INTERNATIONAL EXPORT FLOWS OF VIETNAM: A GRAVITY MODEL APPROACH

THI HANH VU* (DULBEA AND ICITE)

ABSTRACT:
What determines export pattern of Vietnam with its trading partners? The author has employed the gravity model showing one-way exports of Vietnam to the world between 1997 and 2009. Since no proof was found of a possible endogeneity problem, the Hausman-Taylor methodology is therefore an inappropriate solution and the fixed effect model is consequently preferable. The role of the economic size indicator reaffirms its high relevance to export trade. The geographical distance gives an insignificantly negative sign suggesting that trade barrier needs to be examined in the context of Vietnam’s trading with the rest of the world. ASEAN as a regional economic integration organization plays a positive role therefore constitutes a significant absorptive factor. Interestingly, high labor productivity importing countries tend to have a lower demand for export products from Vietnam and the CPI - a proxy for the multilateral resistance factor is positive in respect of export trade.


KEYWORDS: Gravity model, Exports, Vietnam, Instrumental variables.

* Department of Applied Economics – DULBEA and iCite, Solvay Brussels School of Economics and Management, Université Libre de Bruxelles, CP 140, Avenue F.D. Roosevelt, 50, 1050 Bruxelles. Email: hvuthi@ulb.ac.be, Phone: +3 2(0) 494 813 846.
INTRODUCTION

In recent decades, globalization has led to a pattern of multilateral trading between countries or free trade associations. Exporting countries have been able to gain major benefits particularly through offering diverse products giving global consumers a wide choice. Such gains vanish where highly protected economies impose constraints on production and consequently consumption. Albeit, larger importing country markets lead to tougher competition between exporting countries who need to be able to surmount typical trade barriers such as geographic, social and economic hurdles. Some regional economic groups offer member countries preferential treatment for trade and stimulate production capacity as well as technological development which highlight the role of such an area in fostering free trade.

After two decades open to the world’s economy, Vietnam has achieved remarkable growth on some economic indicators such as GDP, Exports and Foreign Direct Investment. Since the US embargo was lifted in February 1994, Vietnam has become an active member of ASEAN\(^1\). Vietnam also became a member of APEC\(^2\) in 1998. Those steps paved the way for this small economy to stimulate bilateral and multilateral trade hence improving exports particularly when it became a full member of the World Trade Organization (WTO) in 2007. Between 1997 and 2009, exports and GDP gradually increased and exports peaked at over 50 billion USD in 2008, falling slightly to 50 billion USD in 2009. Meanwhile, GDP saw a gradual rise over the same period, increasing from 100 billion USD to over 250 billion USD in 13 years. Overall, the export-to-GDP ratio increased remarkably and the rate of GDP growth also increased, reaching 0.25% in 2009 and dropping sharply to 0.2 % partly due to the effect of the world financial crisis.

---

1 Association of Southeast Asian Nations (ASEAN) country members: Brunei Darussalam, Cambodia, Indonesia, Laos PDR, Malaysia, Myanmar, Philippines, Singapore, Thailand, Vietnam.
2 Asia-Pacific Economic Cooperation (APEC) country members: Australia, Brunei Darussalam, Canada, Chile, People's Republic of China, Hong Kong, China, Indonesia, Japan, Republic of Korea (South Korea), Malaysia, Mexico, New Zealand, Papua New Guinea, Peru, Philippines, Russia, Singapore, Chinese Taipei (Taiwan), Thailand, United States of America (USA), Viet Nam.
Many empirical studies on bilateral trade have confirmed the success of the gravity model in explaining the intra-trade among countries involving a single free trade area such as the EU, APEC, CARICOM or the inter-trade between two economic blocks. Serlenqa and Shin (2004); Rault, Sova and Ana Sova (2007); Mitze (2010) found that liberalization within EU tends to lower prices through increased competition and the removal of barriers to the mobility of goods leads to an increase in trade flows within the commodity. However, the role of the EU in the study by Bastos and Silva (2010) changing from positive to negative and insignificant to Portugal exports towards firm-fixed effect, product-fixed effect and firm-product fixed effects implies that productivity of firms frame the export pattern.

In regard to the demand side of a partial equilibrium, the income of each economy measured by GDP, and/or GNP shows that country would expense more if it has the ability to earn more. Further, literature on gravity model demonstrates a higher preference to heterogeneous import products rather than to homogenous goods. To express differently, Nguyen and Arcand (2009) found that homogenous goods are less responsive to changes in income than heterogeneous goods. Regardless of time and space, GDP consistently enhances trade capacity between countries even if many papers tend to confirm endogeneity for these indicators.

Barriers in international trade include geographical distance, social distance and economic distance. Geographical distance is regarded as transport cost. In most of the paper using aggregated data on trade, distance decreases systematically with export and import value. Although distance is a time invariant factor, using it as a proxy is problematic since the underlying assumptions vary over time. To resolve, Brun, Carrère, Guillaumont and Melo (2005) defined the augmented transport cost function that includes indexes for infrastructure, price of oil in a random effect model showing that distance did not vanish and providing support for the distance estimates in a range [-1.3; -0.8]. Bastos and Silva (2010) conversely demonstrated that distance for a firm’s exports does not matter since more competitive firms
could serve more distant markets. Such firms may be able to charge a higher price and/or sell in larger quantities to more distant market which strongly supports the idea that a firm’s productivity is a key factor determining its export pattern.

How does the gravity model explain export flows of Vietnam? If we make some new assumptions on certain typical factors of the full gravity model, this then makes it possible to solve the puzzle as to whether a regional economic association such as ASEAN remains relevant for a Vietnamese exporter. There is little existing literature on the gravity model for Vietnamese bilateral and multilateral trade. Do (2006) has used data on Vietnam’s exports and imports to 23 European countries between 1993 and 2004 to analyze trade potential with individual importing countries. Nevertheless, the paper is limited to fixed and random effect test where an endogeneity problem may arise. Nguyen (2009) examined the intra-trade between Vietnam and the other ASEAN countries under AFTA and applied a similar model but the absence of some trading partners therefore in both cases meant that the conclusions about the determinants were somewhat incorrect. To fill the gap, I attempted to examine the pattern of Vietnam’s exports to all the country trading partners. A significant contribution of this paper is that it deals with the endogeneity problem showing that when being controlled, the estimators of gravity model therefore give the different interesting conclusions for export determinants.

The main findings reveal that high income trading partners are rightly important to Vietnamese exports, whereas high labor productivity markets are not always a suitable target. The greater the difference between the GDP of the potential importing country and that of Vietnam, the greater the labor cost advantage of Vietnam for labor intensive sectors. Moreover, the regional economic integration area – ASEAN, helps Vietnamese exports. The multilateral resistance factor is a better indicator to express trade cost between two countries rather than geographical distance, being insignificant in the model.

The remainder of the paper is structured as follows. Section 1 explains the theoretical view of the gravity model. Section 2 describes the data used in the empirical analysis and the methodology. Section 3 discusses the role of the full gravity model in explaining exports at country level. The last section offers some concluding remarks.

1. THEORETICAL BACKGROUND ON THE GRAVITY MODEL

The main branch of New Trade Theory is the gravity model firstly introduced by Tinbergen (1962) considering the interaction in trade between country pairs. The baseline model is a modified version of Isaac Newton’s Law of Gravitation to predict movement of commodities between countries and continents taking into account the population size of the two things and their distance.

$$\frac{\text{population}_1 \times \text{population}_2}{\text{distance}^2}$$
As Christie (2002) specified that trade flow between any two countries is proportional to the product of each country’s economic mass measured by GDP, each to the power of quantities to be determined, divided by the distance between the countries respective “economic center of gravity” generally their capital, raised to the power of another quantity to be determined. The formulation is generalized to

\[ M_{ij} = KY_i^\beta Y_j^\gamma D_{ij}^\delta \]

Where \( M_{ij} \) is the flow of trade between country i and country j, \( Y_i \) and \( Y_j \) are the country i’s and country j’s GDP and \( D_{ij} \) is the geographical distance among two countries’ capital. The above model could be put in the linear form as follows:

\[ \log (M_{ij}) = \alpha + \beta \log (Y_i) + \gamma \log (Y_j) + \delta \log (D_{ij}) \]

In some cases, a group of dummy variables can be added to the traditional model comprising of the country’s being a member of trade agreement, sharing the common border or speaking the same language, etc.

The theoretical framework of the gravity model later focused on the supply and demand side and how consumers’ demand could be satisfied as to the utility of imported and domestically produced commodities. Linneman (1966), Aitken(1973), Geraci and Prewo (1977) highly criticized that the gravity model is a reduced form of a four—equation partial equilibrium model of export supply and import demand. Goldstein and Khan (1978) indicated that gravity on exports is a function of a simultaneous equation between the supply and demand side of exports and their price based on quarterly export data of 8 European countries.

James E. Anderson (1979) referred to the gravity model as a pure expenditure system applied to commodities where each country completely specializes in producing its own goods. It is also appended to the Cobb-Douglas function and linked to the production endowments including share of country’s income and labor capital. Income per capita is an exogenous demand-side factor and population a supply side factor. However, there would be a tradeoff between bias and efficiency since countries have considerable differences of income, preferences and production capacity which lead to biased estimate of GDP.

The gravity model can be derived from very different models including Ricardian, Heckscher-Ohlin and increasing return to scale (IRS) model (Helpman and Krugman (1985). Relating to the supply and demand function, the Ricardian model assumed income as expenditure focussing on the difference between productivity of labor between countries but labor productivity is based mainly on differences in technology. A country with a comparative advantage in producing a good uses its resources most efficiently when it produces that good rather than producing other goods. The opportunity cost of producing wine is the amount of cheese not produced (trade off). According to Heckscher-Ohlin, a country will export that commodity which uses intensively its abundant factors such as labor, and land while importing a commodity which uses intensively its scarce factors. Helpman
and Krugman (1985) showed that specialization and trade will persist even when countries have identical relative factor endowment for a wide variety of models. They asserted that the theory behind comparative advantage does not predict the relationship in a gravity model. Countries with similar level of income have been shown to trade more. The authors also emphasized that these countries are trading in differentiated goods because of their similarities.

Bergstrand (1985) considered this approach is “loose” and proposed to explain the gravity model in a multiplicative functional form. He found that trade flows are differentiated by country origin and result in imperfect international product substitutability (against the assumption of the previous papers which confirms that the elasticity of substitution between export markets exceed that between for domestic and foreign market, including perfect international product substitutability. Bergstrand (1989) derived the generalized gravity equation, using per capita income as a proxy for the exporter’s K/L ratio which depends on whether the gravity equation is estimated for a capital or labor – intensive industry. He differentiated export products into one-digit categories investigating for labor intensive products such as food & live animals, beverages & tobacco and has a significant negative influence for exports of raw material, fuel and chemical implying that the industry’s output is for luxury good production.

More recently, Deardorff (1995) has proved the link between the gravity equation and the H-O model by assuming that each country could in theory, specialize in producing one goods and if importers do not have the demand for such goods, then trade between the two countries would be zero. Since the gravity equation implies the expenditure system attached to the theory that higher income consumers are ready to pay more for capital intensive products, then exporters would supply the relevant items. Deardorff found the gravity equation relevant to the H-O theory. This explains why capital-intensive countries export less than average to low-income labor-abundant countries. Deardorff’s explanation (1995) may seem implausible since countries do not always require a single goods, but as their preferences change, their requirements may also change.

Almost all theoretical papers on gravity for bilateral trade employ the geographical distance between two countries pair as a proxy for transport cost. A country contiguous to its destination market is not always at an advantage for trade. In reality, transport costs are a function of fuel costs, type of product, volume of trade and therefore variant. It is impossible to identify them in the fixed effect model controlling for the specific effects of country pairs. Suppose that, two countries with a common border were separated from the rest of the world and had the same level of economic development then they would be expected to trade more. In most cases, such a country does not only trade with only one country. This explains why two contiguous countries sharing the same pace of economic development trade less. Anderson and Wincoop (2003) suggested a multilateral resistance term in replacement of the geographical distance and border sharing in the U.S-Canada case (McCallum’s, 1995). Such a multilateral resistance factor implies that if two countries were isolated from the rest of the world on one island in the ocean, far away from the next continent, bilateral average trade costs measured by the iceberg factor might be low and this should guarantee a high trade volume between both
countries. However, in reality, the two “isolated countries” are surrounded by many very large countries therefore multilateral resistance is much lower and thus trade between the two countries is low even if GDPs and trade costs are assumed to be at the same level. Rudolph (2010) has employed the total production index of the exporting country and importing countries belonging to the rest of the world to compute the multilateral resistance of trading countries for 23 OECD countries. However, data on a country’s total production index are not always available.

2. DATA AND METHODOLOGY

2.1. DATA DESCRIPTION

Export volume: The paper uses data on Vietnam’s export trade with countries across five continents over a period of 13 years between 1997 and 2009. Such data are obtained via WITS (World Integrated Trade Solution) on the website of the World Bank which closely collaborated and consulted with UNSD (United Nation Statistical Division). The UN Comtrade data base contains information on exports and imports broken down by commodity and partner country. The data base comprises various relevant items such as gross export value\(^3\) calculated in thousands of USD annually, partner countries by country name and country code from 1997 to 2009. Overall, Vietnam’s total and continental export value gradually increased until the year 2008 and fell back considerably in 2009 as a consequence of the global crisis (See figure 2). Specifically, Vietnam exports more to Asian countries and some other developed markets whereas the African market has been comparatively neglected. In 2002, the number of importing countries peaked at 188 whereas in 2009, there were only 130 countries importing from Vietnam. The country’s export volume peaked in 2008 but declined to about 55 billion USD in 2009.

\(^3\) Gross export value or net export represents the value of a country’s total exports minus the value of its total imports and is used to calculate a country’s aggregate expenditure, or GDP, in an open economy.
### Table 1. Descriptive Statistics of Export Model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>lnExport</td>
<td>8.522255</td>
<td>3.046528</td>
<td>-3.523984</td>
<td>16.29229</td>
<td>N = 2082</td>
</tr>
<tr>
<td></td>
<td>overall</td>
<td>between</td>
<td>within</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.933666</td>
<td>2.422588</td>
<td>1.288242</td>
<td>1.791373</td>
<td>n = 195</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.288242</td>
<td>1.791373</td>
<td>13.28827</td>
<td>T-bar = 10.6769</td>
</tr>
<tr>
<td>lnGDP</td>
<td>24.48665</td>
<td>2.149588</td>
<td>18.74553</td>
<td>30.29106</td>
<td>N = 1962</td>
</tr>
<tr>
<td></td>
<td>overall</td>
<td>between</td>
<td>within</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.264411</td>
<td>18.98828</td>
<td>30.04841</td>
<td>25.45611</td>
<td>n = 177</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.590377</td>
<td>23.52135</td>
<td></td>
<td>T-bar = 11.0847</td>
</tr>
<tr>
<td>lnGDP/Labor</td>
<td>9.476304</td>
<td>1.254469</td>
<td>6.410191</td>
<td>12.14053</td>
<td>N = 1909</td>
</tr>
<tr>
<td></td>
<td>overall</td>
<td>between</td>
<td>within</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.24628</td>
<td>6.492335</td>
<td>11.91702</td>
<td></td>
<td>n = 169</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.957588</td>
<td>8.698967</td>
<td>10.31727</td>
<td>T-bar = 11.2959</td>
</tr>
<tr>
<td>lnDGDP</td>
<td>8.198246</td>
<td>1.896969</td>
<td>-1.101007</td>
<td>11.31331</td>
<td>N = 1962</td>
</tr>
<tr>
<td></td>
<td>overall</td>
<td>between</td>
<td>within</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.613298</td>
<td>4.368188</td>
<td>11.1137</td>
<td></td>
<td>n = 177</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.320088</td>
<td>2.1198</td>
<td>10.36802</td>
<td>T-bar = 11.0847</td>
</tr>
<tr>
<td>CPI</td>
<td>22.59629</td>
<td>568.0742</td>
<td>-10.06749</td>
<td>24411.03</td>
<td>N = 1855</td>
</tr>
<tr>
<td></td>
<td>overall</td>
<td>between</td>
<td>within</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>204.6748</td>
<td>-0.357405</td>
<td>2687.38</td>
<td></td>
<td>n = 172</td>
</tr>
<tr>
<td></td>
<td></td>
<td>532.9437</td>
<td>-2664.047</td>
<td>21746.25</td>
<td>T-bar = 10.7849</td>
</tr>
<tr>
<td>lnDistance</td>
<td>8.943675</td>
<td>6.760568</td>
<td>6.170767</td>
<td>9.856429</td>
<td>N = 2082</td>
</tr>
<tr>
<td></td>
<td>overall</td>
<td>between</td>
<td>within</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>6552468</td>
<td>6.170767</td>
<td>9.856429</td>
<td></td>
<td>n = 195</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>8.943675</td>
<td>8.943675</td>
<td>T-bar = 10.6769</td>
</tr>
<tr>
<td>ASEAN</td>
<td>.0542747</td>
<td>.2266134</td>
<td>0</td>
<td>1</td>
<td>N = 2082</td>
</tr>
<tr>
<td></td>
<td>overall</td>
<td>between</td>
<td>within</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.2103583</td>
<td>0</td>
<td>1</td>
<td></td>
<td>n = 195</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>.0542747</td>
<td>.0542747</td>
<td>T-bar = 10.6769</td>
</tr>
<tr>
<td>OECD</td>
<td>.2050913</td>
<td>.4038653</td>
<td>0</td>
<td>1</td>
<td>N = 2082</td>
</tr>
<tr>
<td></td>
<td>overall</td>
<td>between</td>
<td>within</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.3803945</td>
<td>0</td>
<td>1</td>
<td></td>
<td>n = 195</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>.2050913</td>
<td>.2050913</td>
<td>T-bar = 10.6769</td>
</tr>
<tr>
<td>Landlocked</td>
<td>.1748319</td>
<td>.3799145</td>
<td>0</td>
<td>1</td>
<td>N = 2082</td>
</tr>
<tr>
<td></td>
<td>overall</td>
<td>between</td>
<td>within</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.3889839</td>
<td>0</td>
<td>1</td>
<td></td>
<td>n = 195</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>.1748319</td>
<td>.1748319</td>
<td>T-bar = 10.6769</td>
</tr>
</tbody>
</table>

**Figure 2. Exports by Continents**


**Importing countries:** Data on GDP adjusted to Purchasing Power Parity (PPP) are available on the World Bank’s website. Also, I have obtained data on the labor force of importing countries over this period via the KILM data available from the

---

KILM: Key Indicators of the Labor Market.
International Labor Organization. Simply, by dividing real GDP over labor force, I was able to generate the importer’s productivity index. In regard to expenditure equilibrium, countries with a higher level of income will pay more for goods and services. Although, GDP per capita might seem to be a more plausible proxy for a country's demand capacity than simple GDP since it shows the relative performance of the country, this variable is not included in the model simultaneously with covariate the labor productivity due to a problem of multicollinearity. Depending on the level of GDP per capita of importing countries, imports would tend to be labor or capital intensive goods.

It is noteworthy that the top 15 importing countries account for an average of 80% of Vietnam’s total exports in this period. Although the rank of the major importing partners fluctuated yearly, it includes the United States of America, Japan, China, Australia, Singapore, Germany, France, and the United Kingdom. Importantly, the top 15 highest income countries are also the largest importing countries as mentioned above. In fact, only 25% on average of the highest income countries have the highest labor productivity and are not Vietnam's largest markets.

Data on geographical distance between Vietnam and its country partners is available via CEPII. The data base consists of country code, the distance measured for a given country pair’s capital and other dummy variables showing if exporters and importers share a common border (contig), the common language (comlang), the common colony (comcol) and whether countries are entirely surrounded by land (landlocked).

2.2. METHODOLOGY

This study examines the export trade flows between Vietnam and all importing countries across continents. Responding to the theoretical framework of the gravity model, in this paper, I have supplemented the model so that it now includes the following determinants:

\[ \ln E_{ijt} = \alpha_0 + \beta_1 \ln GDP_j + \beta_2 \ln GDP_j LB_j + \beta_3 \ln GDP_j LB_j + \gamma_1 \ln Distance_{ij} + \gamma_2 ASEA_j + \gamma_3 OECD_j + \gamma_4 Landlocked_j + \epsilon_{ijt} \]

Where:

- \( E_{ijt} \) denotes the aggregated exports in USD by Vietnam to each destination market \( j \) at time \( t \);
- \( \alpha_0 \) is the common intercept;
- \( GDP_j \) represents the Gross Domestic Product of country \( j \) at time \( t \);
- \( (GDP_j / LB_j) \) denotes the Gross Domestic Product Per Capita of importers \( j \) at time \( t \);
- \( ASEA_j \) and \( OECD_j \) are dummy variables indicating whether country \( j \) belongs to the ASEA or OECD, respectively;
- \( Landlocked_j \) is a dummy variable indicating whether country \( j \) is landlocked.

---

5 The author has calculated countries’ labor productivity index by dividing GDP adjusted at power purchasing parity and constant at 2005 by country’s labor force.

6 CEPII stands for Centre d’Etudes Prospective et d’Informations Internationales
$DGDP_{ijt}$ is the difference between GDP per capita of Vietnam and that of the importing partner and is calculated by 

$$DGDP_{ijt} = \frac{GDP_{it} - GDP_{jt}}{POP_{it} - POP_{jt}}$$

$\ln DGDP_{ijt}$ measures the difference in terms of the relative factor endowments between two countries and is zero where there is equality in the relative factor endowments.

$CPI_{jt}$ reflects the annual percentage change in the cost to the average consumer of acquiring a basket of goods and services that may be fixed or changed at specified intervals. This index is a proxy for the multilateral resistance factor since data for the intra-national trade of the exporting importing partners are not available. Unlike the other variables, CPI is expressed in the log-linear form since many values are negative for some countries and in some periods.

$Distance_{ij}$ represents the geographical distance between the capitals of the two countries and is a proxy for the transport costs;

The value 1 is attributed to ASEAN as dummy variable where the importer is an ASEAN member or zero otherwise. This indicates that such export trade is ASEAN intra-export trade;

The value 1 is also attributed to Landlocked a dummy where the importing country is surrounded by land or zero otherwise;

The paper aims to investigate Vietnamese exports to the global market over a period of 13 years from 1997 to 2009. Panel data are relevant to account for the unobservable effects, characteristic of each partner country. Hence, the purpose of employing panel data is to model exports to partner countries as a function of the income and productivity of importing countries. These covariates vary depending on partner country and time. Moreover, there are many time-invariant variables which may affect the independent variables such as colonial history, financial situation, religious affiliation and political regime. Not accounting for country heterogeneity causes serious misspecification. Baltagi (2005) considered that panel data are more informative, provide more variability, reduce colinearity of the variables, give a greater degree of freedom and thereby improve the efficiency of the model. With the additional, more informative panel data, more reliable parameter estimates can be produced. Individual characteristics of importing countries rather than conventional determinants by importing countries can nevertheless influence exports. For comparison purposes, the above proposed model is run using the pooled OLS as follows:

$$Y_{ijt} = \alpha_i + \beta X_{ijt} + e_{ijt}$$

Where $i = 1, 2, \ldots, N$ and $t = 1, 2, \ldots, T$

And satisfied with $e \mid X \sim iid (0, \sigma^2 I_N)$

$$Q = E_N \otimes E_T - I_N \otimes I_T - I_N \otimes \overline{J_T} - \overline{J_N} \otimes I_T + \overline{J_N} \otimes \overline{J_T}$$

This is because there is a strong assumption for the application of the pooled OLS model that $\alpha_0$ should be common to all countries meaning that there is no difference between specific importing countries in terms of unobservable effects. If
This condition is satisfied, pooled OLS is unbiased and consistent. In other words, if there are some individual effects hidden within the residual \( e_{jt} = \hat{e}_{jt} + \alpha_0 \) or \( \alpha_0 \) is correlated with \( X_{jt} \) then \( E[h] = b + E[X'_{jt}X_{jt})^{-1}X'_{jt}\alpha_0] \) and pooled OLS becomes biased.

Once the regression shows unobservable effects, the pooled OLS gives biased estimates and then, the fixed effect model is preferable. In the fixed effect model, \( \alpha_0 \) is thought of as a fixed parameter which enables dummy variables to be introduced for each exporter-importer pairs. These dummy variables allow correction of this error term. Albeit, another problem would arise where the database contains many country partners thus creating a need for many dummies. Estimated N+K parameters for the model would cause a multicolinearity problem. According to Cheng and Wall (2005), the general model of trade between country pairs could be characterized by

\[
\ln X_{kt} = \alpha_0 + \alpha_t + \alpha_k + \beta_{kt}Z_{kt} + \varepsilon_{kt} \quad t = 1, T;
\]

Where \( Z_{kt} \) is the 1 x k row of vector of gravity model specified in detail in the equation (1) above. The intercept normally has three parts, one which is common to all years and trading partners at the level of country \( \alpha_0 \), the other is specific to year \( t \) and common to all trading partners, \( \alpha_t \) and one which is specific to country pairs and common to all years. The disturbance \( \varepsilon_{kt} \) is assumed to be normally distributed with zero mean and constant variance for all observations. To avoid the multicolinearity problem rooted in a large (N+T+K-1) matrix, the fixed effect estimates of \( \beta \) can be obtained by performing the following transformation proposed by Wallace and Hussain (1969):

\[
Q = E_N \otimes E_T = I_N \otimes I_T - I_N \otimes \bar{J}_T - J_N \otimes I_T + J_N \otimes \bar{J}_T
\]

Where \( E_N = I_N - \bar{J}_N \) and \( E_T = I_T - \bar{J}_T \). This transformation removes the country and time specific effect. In fact, \( \tilde{y} = Qy \) has a typical element \( \tilde{y}_{jt} = (y_{jt} - \bar{y}_j - \bar{y}_t + \bar{y}) \) where \( \bar{y} = \sum_j \sum_t y_{jt} / NT \) and the regression of \( \tilde{y} = Qy \) and \( \tilde{X} = QX \) to obtain the estimator \( \tilde{\beta} = (Q'Q)^{-1}Q'y \).

One side effect of the fixed effect model is that the transformation simultaneously removes time invariant effect or time invariant variables. It is not possible to identify the distance effect and the dummies such as “ASEAN” or “landlocked”. Rather, the fixed effect estimator ignores variations across the countries. The characteristics within individual countries may or may not correlate with the explanatory variables. In such cases, the instrumental method is appropriate to treat the correlation between some explanatory variables and unobservable effects.
In contrast with the fixed effect, the random effect could be an alternative solution if the variation across entities is assumed to be random and uncorrelated with the predictor or independent variables included in the model. If $\alpha_i \sim IID(0, \sigma_{\alpha_i}^2)$, $\alpha_t \sim IID(0, \sigma_{\alpha_t}^2)$ and $\nu_{ijt} \sim (0, \sigma_{\nu}^2)$ are independent of each other, then this is a two-way random effects model. In addition, $X_{ijt}$ is independent of $\alpha_i$ and $\alpha_t$ and $\nu_{ijt}$ for all $i$ and $t$. From (1) one can compute the variance-covariance matrix

$$\Omega = E(u'u) = Z_{ai}E(\alpha_i \alpha_i')Z_{ai} + Z_{ai}E(\alpha_i \alpha_t')Z_{ai} + \sigma_{\nu}^2 I_{NT}$$

$$= \sigma_{\alpha_i}^2 (I_N \otimes J_T) + \sigma_{\alpha_t}^2 (J_N \otimes I_T) + \sigma_{\nu}^2 (I_N \otimes I_T)$$

In order to obtain $\Omega^{-1}$, $J_N$ is replaced by $N J_N$, $I_N$ by $E_N + J_N$, $J_T$ by $T J_T$ and $I_T$ by $E_T + J_T$ and collecting terms with the same matrices. This gives

$$\Omega = \sum_{i=1}^4 \alpha_i Q_i.$$  Each $Q_i$ is symmetric and idempotent with its rank equal to its trace. Moreover, $Q_i$ are pairwise orthogonal and sum to the identity matrix. The advantages of this spectral decomposition are that $\Omega' = \sum_{i=1}^4 \alpha_i' Q_i$. Where $r$ is an arbitrary scalar so that $\sigma_r \Omega^{-1/2} = \sum_{i=1}^4 (\sigma_r / \alpha_i^{1/2})Q_i$. And the typical element of $y^* = \sigma_r \Omega^{-1/2} y$ is given by $y^*_{ijt} = y_{ijt} - \theta_1 y_i - \theta_2 y_t + \theta_3 y_{ij}$.  As a result, GLS can be obtained as OLS of $y^*$ on $Z^*$, where $Z^* = \sigma_r \Omega^{-1/2} Z$.

The implication behind random effect is that, all individual differences are captured by the intercept parameters. In the random model, the intercept parameters consist of a fixed part that represents the population average and the individual differences from the population average. In contrast to the fixed effect model, by definition, the random effect approach allows the time-invariant covariates to be estimated. A random effect estimator is a generalized least square procedure and the fixed effect is a least square estimator. In large samples, the GLS estimator has a lower variance than the least square estimator. However, a disadvantage of the random effect sample is that the assumption on the relationship between unobservable effect and explanatory variables goes to zero and if any explanatory variable correlates with unobservable effects, the estimates become biased and inconsistent.

Hausman and Taylor (1981) have proposed another approach which serves to find instruments for the columns of time-variant variables ($X$) and time-invariant
variables (Z) being potentially correlated with $\alpha_i$. In some cases, it is not always easy to find convenient instruments among those variables which are not correlated with $\alpha_i$. If it is possible, the conditions may hold such that all of $\beta_s$'s and $\gamma_s$'s may be consistently and efficiently estimated. The $X_{ijt}$ columns which are not correlated with $\alpha_i$ can serve two functions whereby (i) using deviations from individual means, they produce unbiased estimates of $\beta_s$'s and (ii) using the individual means, they provide a valid instrument in the columns of Zi correlated with $\alpha_i$. Because the only component of the disturbance is time-invariant, any vector orthogonal to a time-invariant vector can be used as an instrument. By construction, $Q_i\alpha_i = 0$ thus, the time-invariant provides at least $TN-N$ linearly independent instruments, namely $TN-N$ basis vectors of the column space of $Q_i$.

Unfortunately, $Q_i$ is also orthogonal to $Z_i$ which violates the requirement that instruments correlate with all of the explanatory variables. If some exogenous k columns are uncorrelated with $\alpha_i$ then k’s can only serve as instruments.

According to the Hausman and Taylor approach, the principles of the test can be outlined as follows: $H_0$: $p \lim \frac{1}{N}X_{ij},\alpha_i = 0$ and $p \lim \frac{1}{N}Z_{ij},\alpha_i = 0$. Under $H_0$, both $\hat{\beta}_w$ and $\hat{\beta}^*$ are consistent, while under the alternative, $p \lim \hat{\beta}^* \neq p \lim \hat{\beta}_w$ thus the deviations of $q = \hat{\beta}^* - \hat{\beta}_w$ from the zero factor cast doubt on the null hypothesis. To form a $\chi^2$ test based on $q$, premultiply the original equation by $Q_Z\Omega^{-1/2} = [I_{NT} - Z_i(Z_i'Z_i)^{-1}Z_i']\Omega^{-1/2}$ and considering the within-group and efficient estimators for $\beta$ in the transformed equation. Let $X^* = Q_Z\Omega^{-1/2}X_{ij}$ we then have

$$q = [(X^*P_A X^*)^{-1} X^*P_A Q_Z -(X^*Q_Z X^*)^{-1} X^*Q_Z Q_Z] \Omega^{-1/2}Y_{ij} = DY^*$$

Where $Q_Z\Omega^{-1/2} = Q_y$ and $Y^* = \Omega^{-1/2}Y$ has a scalar covariance matrix. The specification test statistics are given by

$$\hat{m} = \hat{q}[\text{cov}(q)]^{-1} \hat{q} = \hat{q}[\text{cov}(\hat{\beta}_w) - \text{cov}(\hat{\beta}^*)]^{-1} \hat{q} = \hat{q}[\sigma^{-2}_qDD']^{-1} \hat{q}$$
3. EMPirical Results

3.1. Methodology

The empirical results reported in the table 2 present the estimates of the pooled OLS, fixed effect, random effect and Hausman & Taylor approach. The value of exports, GDP, GDP per labor force and the difference between country’s GDP per capita in USD are shown in the form of natural logarithms. For efficiency, all coefficients are corrected for standard errors. The pooled OLS model shows heteroskedasticity and imperfect auto-correlation. Also, a groupwise heteroskedasticity problem arises in the fixed effect model implying that the variance within country pair exports is not homogenous. The Breusch and Pagan Lagrangian multiplier test for random effects reveals heterogeneous variance between country pair exports. Unlike the fixed effect model which treats the explanatory variables to be correlated with the unobservable effect \( ui \), the random effect model assumes that there is no correlation between covariates and unobservable effects. Since \( ui \) is considered to be random, the GLS approach is an appropriate tool to control the effects of the model for both within group and between group countries.

Pooled OLS will be biased and inconsistent if unobservable effects are found to correlate with independent variables. Unobservable effects may be the other time-invariant factors that are not included in the model such as political or social characteristics within each partner country. Although the Hausman test does not reject the null hypothesis that the random effect test is appropriate, the results are not identical with Hausman's two options which are sigmamore and sigmaless. The former option is based on both (co) variance matrices on disturbance variance estimate derived from the efficient estimator, the latter based on both (co) variance matrices on the disturbance variance estimate derived from the consistent estimator. Hence, the Hausman test with the two options favors the fixed effect model.

The Breusch and Pagan Lagrangian multiplier test result also shown in table 2 advocates the choice of the random effect model since p-value is significant enough to reject the null hypothesis that the variance of the specific errors is equal to zero. However, it is generally assumed in a random model that there are uncorrelated relationships between covariates and unobservable effects, so the estimators are biased if some of the independent variables are found to be related to unobservable effects. If it is the case, for an all or nothing correlation in a fixed or random world, neither models is valid. This being so, using the Sargen and Hansen, calibrate the validity of the model in a random world and found that the model gave a significantly overestimated result.

Column 4 of table 2 includes all necessary variables including time-variant and time-invariant variables identifiable by a random model. Four independent variables GDP, GDP per labor, DGDP and CPI are assumed to be endogenous which could be correlated with the unobservable effects. The Hausman and Taylor world bring

\[ \chi^2 (15) = 47.69 \] and significant at 1%.

\[ \chi^2 (15) = 48.60 \] and significant at 1%.
two advantages: Firstly, unlike a fixed effect model, it allows unobservable effects to be treated as random enable the time-invariant regressors to be estimated. Secondly, it avoids the null correlation assumption for covariates and unobservable effects. These four independent variables are assumed to be endogenous on the basis of the testable hypothesis. The remaining regressors are set as exogenous variables. The Sargan and Hansen test is an appropriate application to test for the null hypothesis that the excluded instruments are valid instruments for endogenous variables and uncorrelated with the individual effect and correctly excluded from the estimated coefficients of gravity model. As noted in the table 2, for this random effect model, the P-value indicates that the random estimator requires a larger number of orthogonal conditions, while the P-value\(^8\) of the Sargan and Hansen test for HT/IV is insignificant and therefore rejects the null hypothesis and also determines the validity of the excluded instruments.

### 3.2. DISCUSSION OF RESULTS

As expected, the coefficients for the income of importing countries are positive and significant at 1% implying that higher income countries import more. According to the result shown in the table 2, coefficient of importers’GDP is under-estimated. According to the Hausman and Taylor test, holding other regressors constant, when the income of Vietnamese trading partners increases by one percent, the export volume of Vietnam to its importing countries would increase by 1.35%. If gravity model is a type of the reduced expenditure equation, it explains why it confirms the theoretical notion that higher income induces higher expenditure. For Bergstrand’s (1985) utility function where the total good and service utility of one country is subject to its income, it means that if that country is able to satisfy its demand for goods then it would expand its ability to increase consumption of domestic and imported goods. If consumers make a trade off either between these two categories of goods or their quantity, the country utility would be improved if incomes rises and the demand curve would shift to the right. Empirical studies on country level data such as Zarzoz and Lehman (2003), Papazouglow (2007), Carrère (2006), Zhang and Kristensen (1995) have proved the importance GDPs contribution to the intensity of trade among developed and developing countries. Furthermore, Bastos and Silva (2010) and Nguyen (2009) at firm and product level also highlight the important and positive role of GDP. The findings of Nguyen (2009) indicates importing countries’ income determines export volume of particular product items. The sensitivity of the coefficient GDP for heterogeneous products is higher than that of homogenous ones implying that richer countries would absorb more products which are more heterogeneous.

---

\(^8\) Over identification statistic is robust to arbitrary heteroskedasticity and within-group correlation. So Hausman and Taylor estimator is forced by robust-cluster standard errors.
### TABLE 2. GRAVITY MODEL FOR EXPORT FLOWS
(Dependent variable: export value in USD, robust S.E)

<table>
<thead>
<tr>
<th>Independent variables</th>
<th>Pooled OLS</th>
<th>FE</th>
<th>GLS random</th>
<th>HT-IV/GLS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>GDPjit</td>
<td>0.998***</td>
<td>1.892**</td>
<td>1.031***</td>
<td>1.346***</td>
</tr>
<tr>
<td></td>
<td>(18.25)</td>
<td>(2.40)</td>
<td>(19.91)</td>
<td>(9.85)</td>
</tr>
<tr>
<td>GDP Per Laborjt</td>
<td>-0.254</td>
<td>-1.245</td>
<td>-0.330**</td>
<td>-0.635***</td>
</tr>
<tr>
<td></td>
<td>(-1.56)</td>
<td>(-1.57)</td>
<td>(-2.41)</td>
<td>(-3.52)</td>
</tr>
<tr>
<td>DGDPjit</td>
<td>0.206**</td>
<td>0.149*</td>
<td>0.183***</td>
<td>0.166***</td>
</tr>
<tr>
<td></td>
<td>(2.06)</td>
<td>(1.71)</td>
<td>(2.61)</td>
<td>(2.84)</td>
</tr>
<tr>
<td>CPIjt</td>
<td>0.00124</td>
<td>0.00141***</td>
<td>0.00159***</td>
<td>0.00153*</td>
</tr>
<tr>
<td></td>
<td>(0.79)</td>
<td>(3.37)</td>
<td>(4.19)</td>
<td>(1.92)</td>
</tr>
<tr>
<td>lnDistanceij</td>
<td>-0.460**</td>
<td>-</td>
<td>-0.401**</td>
<td>-0.336</td>
</tr>
<tr>
<td></td>
<td>(-2.33)</td>
<td></td>
<td>(-2.08)</td>
<td>(-0.93)</td>
</tr>
<tr>
<td>ASEAN</td>
<td>1.834***</td>
<td>-</td>
<td>1.902***</td>
<td>1.711*</td>
</tr>
<tr>
<td></td>
<td>(3.13)</td>
<td></td>
<td>(2.85)</td>
<td>(1.72)</td>
</tr>
<tr>
<td>Landlockedij</td>
<td>-0.909***</td>
<td>-</td>
<td>-1.067***</td>
<td>-0.960**</td>
</tr>
<tr>
<td></td>
<td>(-2.69)</td>
<td></td>
<td>(-3.22)</td>
<td>(-2.08)</td>
</tr>
<tr>
<td>Constant</td>
<td>-11.99***</td>
<td>-28.29**</td>
<td>-12.49***</td>
<td>-17.70***</td>
</tr>
<tr>
<td></td>
<td>(-5.07)</td>
<td>(-2.99)</td>
<td>(-3.50)</td>
<td>(-3.90)</td>
</tr>
</tbody>
</table>

| Year dummies          | yes        | yes     | yes        | yes       |
| Observations          | 1794       | 1794    | 1794       | 1794      |
| No of groups          | -          | 164     | 164        | 164       |
| R-squared             | 0.72       | 0.50    | 0.72       | -         |
| LM (p-value)⁹         | 23.97***   | -       | -          | -         |
| Breusch-Pagan / Cook-Weihsberg¹⁰ | 52.35*** | -       | -          | -         |
| Wald test (p-value)¹¹ | 4.6e+09*** | -       | 3599.38*** | -         |
| Breusch-Pagan LM¹²    | -          | 0.87    | -          | -         |
| Hausman test          | -          | -       | -          | -         |
| Sargan and Hansen test| -          | χ² = 62.91*** | -          | -         |

**Notes:** * significant at 10%; ** significant at 5%; *** significant at 1%.
Absolute of t statistics are in parentheses.

Overall, GDP per labor force effect is significantly negative across four types of regression, showing that an increase in labor productivity of the importing country causes the Vietnam’s exports to decrease. The coefficients are not significant in the pooled OLS and fixed effect regressions. All else constant, HT/IV noted that for each percentage point increase in importer’ labor productivity, Vietnam’s export volume drops by 0.63%. Sanberg et al (2002) discussed that the GDP per capita by importing country captures absorptive capacity and sophistication of demand for imported goods since it is a measure of the capital labor ratio of the importing country. However, since not all citizens can enter the country’s labor market, income per potential worker is more accurate indicator to compare two countries having the same level of income but different productivity. In fact, if this term

---

⁹ Wooldridge test for autocorrelation in panel data with H0: no first-order autocorrelation.
¹⁰ Breusch-Pagan / Cook-Weihsberg test for heteroskedasticity with H0: Constant variance.
¹¹ Modified Wald test for groupwise heteroskedasticity in fixed effect regression model with H0: sigma(i)^2 = sigma^2 for all i.
¹² Breusch and Pagan Lagrangian multiplier test for random effects with var(u) = 0 by lnExportValue[cnum,t] = Xb + u[cnum] + ε[cnum,t].
could account for unemployed workers, the result would be more plausible. Bastos and Silva (2010) found that income per worker of an importing country may affect the quality of import products and the competitiveness of firms in importing countries. As such, an increase in the importing countries’ labor productivity induces a growth in export quantity and unit value on the part of exporting countries. Locally-based firms in importing countries may be a challenge to competing outside firms. This finding clearly shows the importance of export pricing in dealing with the difference in purchasing power or low cost labor in local production sectors. For country level export data, it is not possible to evaluate the effect of the country’s labor productivity on the export unit price. In the case of Vietnam, if country partners are more productive, they tend to import less.

In regard to the difference in income per capita of Vietnam and its trading partners, the coefficient is found to be positive and significant across the regressions indicating that the higher the difference in income per capita by Vietnam and its trading partners, the higher the export volume. This finding is consistent with that of Rault et al (2007) while this effect in the study of Serlanga and Shin (2004) is somewhat ambiguous. In contrast, Nguyen (2009) found a negative effect for the GDP difference between Vietnam and its 23 trading partners comprising 6 ASEAN country partners and 17 non-ASEAN countries over a 16 year period from 1990 to 2005 which fits the Linder hypothesis. In this study, the finding confirms Heckscher and Ohlin (H-O) theorem that countries will export products that make use of their abundant and cheap factors of production. Furthermore, Vietnam’ trading partners include most of the developed countries with a high level of income whereas the exporting country is low in the income and labor cost. Meanwhile, as found by Leontief (1954), a capital intensive industry such as the US, tends to export labor-intensive products which appears to run counter to the philosophy of H-O model since this phenomenon can be explained by the relatively high capital investment per worker in the United States.

According to Linder (1961) the international trade pattern can be conjectured on the basis of the demand structure of countries. That is to say, the more similar the demand structure of countries, the more they will trade with one another. Further, international trade will still occur between two countries having identical preference and factor endowments. The findings by Baltagi et al (2003) advocate several new ideas on international trade theory which covers Linder’s hypothesis of a preferred model with full interaction effects and the classical H-O model with incomplete interactions model. The finding is also in line with that in Marie’s paper (2009) investigating whether the GDP per capita difference between 12 EU countries and OECD countries depresses export volumes.

Another basic factor inherent to the gravity is geographical distance between two trading countries depressing bilateral trade. Empirical testing for the effect of distance is popular in a wide range of gravity papers such as Longo and Sekkat (2004), Sohn (2005), Elliot (2007). Effect of distance within export trade by Vietnam and its over 150 countries is somewhat ambiguous among the regressions. For the OLS and random effects model, it is negative and significant while for the Hausman and Taylor regression, it is still negative but insignificant. That is, when controlling for possible relationship between covariates and unobservable effects,
the distance sign appears to be biased in the pooled OLS model and the all or nothing choice of correlation between the regressors and unobservable effects suggested by Hausman and Taylor (1981). The problem is that distance does not change over time while transport costs such as fuel costs, or any fee applied to certain export goods is variant. Consequently, Brun et al (2005) have treated distance as a function of oil price and infrastructure variables. Thereby, they demonstrated that the effect of distance does not diminish over time even though this effect changes from negative to positive. And using firm level data, Bastos and Silva (2010) proved that distance is significantly positive for export volumes and export unit prices suggesting that more competitive firms are better able to serve more difficult markets.

There is several reason why geographical distance is not always a good proxy for trade obstacle. Firstly, this covariate does not change over time and it can not be identified in the fixed effect model. Secondly, this covariate results in dubious estimates where two states have reasonably large economies and consequently a high level of mutual trade even though. Thirdly, the level of mutual trade between two contiguous countries depends on their degree of trade with the rest of the world. Fourthly, foreign trade rests on how much domestic trade takes place within exporting or importing countries. If domestic trade accounts for most of a country’s production capacity, its export volume would fall correspondingly. The empirical result shows an insignificantly negative sign of distance corresponding to Hausman-Taylor approach. It is interesting to note that Nguyen (2010) found distance between country pairs to be a hindrance to trade.

13 If the regression is manipulated by OLS on the basis of separate years (1997-2009), effect of distance is negative and significant. The coefficients are not fixed and fluctuates in a range [-0.75: -0.48]
Anderson and Van Wincoop (2003) have proposed a multilateral resistance term (MRT)\(^{14}\) as another term for trade hindrance. Trade cost in the context of the MRT considers a country's trade cost with the rest of the world, which includes the intra-national and international trade and the CPI. The general equilibrium structure with many countries trading a variety of goods differentiated by country of origin is expressed as follows:

\[ X_{ij} = \frac{Y_i Y_j}{Y_{w}} \left( \frac{t_{ij}}{p_i p_j} \right)^{1-\sigma} \]

Where \(Y_i, Y_j\) are total income of country \(i\) and country \(j\), \(Y_w\) is the GDP of the whole world, \(t_{ij}\) is the trade cost incurred by the exporter and/or importer and is assumed to be symmetric (\(t_{ij} = t_{ji}\)). If the trade cost inside the exporting country is assumed to be 1 (\(t_{ii} = 1\)), then the trade cost that country \(i\) passes on to country \(j\) is \(t_{ij} - 1\), and transport costs between the two countries are \(t_{ij} = t_{ji} P_j\), where \(P_j\) designates the exogenous multilateral resistance of the exporting and/or the importing country. As explained in the previous section, because of the unavailability of the intra-national trade data, the study uses the CPI data instead. Theoretically for the MRT, if the multilateral resistance \(P_i\) is high, this means high trade barriers for country \(i\) and demand for country \(i\)'s goods is lower. In contrast, higher trade barriers for country \(j\) characterized by high multilateral resistance \(P_j\), leads to higher demand for country \(i\)'s goods. The result shows that multilateral resistance of the importing country \((P_j)\) fosters demand for Vietnamese products and is significant regarding the fixed, random and HT-IV models. Since, the CPI

\(^{14}\) Starting from the idea of utility function, Anderson and Van Wincoop assumed that if \(c_{ij}\) is consumption by region \(j\) consumers of goods from region \(i\), consumers in region \(j\) maximize

\[ \sum_i \beta^{(1-\sigma)/\sigma} c_{ij}(\sigma - 1)/\sigma \] subject to the budget constraint \(\sum_i p_{ij} c_{ij} = y_j\) Where \(\sigma\) is the elasticity of substitution between all goods, \(\beta\) is a positive distribution parameter, \(y_j\) is a nominal income. If \(p_i\) denotes the exporter’s supply price and \(t_{ij}\) is the trade cost factor between \(i\) and \(j\), then \(p_{ij} = p_i t_{ij}\). For each goods shipped from \(i\) to \(j\), and if the exporter incurs export cost equal to \(t_{ij} - 1\) of country \(i\)'s goods, the sum of value of production of the origin \(p_i c_{ij}\) and the trade cost \((t_{ij} - 1) p_i c_{ij}\) that the exporter passes on to the importer. And total income of country \(i\) is therefore \(y_j = \sum_{i} X_{ij}\). As such, the nominal demand for country \(i\) goods by country \(j\), consumers satisfy by

\[ X_{ij} = \left( \frac{\beta_i p_j t_{ij}}{P_j} \right)^{1-\sigma} y_j \]

Where \(P_j\) is the consumer price index of \(j\)

given by

\[ P_j = \left[ \sum_i \left( \beta_i p_i t_{ij} \right)^{1-\sigma} \right]^{1/(1-\sigma)} \]
captures some negative values over the period, it is not regressed in the form of a natural logarithm. According to the regression results, when the importing country has a 1% increase in CPI, Vietnam’s exports will increase by around 14%. The positive sign of $p_j$ is in line with the findings of Rahman (2003) and Rudolph (2010) fitting the theoretical view that the rise in an importing country’s trade cost shifts demand toward the international market.

The regression results show that in the intra trade context of the ASEAN, exports are significantly positive, confirming the important role of the ASEAN free trade area in stimulating Vietnamese exports. The ASEAN coefficient indicates that Vietnam’s membership of the ASEAN group has contributed to a 568% increase in export volume between 1997 and 2009. The effect of the ASEAN dummy is not in line with Nguyen, X (2010) and Nguyen, K (2009) where the negative sign of the ASEAN dummy raise a suspicion of its erosion. This discrepancy could be explained by the type of model used. Nguyen, K (2010) examined the ASEAN effect for trade between Vietnam and the other ASEAN countries and Nguyen, X (2009) observed Vietnam’s trading not only with ASEAN countries but also with non-ASEAN countries. In fact, Vietnam has become more integrated into the global market and thus less involved in intra-regional trade. This negative sign result may also be partly due to the random sample taken for 15 largest trading partners of Vietnam. For my export model, the ASEAN effect remains important. Geographical distance is not important for neighboring ASEAN markets, however, what is important for Vietnam is the effect of trade liberalization through tariff and non-tariff removal.

Lastly, another trade barrier term is also examined is the landlocked dummy. The negative sign of the landlocked dummy indicates that Vietnam finds it a greater challenge to export products to these countries. The reason may be that over 90% of export contracts are made in FOB15 price which are less favorable to landlocked countries. The result in table 2 suggests that the median value of Vietnamese exports to landlocked countries is about 95%16 lower than non-landlocked countries.

3.3. ENDOGENEITY TEST

Using the Hausman and Taylor model, the author has attempted to test whether the income (GDP) of the importing countries correlates with the unobservable effects. If this is the case, the all or nothing correlation between independent variables and unobservable effects in the fixed and random effects world would accordingly be biased. In the first stage, we applied the Hausman and Taylor approach, where the income (GDP) of the importing countries is assumed to be endogenous. In the second stage, I used the extended IV/SLS approach to determine whether this variable is endogenous and the instrumental variables can be valid.

---

15 FOB stands for Free on Board and this term is applied to sea freight export contract.
16 Technically, the figure is obtained by taking the antilog (to base e) of the estimated dummy coefficient then subtracting 1 (*100). For further detail, see the econometric book by Gujarati (2004, 321).
As discussed above, the validity of the excluded and instrumented variables must be confirmed for the application of the Hausman and Taylor approach. In some empirical papers such as Rault, et al (2008), the authors did not clearly indicate which variable should be treated as endogenous. Serlanga and Shin (2004) did not give a persuasive test to determine whether the GDP variable was endogenous merely assuming that it was. Although Egger (2004) suggested a list of highly relevant instruments including exporter and importer viability of contracts, exporter and importer rule of law, exporter and importer taxes as a percentage of trade, the use of canonical correlations criticized by Shea (1996), does not identify the poor conditioning of the independent variables. In fact, he shows that the variables belonging to two sets of variables are not necessarily correlated (multi co-linearity), hence the canonical correlation in such cases would give ambiguous results for the linear combinations of the x’s and the y’s which have max correlation with each other. As such, Egger (2004) using canonical correlation did not resolve the multicolinearity problems of the instrumented variables in his augmented gravity model. Although Egger has applied the over-identification test to the validity of the instruments, he neglected testing the validity of endogenous variables for providing clear confirmation of the possible relationship between endogenous variables and their excluded instruments.

Intuitively, I believe that the importing countries’ labor productivity can influence the income of those countries thus being a possible relevant excluded instrument. The H-T regression is shown in column 4 table 2, and gives all coefficients lying between those in the fixed and random regressions. Table 3 reports the result of endogeneity testing. Although Shea correlation and partial R-squared show a high correlation for the relationship between the endogenous and instrumental variables, using the Hansen J test significantly rejects the null hypothesis that the instrument is valid. However, we were not able to reject the null hypothesis that the endogenous variable GDP can actually be treated as exogenous in our case, therefore accepting that the importing countries’ income is exogenous rather than endogenous as we had previously assumed.

This is being so, I need to determine whether fixed or random effects approach is more suitable for my case. In a fixed effects model, the condition of the regressors being orthogonal to the idiosyncratic term must be satisfied while in the random effects model, the additional condition of orthogonality between the regressors and the group-specific errors is a requirement. Table 2 presents the test results for the fixed or random model, using two types of specification test including Hausman (1978) and the xtoverid test described by Arellano (1993) and Wooldridge (2002). According to the Hausman test, it is not possible to reject the null hypothesis that the difference in coefficients is not systematic therefore the random effects model is preferred. However, the result of this model is inconsistent with the xtoverid test whereas the fixed effects model is significant at 1%. The xtoverid test outperforms the Hausman test in that it is robust to arbitrary heteroskedasticity and within-group correlation when the cluster option is used. As such, in this case the xtoverid test is more effective than the Hausman identification test.
TABLE 3. ENDOGENEITY TEST

Warning - singleton groups detected. 2 observation(s) not used.
Warning - collinearities detected
Vars dropped: lnDistance ASEAN

FIXED EFFECTS ESTIMATION

Number of groups = 162 Obs per group: min = 2
          avg = 11.1
          max = 13

Summary results for first-stage regressions

<table>
<thead>
<tr>
<th>Variable</th>
<th>Shea Partial R2</th>
<th>Partial R2</th>
<th>F( 1, 161)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>lnGDP_PPP</td>
<td>0.6933</td>
<td>0.6933</td>
<td>324.86</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

NB: first-stage F-stat cluster-robust

Underidentification tests
Ho: matrix of reduced form coefficients has rank= K1-1 (underidentified)
Ha: matrix has rank= K1 (identified)
Kleibergen-Paap rk LM statistic = 32.15 P-val = 0.0000
Kleibergen-Paap rk Wald statistic = 329.45 P-val = 0.0000

Weak identification test
Ho: equation is weakly identified
Kleibergen-Paap Wald rk F statistic = 324.86
See main output for Cragg-Donald weak id test critical values

Weak-instrument-robust inference
Tests of joint significance of endogenous regressors B1 in main equation
Ho: B1=0 and overidentifying restrictions are valid
Anderson-Rubin Wald test F(1,161) = 0.31 P-val = 0.5773
Anderson-Rubin Wald test Chi-sq(1) = 0.32 P-val = 0.5739
Stock-Wright LM S statistic Chi-sq(1) = 0.32 P-val = 0.5699

NB: Underidentification, weak identification and weak-identification-robust test statistics cluster-robust

Number of clusters N_clust = 162
Number of observations N = 1792
Number of regressors K = 15
Number of instruments L = 15
Number of excluded instruments L1 = 1

IV (2SLS) estimation

Total (centered) SS = 2791.722873 Centered R2 = 0.4960
Total (uncentered) SS = 2791.722873 Uncentered R2 = 0.4960
Residual SS = 1407.064264 Root MSE = 0.9291

Estimates efficient for homoskedasticity only
Statistics robust to heteroskedasticity and clustering on cnum

Number of clusters (cnum) = 162
Number of obs = 1792
F( 15, 161) = 58.95 Prob > F = 0.0000
Empirical studies using the gravity model have successfully explained the international trade pattern. General trade between a country pair is shown to be proportional to their income level and inversely proportional to the distance between them.

This study specifically employs Vietnam's export trade with all countries across continents over the period 1997-2009 with the purpose of examining the application of the gravity model on the country trade. Among four types of regressions, generalized least square with instrumental variables is found to be the most relevant. Significantly, the 15 largest importing countries of Vietnamese goods in this period dominate the export market, accounting for some 80% of the annual

### Table 1: Robust lnExportValue Regressions

| lnExportValue | Coef.   | Std. Err. | z     | P>|z|     | [95% Conf. Interval] |
|--------------|---------|-----------|------|--------|---------------------|
| lnGDP_PPP    | 0.3093001 | 0.5501963 | 0.56 | 0.574  | -0.7690647          | 1.387665 |
| lnRLF       | 0.1769208 | 0.0907648 | 1.95 | 0.051  | -0.0009749          | 0.3548165 |
| cpi          | 0.0015278 | 0.0004108 | 3.72 | 0.000  | -0.0007226          | 0.0023329 |
| year2        | 1.137163  | 0.2482436 | 4.58 | 0.000  | 0.650615            | 1.623712 |
| year3        | 0.633033  | 0.1443147 | 4.39 | 0.000  | 0.3501814           | 0.9158845 |
| year4        | 0.7305751 | 0.1314503 | 5.94 | 0.000  | 0.5227372           | 1.038013 |
| year5        | 0.8344434 | 0.164588  | 5.07 | 0.000  | 0.5118569           | 1.15703 |
| year6        | 0.745543  | 0.1890311 | 3.94 | 0.000  | 0.374947            | 1.116139 |
| year7        | 0.9544061 | 0.2105705 | 4.53 | 0.000  | 0.541695            | 1.367117 |
| year8        | 1.385668  | 0.2579409 | 5.37 | 0.000  | 0.880113            | 1.891223 |
| year9        | 1.627274  | 0.3023218 | 5.38 | 0.000  | 1.034735            | 2.219814 |
| year10       | 1.9900124 | 0.337095  | 5.90 | 0.000  | 1.32943             | 2.650818 |
| year11       | 2.183265  | 0.3844374 | 5.68 | 0.000  | 1.429781            | 2.936748 |
| year12       | 2.703081  | 0.426947  | 6.39 | 0.000  | 1.874615            | 3.531548 |
| year13       | 2.671522  | 0.4305472 | 6.20 | 0.000  | 1.827664            | 3.515379 |

**Underidentification test (Kleibergen-Paap rk LM statistic):** 32.147

**Chi-sq(1) P-val =** 0.0000

**Weak identification test (Kleibergen-Paap rk Wald F statistic):** 324.860

**Stock-Yogo weak ID test critical values:**
- 10% maximal IV size: 16.38
- 15% maximal IV size: 8.96
- 20% maximal IV size: 6.66
- 25% maximal IV size: 5.53

**Source:** Stock-Yogo (2005). Reproduced by permission.

**NB:** Critical values are for Cragg-Donald F statistic and i.i.d. errors.

Hansen J statistic (overidentification test of all instruments): 0.000

(equation exactly identified)

Collinearity/identification problems in eqn. excl. suspect orthog. conditions:
- C statistic not calculated for -orthog- option
- endog- option:

Endogeneity test of endogenous regressors:

<table>
<thead>
<tr>
<th>Regressors tested</th>
<th>Chi-sq(1) P-val =</th>
</tr>
</thead>
<tbody>
<tr>
<td>lnGDP_PPP</td>
<td>1.704</td>
</tr>
</tbody>
</table>

Regressors tested: lnGDP_PPP

**Instrumented:** lnGDP_PPP

**Included instruments:** lnRLF cpi year2 year3 year4 year5 year6 year7 year8 year9 year10 year11 year12 year13

**Excluded instruments:** lnGDP_PPPPerLB

Dropped collinear: lnDistance ASEAN

**Notes:** Heteroskedasticity and correlation are robust to the standard errors with country-time clusters.

**CONCLUSION**

Empirical studies using the gravity model have successfully explained the international trade pattern. General trade between a country pair is shown to be proportional to their income level and inversely proportional to the distance between them.

This study specifically employs Vietnam's export trade with all countries across continents over the period 1997-2009 with the purpose of examining the application of the gravity model on the country trade. Among four types of regressions, generalized least square with instrumental variables is found to be the most relevant. Significantly, the 15 largest importing countries of Vietnamese goods in this period dominate the export market, accounting for some 80% of the annual
volume of exports. Nevertheless, high income countries remain important to Vietnamese export trade. However, the augmented gravity model highlights the fact that high-labor productivity countries are not significant to Vietnam’s export trade. The difference in country’s per capita income strengthens the H-O theory, suggesting that Vietnam can continue to use its low labor cost advantage.

Unlike in previous literature on Vietnamese trade, geographical distance is shown not to be a good proxy for trade cost although the sign is negative. To replace it, the author suggests using the CPI of importing countries as a proxy for the multilateral resistance factor since data on intra country production is inadequate. The positive effect of CPI acts as a stimulus on imports from Vietnam. The positive ASEAN sign emphasizes the role of regional economic integration where tax cuts help to encourage imports by the ASEAN countries from Vietnam.

Although the results fit well with the export data, export pattern of Vietnam could not confirm the role of firms or export products on international markets. If firm or product level data becomes publicly available, it is expected that export volume by firm or product categories would be integrated into the gravity model. Bastos and Silva (2010) have successfully calibrated gravity model at firms level and investigated the influence of firm productivity on trade of Portugal with its distant countries. Nguyen and Ancard (2009) compared the export value data of various categories of gravity model factors. In future, I hope to study the gravity model for Vietnam at firm level with the aim of deeply investigating whether exports by firms could give more interesting results.

REFERENCES


