure. Thus:

$$s_d = \frac{Y_d}{Y_W} \tag{13}$$

So export value from country o to country d in case of no trade cost is:

$$v_{od} \quad = \quad s_d Y_o \tag{14}$$

Similarly, export value from country d to country o is $v_{do} = s_o Y_d$.

Thus, the geometric mean of trade volume between country o and country d in case of no trade cost would be:

$$\sqrt{v_{od} v_{do}} = \sqrt{s_d Y_o s_o Y_d} \tag{15}$$

$$= \sqrt{\frac{Y_d Y_o}{Y_W} \frac{Y_o Y_d}{Y_W}} \tag{16}$$

$$= \frac{Y_o Y_d}{Y_W} \tag{17}$$

In the presence of trade cost $\tau(\tau > 1$, viewed as an iceberg trade cost) due to distance between exported and imported countries, the mean of trade values between the two countries is:

$$V_{od} = \frac{1}{\tau} \frac{Y_o Y_d}{Y_W} \tag{18}$$

(V_{od} is the geometric mean of trade volume between country o and country d in the presence of trade cost.)

Since the purpose of this paper is to estimate gravity equations different product groups, we would have to run gravity equations separately for each product group as suggested by Rauch (1999). Let h =1, 2, 3 denotes different product groups with:

1 refers to product group that have goods traded on organized exchanges or homogenous goods

2 refers to product group that have goods not traded on organized exchanges but with some referenced prices

3 refers to product group that have goods not traded on organized exchanges or heterogeneous goods.

Export value from country o to country d for product group h in the presence of trade cost would be:

$$v_{odh} = \frac{1}{\tau} s_d w_{dh} Y_o \tag{19}$$

w_{nh} is commodity group h's share of country n's output. w_{nh} may vary for different countries due to comparative advantage.

The geometric mean of trade volume between two countries o, d would be:

$$V_{odh} = \sqrt{v_{odh} v_{doh}} \tag{20}$$

$$V_{odh} = \sqrt{(\frac{1}{\tau} s_d w_{dh} Y_o)(\frac{1}{\tau} s_o w_{oh} Y_d)} \quad (21)$$

$$V_{odh} = \sqrt{w_{oh} w_{dh}} \frac{1}{\tau} \frac{Y_o Y_d}{Y_W} \quad (22)$$

Rauch assumes that the term $\sqrt{w_{oh} w_{dh}}$ in the equation above is absorbed into the multiplicative error term.

Equation (22) states that trade volume of commodity group h between country o and country d is proportional to the product of GDPs and the share of expenditure spent on that group, and disproportional to trade cost between the two countries.

In logarithm form, the equation can be written:

$$\ln V_{odh} = \ln (Y_o Y_d / Y_W) - \ln (\tau) + \ln (\sqrt{w_{oh} w_{dh}}) \quad (23)$$

In econometric form:

$$\ln V_{odh} = \alpha_h + \beta_h \ln(Y_o Y_d) + \gamma_h \ln(\tau) + u_{odh} \quad (24)$$

There will be 3 separate gravity equations for 3 product groups ($h = 1, 2, 3$). Gravity equation would assume that β_h has a positive sign while γ_h has a negative sign. If running equation (24) using country fixed effect, it would yield exactly the same result as equation(11) for differentiated products, that is derived by Anderson and Wincoop (2001). Or using fixed effect estimation, equation (24) also takes into account the effect of the "remoteness" term G in equation (8).

In terms of estimation of equation (24), Rauch (1999) and Feenstra, Markusen and Rose (2001) sum the bilateral exports of all goods at 4-digit SITC level into the categories of homogeneous, reference priced and differentiated goods for each country. Then they run the regression with the left hand side (LHS) variable at this aggregated level of bilateral exports. Problem with this method is that when these authors aggregate all homogeneous goods into one number for bilateral exports, that number does not represent each homogeneous good any more. Instead, it represents the sum of all different homogeneous goods, which is similar to the flow of one differentiated product. These aggregating problem might be severe for the case of homogeneous goods because the response of each homogeneous good to changes in explanatory variables can be significantly different from the response of the sum of all homogeneous goods together.

IV: Methodology

In this paper, I would also run gravity equation (24) for three product groups, but the level of observation is at 4-digit SITC goods. Each observation in my equation is not the sum of each product group's bilateral exports, but is the trade value of each product at 4-digit SITC level itself.

My specification would be:

$$\begin{aligned}\ln V_{odgh} &= \alpha_{0h} + \alpha_{1h}\ln(Y_oY_d) + \alpha_{2h}\ln(\tau) + \\ &\quad \alpha_{3h}Contig + \alpha_{4h}ComLange + \\ &\quad \alpha_{5h}Colony + \lambda_{oh}D_o + \lambda_{dh}D_d + u_{odgh} \qquad (25)\end{aligned}$$

with h : product group ($h = 1, 2,$or 3)

$\ln V_{odgh}$: natural logarithm of exports of good g from country o to d

$\ln(Y_oY_d)$: natural logarithm of product of GDPs of country o and country d

$\ln(\tau)$: Natural logarithm of distance between country o and country d

$Contig$: dummy equal to 1 if country o share a common border with d

$ComLange$: dummy equal to 1 if country o and d share a common official language

$Colony$: dummy equal to 1 if country o and country d have colonial relationship

$D_o,\ D_d$: country specific dummy.

The reason that many economists add other variables such as common border, common language and colonial relationship to gravity equation is because they believe those variables positively affect exports of goods. Here, I would like to assess whether these different measures of distance affect exports of all product groups the same way. Common language and colonial relationship are two social distance measures that signify how socially close the two trading partners are. If two partners share a common language or a colonial relationship, then communication between the two countries should be relatively higher than with other countries. Communication is important because it helps to reduce information asymmetry, thus allows consumers in one country to understand more about the reputation or quality of a product that they are going to buy. Communication may matter differently for different product groups. In particular, communication is more necessary when a product is heterogeneous. That product has so many functions, and different quality that consumers are often puzzled of which type to buy. Should an Algerian buys a Samsonite handbag imported from its U.S. manufacturer or buy one named Tous from France. Since the customer understand French, she can always read reviews of the product and see that the price of Tous is cheaper and its design is more suitable for her. Her social connection with France plays a significant role in determining whether she chooses Tous over Samsonite. However, if the product is corn instead of handbag, social distance measure may not matter at all. Corn is a homogeneous good, so a Vietnamese customer would know already how the product is even though he does not speak a common language with the corn exporting countries. Speaking a common language with foreign exporters or not does not help him in deciding from whom he should buy the homogeneous product, corn.

Hence, we can see that homogeneous goods are somewhat information self-revealing, while heterogeneous goods are information-asymmetric. As a result, having a social connection such as speaking a common language and having a colonial relationship would help importers of heterogeneous goods to understand more about quality and reputation of different brands of a product. Thus, the importers tend to import more from the exporters whom they know better thanks to the social link between them. On the supplying side, exporters are likely to export more to countries with which they have social link, so as to reduce communication costs such as advertising and marketing costs. Differentiated products have much more need to be advertised than homogeneous products.

Data Description and Estimation

Data on exports of all products at SITC-4 digit level is available from the United Nations' COMTRADE database. These data will then be merged with Rauch's classification to determine which good at SITC-4 digit level belongs to which product type. Rauch has two classifications of goods, which is conservative and liberal. Following Feenstra, Markusen and Rose (2001), I use the more conventional classification, or the conservative. Data on GDPs is obtained from the United Nations Statistics Division. Regarding data on distance, CERII's database provides complete information. The CERII have different measures of geographical distance such as simple geographical distance between two countries' capitals using Great Circle formula, and weighted distance between two countries' economic centers. From my experience of trying both distance measures, it does not create much difference as to whether I use the simple or the weighted distance measure. Thus, I then choose to use the simple geographical distance variable because it has more complete number of observations.

My sample size has 66 countries trading with each other over 745 products in the year 2000. Among 745 products, 45 (6%) of them were not classified by Rauch in 1999, partly because there were some new products traded in 2000. This 6% of unclassified products would not be considered in the final dataset. List of countries is in the Appendix. 66 countries in the data account for 83% of total world trade.

There are totally 36,157 observations reported for homogeneous goods, 169,367 observations for reference priced and 502,955 observations for heterogeneous goods.

First I would run equation (25) for all observations that are reported using fixed effect estimation at country level. The estimation based on this sample, though normally distributed, is biased because it does not take into account the zero trade values which are not reported by countries. So I will then code any unreported trade value between any pair of countries as zeros and use fixed effect quantile regression developed by Koenker (2004) for longitudinal data. A median fixed effect quantile regression is, in fact, a fixed effect Tobit regression developed by Bo Honoré (1992) for data whose observations concentrate above zeros.

Country fixed effect estimation will be carried out using "within" transformation. Considering the fact of huge dataset for reference priced and heterogeneous goods, the usual transformation of pre-multiplying both sides of equation

(25) by the annihilator matrices is not possible, since the dimension of the annihilator matrices would be too large.[9] Thus, my "within" estimation will be implemented by first taking deviation of all variables from country o's mean, then from country d's mean.

Result

Table 1 reports fixed effect estimation for dataset with trade values being greater than zero. All standard errors are first corrected for the degree of freedom for the number of group's means that are used to take deviation from country's means transformation. There are 66 exporting and 66 importing countries, thus the total number of degree of freedom to be corrected is 132. The standard errors are also then clustered by a country-pair variable to correct for any contemporaneous correlation due to common characteristics of products traded between each pair of countries.

The estimation reveals that coefficients of the product of GDPs and simple distance for homogeneous goods are smaller than that for referenced price and heterogeneous goods. Furthermore, Figure 2 reveals that the effect of border, common language and colonial relationship on the trade flow of goods also seems to increase as goods move from homogeneous to heterogeneous. At 5% significant level, social distance variables (common language and colonial relationship) do not have any effect on exports of homogeneous goods. However, they all have significant impact on increase the volume of exports of heterogeneous goods.

[9] In matrix notation, we can remove the effect of D_o, D_d from equation below:

$$\ln V_{odgh} = \alpha_{0h} + \alpha_{1h} \ln (Y_o Y_d) + \alpha_{2h} \ln (\tau) + \lambda_{oh} D_o + \lambda_{dh} D_d + u_{odgh} \tag{26}$$

by firstly by pre-multiplying both sides of equation (26) by M_o or the annihilator matrix of D_o.

Remember that: $M_o D_o = 0$, but $M_o Y$ (with Y is any particular matrix) creates the residuals matrix of Y, or the deviation of Y from its country o's mean.

Thus, after pre-mutiplying both side by M_o, equation (26) becomes:

$$M_o \ln V_{odgh} = \alpha_{1h} M_o \ln (Y_o Y_d) + \alpha_{2h} M_o \ln (\tau) + \lambda_{dh} M_o D_d + M_o u_{odgh} \tag{27}$$

$M_o D_d$ creates the residual matrix of D_d, or the deviation from country o's mean for the dummy matrix of destination country.

To remove the effect of $M_o D_d$, I create another annihilator matrix, which is $M_d = (I - M_o D_d[(M_o D_d)'(M_o D_d)]^{-1}(M_o D_d)'$.

This annihilator matrix creates the deviation (of the deviation from country o's mean) from country d's mean.

$$M_d M_o \ln V_{odgh} = \alpha_{1h} M_d M_o \ln (Y_o Y_d) + \alpha_{2h} M_d M_o \ln (\tau) + M_d M_o u_{odgh}$$

"Within" Estimation Result			
	Homogeneous	Reference	Heterogeneous
lnGDPDev	0.420 **	0.526 **	0.596 **
	(0.087)	(0.055)	(0.047)
lnDistanceDev	-0.470 **	-0.659 **	-0.741 **
	(0.036)	(0.024)	(0.031)
ContigDev	0.624 **	0.632 **	0.753 **
	(0.154)	(0.117)	(0.130)
ComLangeDev	-0.005	0.059	0.272 **
	(0.085)	(0.055)	(0.059)
ColonyDev	-0.108	0.077	0.329 **
	(0.108)	(0.078)	(0.100)
Adjusted R-Squared	0.0314	0.0578	0.0808
N	36157	169367	502955

** Significant at one percent level. *Significant at five percent level
This dataset does not include zero export values.
Adjusted standard errors are reported in parenthesis.

Table 1: Country Fixed Effect Estimation

Residuals obtained from above regression are normal distributed.

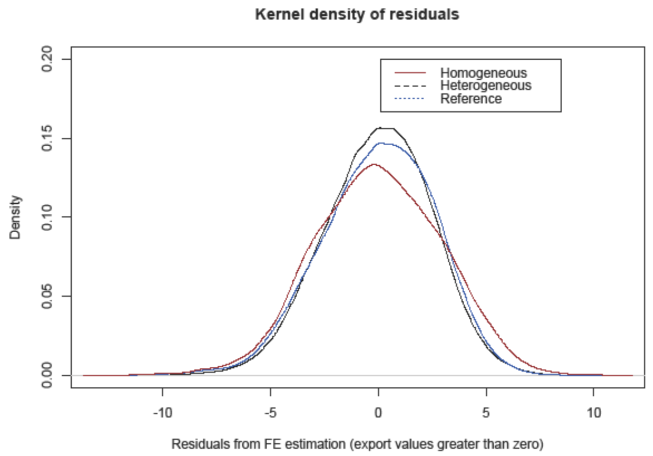


Figure 1: Residuals obtained from fixed effect regression

However, this regression result is based on a biased trade dataset because unreported/missing trade values are not taken into account. Countries only report how much they export when they do export to a particular country. If one country does not export any to another partner, the export value between these two should be coded as zeros. Then we can evaluate how explanatory variables impact the decision of a country on whether it should export to a certain country or not. Thus, all unreported trade values in the original dataset are coded as zeros. Problem with the new dataset which includes zero export values is that the response variable is not normally distributed but concentrated above zero.

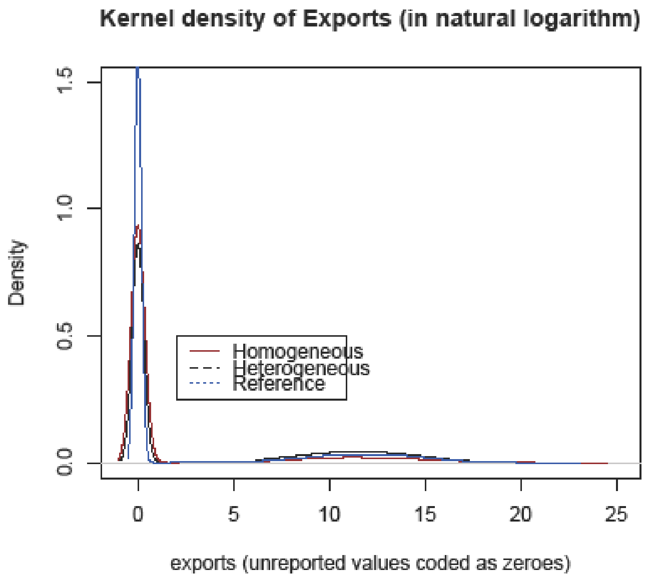


Figure 2: Distribution of exports (zero trade values included)

Since trade values in the data is bounded by zeros, a fixed effect Tobit regression should to be applied for this truncated dataset. A method for estimating a fixed effect Tobit regression using trimmed least-absolute deviation (LAD) estimation was developed by Bo Honoré in 1992. This LAD estimation is, in fact, a median quantile regression which minimizes the least absolute deviation of the median, instead of the mean, of the response variable for a given level of predictor variables. Thus, if using fixed effect median quantile regression, we are in fact estimating a fixed effect Tobit regression. More conveniently, we can always apply quantile regression for quantiles other than 0.50th. Quantile estimation at levels other than the median is defined as a problem that minimizes the weighted sum of the absolute value of the residuals.

$$\hat{\beta}_\tau = argmin \left\{ \sum_{i:y_i \geq x_i'\beta_\tau} \tau|y_i - x_i'\beta_\tau| + \sum_{i:y_i < x_i'\beta_\tau} (1-\tau)\,|y_{i-}x_i'\beta_\tau| \right\} \tag{28}$$

When τ is higher than .50th, it means that we put more emphasis on the upper quantile (where $y_i \geq x_i'\beta_\tau$) than the lower quantile ($y_i < x_i'\beta_\tau$). I will use fixed effect quantile regression for .50th, .60th, .70th, .80th and .90th quantiles. The reason for a starting point of .50th quantile is because more than half of trade values in the dataset are coded as zeros. Zero trade values are comprised of 83% of homogeneous goods, 75% of reference priced goods, and 66% of heterogeneous goods. As a consequence, I will particularly focus on the regression result for the .80th and .90th quantile for homogeneous goods, 0.7th, 0.8th and 0.9 for reference priced goods, and 0.6th, 0.7th, 0.8th, 0.9th quantile for heterogeneous good. Coefficients at median (50th percentile), 60th or 70th percentile in the regression for homogeneous goods can tell us whether they affect the decision of a country to export this type of goods or not. However, coefficients at .80th and 0.90th quantile (where trade values are greater than zero) can tell us a more complete story, that is we will be able to know to whom within its already trading partners a country exports more.

Koenker (2004) states that fixed effects at all quantile levels should be penalized to be the same. That is to say, for datasets used in this paper, country fixed effect ($\lambda_{oh}, \lambda_{dh}$) should be the same at all 0.50th, 0.60th, 0.70th, 0.80th and 0.90th quantiles. Tables below show quantile regression result with bootstraped standard errors in parenthesis for all three product groups. Standard errors are bootstrapped because there are concern about heterogeneity at country level.

Fixed Effect Quantile Regression for Homogeneous Goods

Quantile	0.5	0.6	0.7	0.8	0.9
lnGDPDev	0.307 **	0.307 **	0.307 **	0.621 **	1.277 **
	(0.056)	(0.056)	(0.056)	(0.028)	(0.191)
lnDistanceDev	-0.097	-0.098	-0.099	-1.781 **	-2.583 **
	(0.380)	(0.380)	(0.380)	(0.079)	(0.252)
ContigDev	7.840 **	9.671 **	10.832 **	6.618 **	2.759 **
	(1.373)	(1.833)	(2.126)	(1.077)	(0.144)
ComLangeDev	0.011	0.011	0.011	0.352 **	2.475 **
	(0.046)	(0.046)	(0.046)	(0.064)	(0.591)
ColonyDev	0.048	0.049	3.171 **	3.158 **	-0.086
	(0.192)	(0.192)	(0.662)	(0.625)	(0.307)
N			216125		

** Significant at one percent level. *Significant at five percent level
Coefficients and standard errors are bootstrapped within each country-pair cluster.
Bootstrap is performed with 200 replications.

Table 2: Penalized Fixed Effect Quantile Regression for Homogeneous Goods

Fixed Effect Quantile Regression for Reference-priced Goods

Quantile	0.5	0.6	0.7	0.8	0.9
lnGDPDev	0.806 **	0.856 **	1.064 **	1.455 **	1.835 **
	(0.014)	(0.009)	(0.066)	(0.181)	(0.293)
lnDistanceDev	-0.778 *	-0.933 **	-1.675 **	-2.435 **	-2.591 **
	(0.335)	(0.290)	(0.077)	(0.158)	(0.202)
ContigDev	6.164 **	6.790 **	5.619 **	3.408 **	2.066 **
	(1.185)	(1.371)	(1.027)	(0.377)	(0.066)
ComLangeDev	-0.064	-0.006	0.256 **	1.692 **	2.571 **
	(0.142)	(0.125)	(0.048)	(0.382)	(0.640)
ColonyDev	1.396 **	2.880 **	2.934 **	0.808 **	-1.232
	(0.137)	(0.511)	(0.525)	(0.131)	(0.715)
N			672490		

** Significant at one percent level. *Significant at five percent level

Coefficients and standard errors are bootstrapped within each country-pair cluster.

Bootstrap is performed with 200 replications.

Table 3: Penalized Fixed Effect Quantile Regression for Reference-Priced Goods

Fixed Effect Quantile Regression for Heterogeneous Goods

Quantile	0.5	0.6	0.7	0.8	0.9
lnGDPDev	1.193 **	1.363 **	1.581 **	1.762 **	1.899 **
	(0.019)	(0.083)	(0.168)	(0.238)	(0.291)
lnDistanceDev	-1.544 **	-1.822 **	-2.108 **	-2.311 **	-2.263 **
	(0.106)	(0.013)	(0.114)	(0.192)	(0.174)
ContigDev	3.168 **	2.974 **	2.579 **	2.110 **	1.836 **
	(0.588)	(0.514)	(0.362)	(0.181)	(0.082)
ComLangeDev	-0.076	0.173	0.436 **	1.257 **	1.529 **
	(0.278)	(0.183)	(0.082)	(0.241)	(0.345)
ColonyDev	2.805 **	2.588 **	2.019 **	0.964 **	0.005
	(0.565)	(0.481)	(0.261)	(0.154)	(0.523)
N			1460875		

** Significant at one percent level. *Significant at five percent level

Coefficients and standard errors are bootstrapped within each country-pair cluster.

Bootstrap is performed with 200 replications.

Table 4: Penalized Fixed Effect Quantile Regression for Heterogeneous Goods

From regression results, we can see that income level positively affects exports of homogeneous goods. Thus, homogeneous goods are clearly traded more as income level rises, just as heterogeneous goods. However, the income elasticity of homogeneous goods with respect to the product of two trading partners' income is significantly smaller than that of referenced priced and heterogeneous goods in all quantiles. Engle's law and all characteristics of homogeneous mentioned in Section II helps explain the regression result.

The effect of different distance measures on the exports of different product groups are interesting. Firstly, simple distance between two countries does not have any impact on the exports of homogeneous goods in either .50th, .60th, or .70th quantile. Keeping in mind that zero trade values account for 83% in the sample for homogeneous goods, which means there is no export in 83% of the sample. If we put emphasis on export values from 80th percentile onwards, then the regression result in Table 2 shows that simple distance has negative impact only on exports which actually take place. As a result, we can interpret that simple distance does not play a significant role on whether a country would export homogeneous goods to another or not. However, once a country exports, then within all partners to which the country exports, the country would export less to those which are farther away. Different from homogeneous goods, both referenced priced and heterogeneous goods are negatively correlated with simple distance at the median quantile, and all other higher quantiles.

The effects of the other geographical distance measure, common border, are strong on exports of all product types at all quantile levels. Hence, having a common border positively increases the possibility of countries to exports, and significantly impacts to which partner the country should export more. More intuitively, the effect of having a common border is much stronger on exports of homogeneous goods than that of reference-priced and heterogeneous goods. That is because most of homogeneous goods are food products which usually cannot travel for a long time, and fuels and minnings which are large in volume and size, thus cost much more if transportation distance increases. Common border matters the least to the exports of heterogeneous goods. This result contradicts the OLS fixed effect regression in Table 1.

Regarding the two social distance measures, median quantile estimation suggests that common language does not have impact on exports of any product type, while colonial relationship significantly influences export of reference-priced and heterogeneous goods. At higher levels of quantiles when higher weights are given to export values that are greater than zeros, then we can see the impact of common language on all product types, and even strongest on exports of homogeneous goods at 90th percentile. That is to say within the group of all partners that a country exports to, common language plays a significant role in differentiating which partners the country would export more to. Examining the data on homogeneous goods carefully, it turns out that export volumes at the 0.90th quantile level are of 5 products: crude oil, gold, aluminum, motor spirit and other oil, and meat of bovine animals. This high level of exports volume for those products are due to trade between the U.S. and Canada (crude oil, meat, aluminum, motor spirit), Nigeria and India (crude oil), the U.S. and the U.K. (gold, crude oil), and between other oil exporting countries (Saudi Arabia, Kuwait, Iran, Norway) to highly developed countries like the U.S., Canada, the U.K., and Japan. One-third of all countries whose export volume are at 0.90th quantile are English speaking countries. These countries are normally regarded as the economic and political powers of the world. Their economies dominate world trade in homogeneous goods and their language, English, plays a very strong role in explaining such high volume of trade among them.

These countries trade also at 0.90th quantile level in heterogeneous products. Heterogeneous products being traded at this high level of quantile are mostly electrical equipments and machineries, as well as motor vehicles of all kinds. There are also participation of other English speaking countries in this level of trading such as Hong Kong, Singapore. However, the participation of many other non-English speaking countries makes the impact of common language less important for exports of heterogeneous goods. Those non-English speaking exporters of heterogenous goods at 0.90th quantile are China, Japan, Korea, Malaysia, and Mexico. Nevertheless, the effect of common language on exports of heterogeneous goods are still highly significant.

At the median, having colonial relationship increases the possibility to export referenced priced and heterogeneous goods of countries, but not homogeneous goods. Only at 0.70th, and 0.80th quantiles, colonial relationship starts to have strong impact on exports of homogeneous goods. Remembering that 0.83 is the threshold between no export or export of homogeneous goods. Thus, the strong impact of colonial relationship shown around this level of quantiles is somewhat understandable. At 90th percentile, all good types are no longer affected by colonial relationship. It is because at such high volume of trade, there are only economic powers trading with each other. Very few of them trade extensively with their former colonies. Those are Japan and Korea, the US and the Philippines, the US and the UK.

From this result, we can see that economic superpowers dominate world trade in all three types of goods. Many of them are English speaking countries, thus we perceive a high correlation between language and high trade volume. Also, countries having highest volume of trade (above 90 percentile) between each other do not usually have colonial relationship.

Robust Check: Adding product fixed effect

Quantile estimation reveals that the coefficients of each explanatory variable at different quantiles are very different from each other. For instance, the effect of border on exports of homogenegous goods is very high at the median and 0.60th, 0.70th quantiles. However, it then becomes much smaller at the 0.80th and 0.90th quantiles. This variation then suggests that there might be a relatively significant level of heterogeneity, which drives the coefficients to vary at different quantiles. The problem might be that when choosing the estimation method at product level, this paper is imposing the coefficients (of each explanatory variable) of all products within one group to be the same. However, the responsiveness of each product to changes in income and different measures of distance may be very different from each other. For instance, some products are inferior goods, thus respond negatively to an increase in income, while most other products are normal goods. Also, the response of crude oil to changes in simple distance may different much from that of corn. To confirm such reasoning, I then run 700 gravity equations for 700 individual products and obtaining the coefficients from the equations. Graphs of these coefficients are shown in Appendix E. We can see that there is large variation among the coefficients of each variable in each product group, especially of homogeneous and reference-priced goods. Many homogeneous and referenced priced goods have largely positive distance coefficients, while most have negative coeffients

Thus, unobserved effect at product level should be taken into account. Rerunning the "within estimation" for three product groups with country and product fixed effect, I obtain the results as in Table:

Country and Product Fixed Effect Estimation

	Homogeneous	Reference	Heterogeneous
lnGDPDev	0.756 **	0.910 **	1.107 **
	(0.093)	(0.063)	(0.057)
lnDistanceDev	-0.508 **	-0.784 **	-0.867 **
	(0.038)	(0.027)	(0.034)
ContigDev	0.728 **	0.689 **	0.807 **
	(0.168)	(0.127)	(0.140)
ComLangeDev	0.040	0.086	0.335 **
	(0.088)	(0.060)	(0.065)
ColonyDev	-0.063	0.162	0.423 **
	(0.116)	(0.087)	(0.111)
Adjusted R-Squared	0.0440	0.0880	0.1281
N	36157	169367	502955

** Significant at one percent level. *Significant at five percent level

This dataset does not include zero export values.

Adjusted standard errors are reported in parenthesis.

Table 5: Fixed Effect Regression at Country and Product Level

This table still suggest that homogeneous goods are less responsive to changes in income. Geographical distance have significant impact on all types of goods, while cultural distance only significantly affect exports of heterogeneous goods.

Conclusion

In conclusion, this research finds that homogeneous goods are less responsive to changes in income than heterogenous goods. I also find that export volume of all product types is significantly hindered by geographical distance between countries. Fixed effect "within" estimation suggests that the two social distance measures, common language and colonial relationship, do not have any impact on homogeneous goods while they positively affect exports of heterogeneous goods. Whereas, fixed effect Tobit or median quantile regression suggests that only colonial relationship improves exports of heterogeneous goods, not common language.

At higher levels of quantiles (0.70th, 0.80th) when higher weights are given to export values which are greater than zeros, we then see the impact of common language and colonial relationship on exports of homogeneous goods. That is to say, common language and colonial relationship play a significant role in differentiating among partners in the group that one country is exporting to. Countries would export more to those that speak the same language and have colonial relationship. At the highest level of quantile, 0.90th, there are mainly world economic powers trading with each others. Most of them are English speaking countries, thus common language is strongly correlated with exports between countries. Income, distance and border effect are also very significant. However, colonial relationship loses all its impact at this level of high trade volume.

References

[1] Anderson, James E, 1979. "A Theoretical Foundation for the Gravity Equation," American Economic Review, American Economic Association, vol. 69(1), pages 106-16, March.

[2] Baldwin, Richard & Taglioni, Daria, 2006. "Gravity for Dummies and Dummies for Gravity Equations," CEPR Discussion Papers 5850, C.E.P.R. Discussion Papers.

[3] Deaton, A. (1997): The Analysis of Household Surveys. A Microeconomic Approach to Development Policy. Johns Hopkins University Press, Baltimore, MD

[4] Feenstra, Robert & Markusen, James R. & Rose, Andrew K, 1998. "Understanding the Home Market Effect and the Gravity Equation: The Role of Differentiating Goods," CEPR Discussion Papers 2035, C.E.P.R. Discussion Papers.

[5] Feenstra, Robert & Markusen, James R. & Rose, Andrew K, 2001. "Using the gravity equation to differentiate among alternative theories of trade," Canadian Journal of Economics, Canadian Economics Association, vol. 34(2), pages 430-447, May.

[6] Glick, Reuven & Rose, Andrew K., 2002. "Does a currency union affect trade? The time-series evidence," European Economic Review, Elsevier, vol. 46(6), pages 1125-1151, June.

[7] Honore, Bo E, 1992. "Trimmed LAD and Least Squares Estimation of Truncated and Censored Regression Models with Fixed Effects," Econometrica, Econometric Society, vol. 60(3), pages 533-65, May

[8] James E. Anderson & Eric van Wincoop, 2000. "Gravity with Gravitas: A Solution to the Border Puzzle," Boston College Working Papers in Economics 485, Boston College Department of Economics.

[9] Koenker, R., and G. J. Bassett (1978): "Regression Quantiles," Econometrica, 46(1), 33–50.

[10] Koenker, Roger, 2004. "Quantile regression for longitudinal data," Journal of Multivariate Analysis, Elsevier, vol. 91(1), pages 74-89, October.

[11] OECD Economic Outlook 2002, Chapter VI: Intra-Industry and Intra-Firm Trade and the Internationalization of Production.

[12] Rauch, James E., 1999. "Networks versus markets in international trade," Journal of International Economics, Elsevier, vol. 48(1), pages 7-35, June.

[13] Wooldridge, J. M. (2002): Econometric Analysis of Cross Section and Panel Data. MIT Press.

Appendix A

List of Countries

Algeria
Argentina
Australia
Austria
Belgium
Bolivia
Brazil
Canada
Chile
China
China, Hong Kong SAR
China, Macao SAR
Colombia
Denmark
Ecuador
Egypt
Ethiopia
Finland
France
Ghana
Greece
Hungary
Iceland
India
Indonesia
Iran
Ireland
Israel
Italy
Japan
Kenya
Kuwait
Luxembourg
Malaysia
Mexico
Morocco
Netherlands
New Zealand
Nigeria
Norway
Pakistan
Paraguay
Peru
Philippines
Poland
Portugal
Qatar
Rep. of Korea
Russian Federation
Saudi Arabia
Singapore
South Africa
Spain
Sudan
Sweden
Switzerland
Thailand
Tunisia
Turkey
USA
Ukraine
United Kingdom
Uruguay
Venezuela
Viet Nam
Zimbabwe

Appendix B

Derivation of Gravity Equation by Anderson and Wincoop (2001)
(Notations in this derivation closely follow Baldwin and Taglioni (2005))

Anderson and Wincoop (2001) model assumes that each country produces a unique variety of goods. It starts with an expenditure share equation:

$$p_{od}C_{od} = s_{od}E_d \quad (29)$$

On the RHS is the expenditure of country d (destination) on the good imported from country o (origin), which is equal price p_{od} times quantity C_{od}.

C_{od} is export quantity of a single variety from country o to country d.

s_{od} is the share of expenditure in country d on a typical variety made in country o. E_d is country d's total expenditure.

Remember from Dixit-Stiglitz's problem of maximize the utility function:

$$U = [\Sigma_{i=1}^{N}(C_i)^{1-\frac{1}{\sigma}}]^{\frac{1}{1-\frac{1}{\sigma}}} \text{ subject to a budget constraint } \quad \Sigma_{i=1}^{N}p_iC_i = E \quad (30)$$

Using Lagrangian Multiplier and maximizing the following function with respect to consumption of a particular variety j C_j,

$$\text{L} = [\Sigma_{i=1}^{N}(C_i)^{1-\frac{1}{\sigma}}]^{\frac{1}{1-\frac{1}{\sigma}}} + \lambda(E - \Sigma_{i=1}^{N}p_iC_i) \quad (31)$$

We can derive the following result for direct demand function:

$$p_j = \frac{(C_j)^{-\frac{1}{\sigma}}}{\Sigma_{i=1}^{N}(C_i)^{1-\frac{1}{\sigma}}}E \quad (32)$$

And the indirect demand function as:

$$C_j = \frac{p_{j-\sigma}}{\Sigma_{i=1}^{N}p_i^{1-\sigma}}E \quad (33)$$

Multiply by p_j,

$$p_jC_j = \frac{p_j^{1-\sigma}}{\Sigma_{i=1}^{N}(p_i)^{1-\sigma}}E \quad (34)$$

Or:

$$\frac{p_jC_j}{E} = (\frac{p_j}{\Sigma_{i=1}^{N}p_i})^{1-\sigma} \quad (35)$$

This is the expenditure share on good j out of total expenditure of one country. It implies that expenditure share depend upon relative price and income level.

Returning to equation (29), assuming that all goods are traded, we can apply equation (35) to derive the expenditure share that nation d spend on a typical variety made in nation d as:

$$s_{od} = (\frac{p_{od}}{P_d})^{1-\sigma} \qquad with \qquad P_d = (\Sigma_{r=1}^{R} n_{r(p_{rd})^{1-\sigma}})^{\frac{1}{1-\sigma}} \tag{36}$$

$\frac{p_{od}}{P_d}$ is the real price of p_{od}.P_d is nation d 's ideal CES price index. R is the number of nations from which nation d buy things. n_r is the number of varieties exported from nation r.

Taking into account the effect of trade cost τ_{od}, and bilateral mark-up μ in this world of imperfect competition, product produced in country o landed in country d with a price:

$$p_{od} = \mu p_o \tau_{od} \tag{37}$$

Assuming $\mu = 1$, aggregate across all individual goods we have:
$V_{od} = s_{od} E_d$

$$V_{od} = (\frac{p_{od}}{P_d})^{1-\sigma} E_d = n_o (\frac{p_o \tau_{od}}{P_d})^{1-\sigma} E_d \tag{38}$$

In reality, we do not have complete data on n_o and p_o.

But we can use the assumption that income of country o will be equal to sale from o to all other nations:

$$Y_o = \Sigma_{d=1}^{R} V_{od} \tag{39}$$

Thus,

$$Y_o = \Sigma_{d=1}^{R} n_o (\frac{p_o \tau_{od}}{P_d})^{1-\sigma} E_d = n_o p_o^{1-\sigma} (\Sigma_{d=1}^{R} \tau_{od}^{1-\sigma} \frac{E_d}{P_d^{1-\sigma}}) \tag{40}$$

From (40),

$$n_o p_o^{1-\sigma} = \frac{Y_o}{\Sigma_{k=1}^{R} \tau_{ok}^{1-\sigma} \frac{E_k}{P_k^{1-\sigma}}} \tag{41}$$

(k except i and j is the rest of the world (RoW))

So we have successfully express $n_o p_o^{1-\sigma}$ as a function of what we can observe:
Y_o: Income of country o
τ_{ok} : Trade cost between country o and countries other than country d
E_k : Expenditure or income of countries other than o and d
P_k : Perfect Price Index in countries other than o and d
Plug in the result from (41) into (38) :

$$V_{od} = n_o p_o^{1-\sigma} \tau_{od}^{1-\sigma} \frac{E_d}{P_d^{1-\sigma}}$$

$$V_{od} = \frac{Y_o}{\Sigma_{k=1}^{R} \tau_{ok}^{1-\sigma} \frac{E_k}{P_k^{1-\sigma}}} \tau_{od}^{1-\sigma} \frac{E_d}{P_d^{1-\sigma}} \tag{42}$$

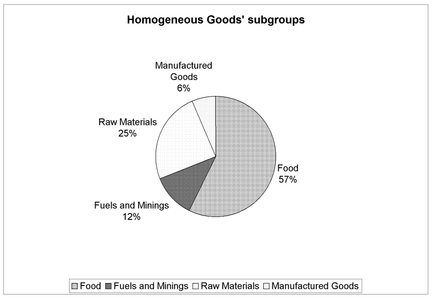
$$V_{od} = \frac{Y_o}{\Omega_o} \tau_{od}^{1-\sigma} \frac{E_d}{P_d^{1-\sigma}} \tag{43}$$

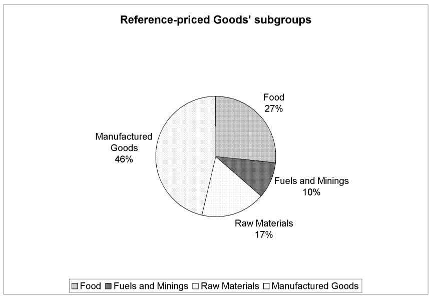
$$V_{od} = \frac{1}{\Omega_o P_d^{1-\sigma}} \frac{Y_o E_d}{\tau_{od}^{\sigma-1}} \tag{44}$$

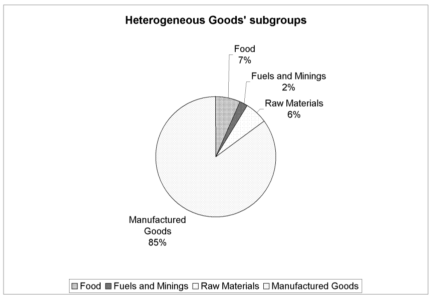
$$V_{od} = G \frac{Y_o Y_d}{(\text{distance})^{\sigma-1}} \left(G = \frac{1}{\Omega_o P_d^{1-\sigma}} = \frac{1}{\Sigma_{k=1}^{R} \tau_{ok}^{1-\sigma} \frac{E_k}{P_k^{1-\sigma}}} \frac{1}{P_d^{1-\sigma}} \right) \tag{45}$$

Appendix C

Product Types' Subgroups in Sample of 700 Products from COMTRADE



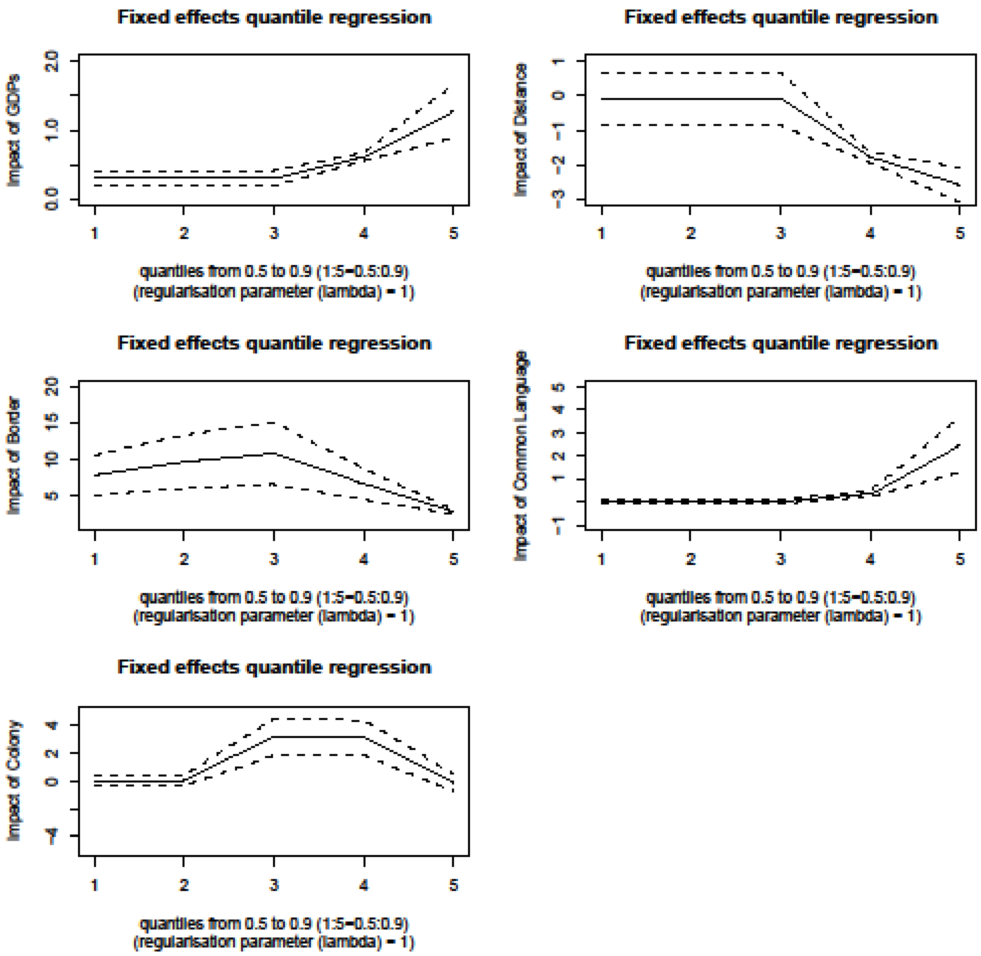


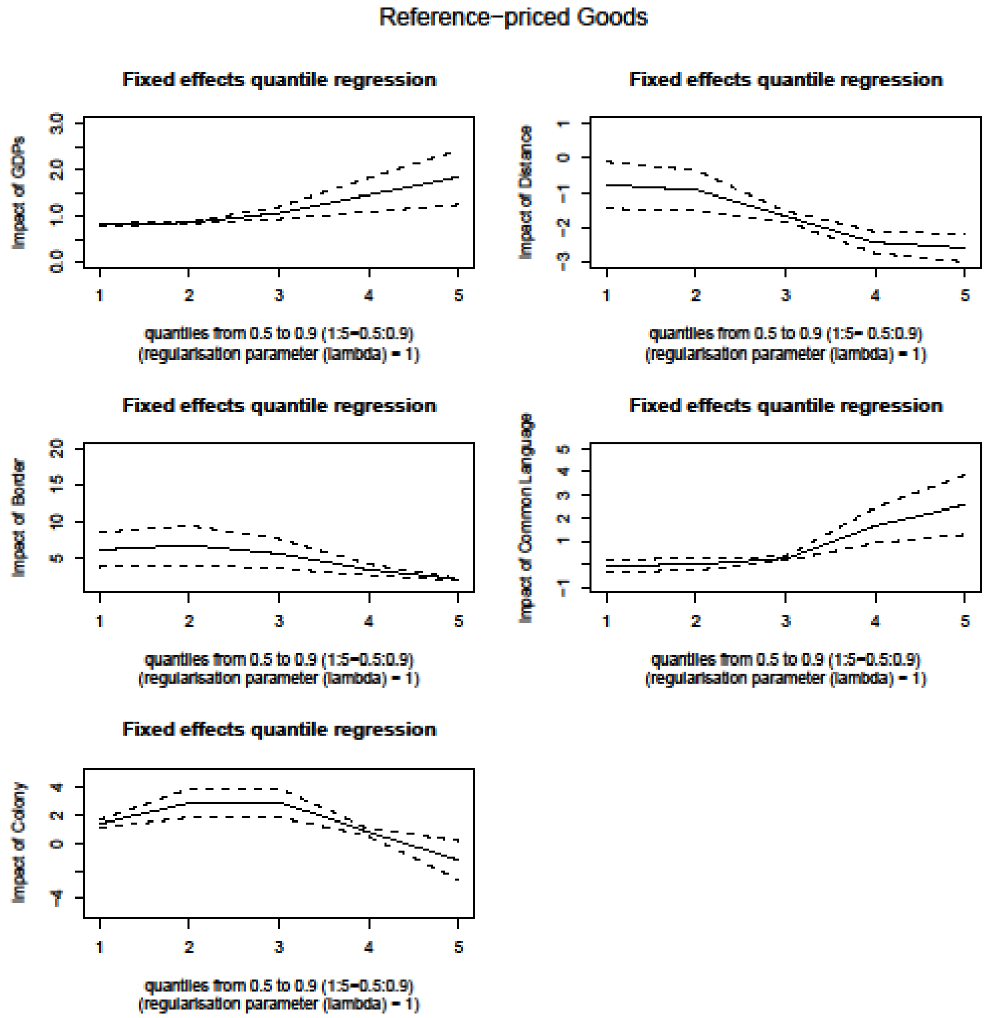


Appendix D

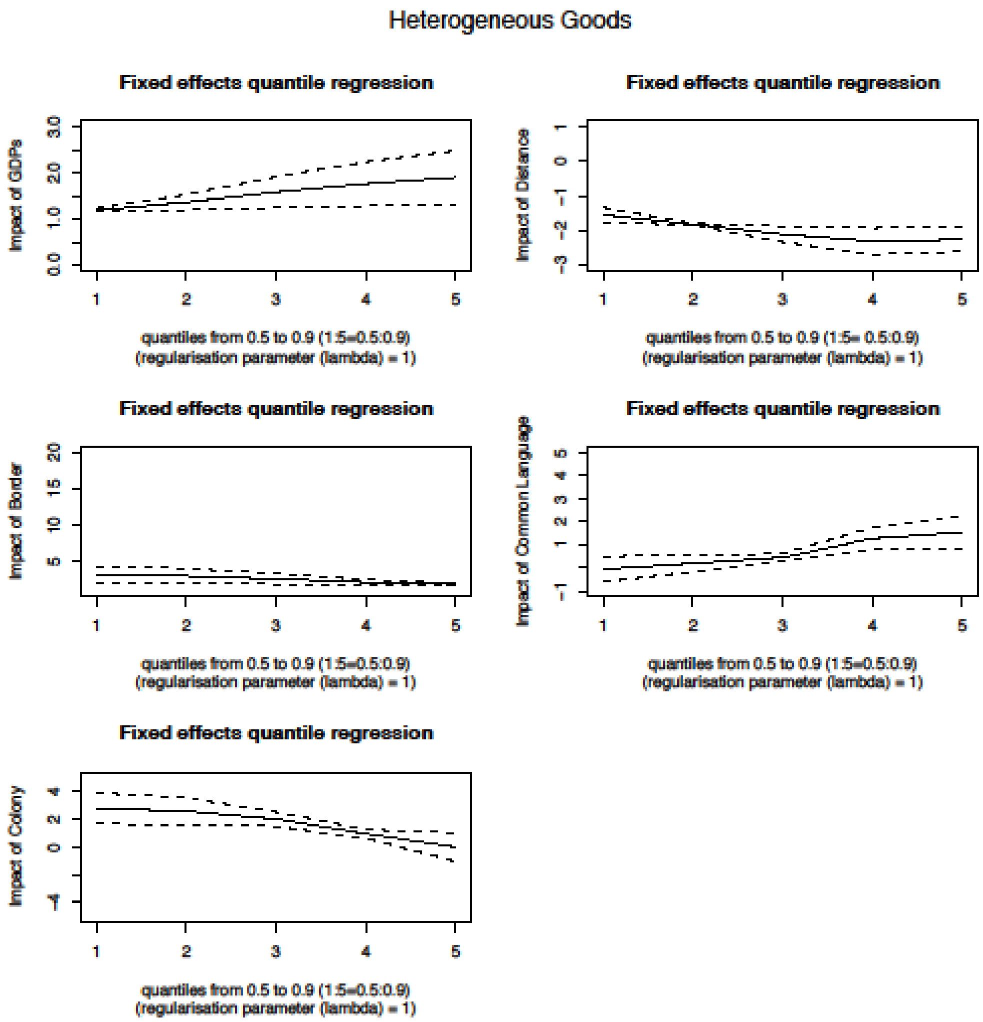
Graphs of Quantile Fixed Effect Estimation of Three Product Groups

Homogeneous Goods





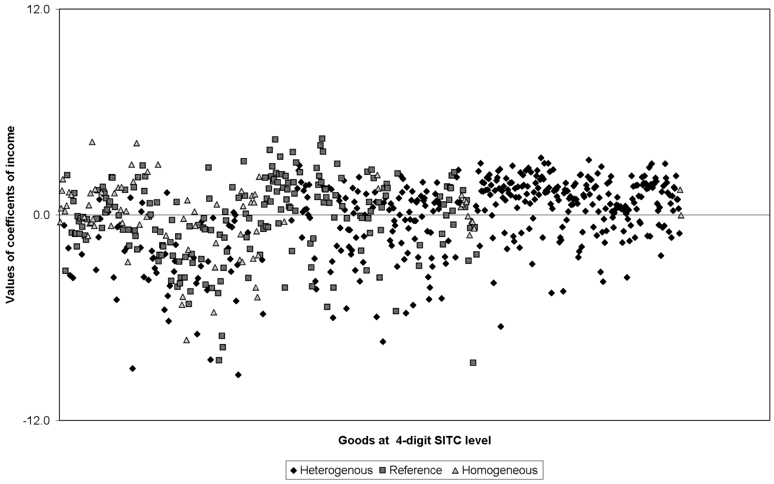
Reference−priced Goods
Fixed effects quantile regression
Impact of GDPs
Impact of Distance
Impact of Border
Impact of Common Language
Impact of Colony
quantiles from 0.5 to 0.9 (1:5=0.5:0.9)
(regularisation parameter (lambda) = 1)



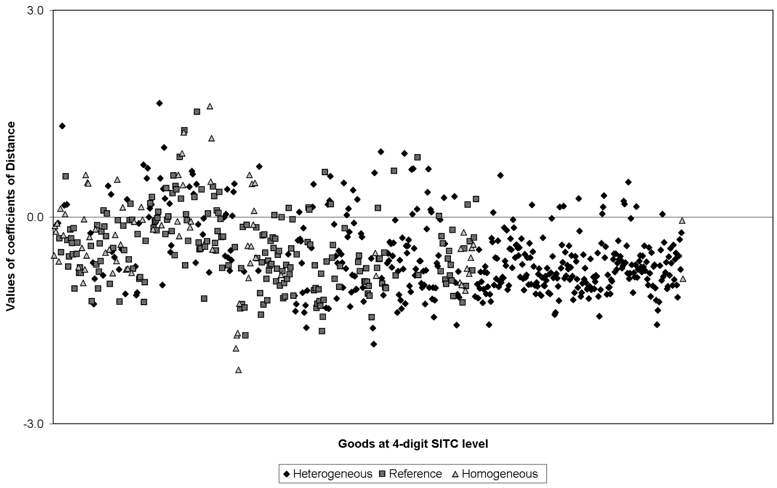
Appendix E

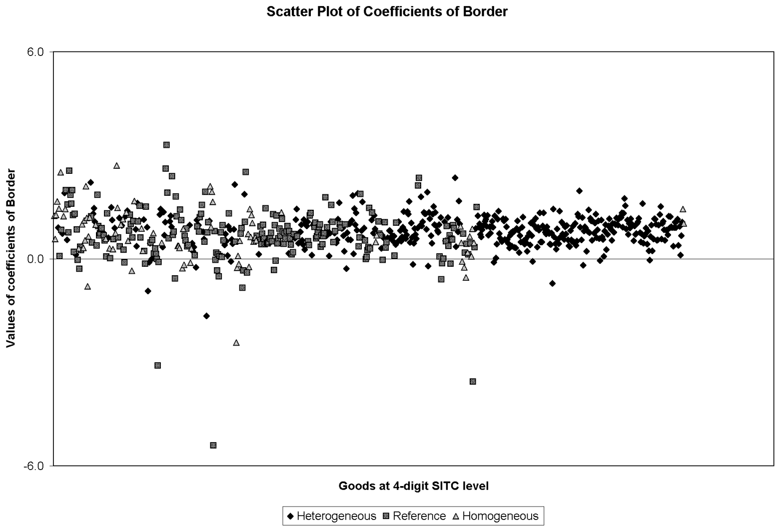
Beta Coefficients Obtained from Running 700 Gravity Equations for 700 Products

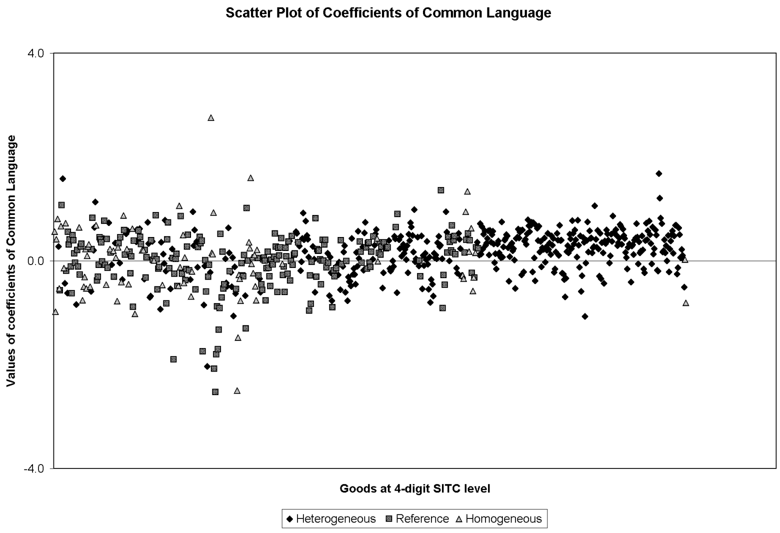
Scatter Plot of Coefficients of Income

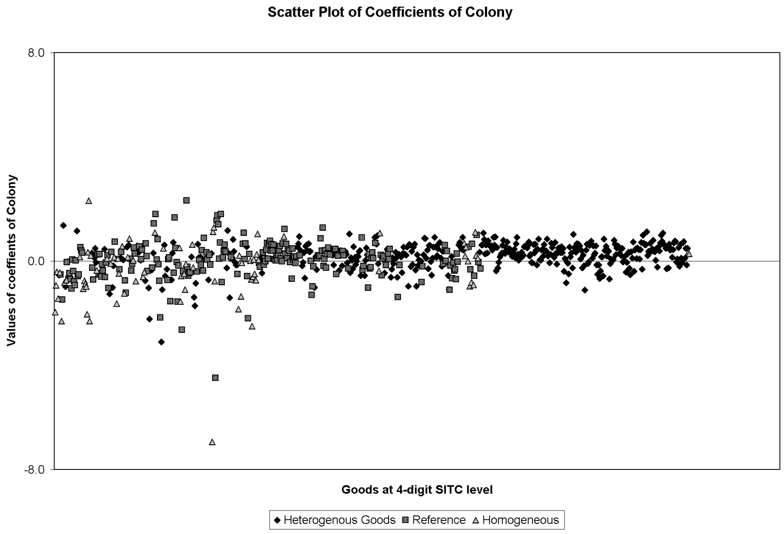


Scatter Plot of Coefficients of Distance




Scatter Plot of Coefficients of Border
6.0
0.0
-6.0
Values of coefficients of Border
Goods at 4-digit SITC level
Heterogeneous
Reference
Homogeneous


Scatter Plot of Coefficients of Common Language
4.0
0.0
-4.0
Values of coefficients of Common Language
Goods at 4-digit SITC level
Heterogeneous
Reference
Homogeneous


Scatter Plot of Coefficients of Colony
8.0
0.0
-8.0
Values of coeffients of Colony
Goods at 4-digit SITC level
Heterogenous Goods
Reference
Homogeneous